LIFE CYCLE, NOVEL PROTEINS AND STRUCTURAL ORGANIZATION OF
THE SYNECHOCOCCUS CYANOPHAGE SYN5

A dissertation presented by

Desislava Andreeva Raytcheva

to

The Department of Biology

In partial fulfillment of the requirements for the degree of
Doctor of Philosophy

in the field of

Biology

Northeastern University
Boston, Massachusetts
April 2012
LIFE CYCLE, NOVEL PROTEINS AND STRUCTURAL ORGANIZATION OF
THE SYNECHOCOCCUS CYANOPHAGE SYN5

by

Desislava Andreeva Raytcheva

ABSTRACT OF DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Biology in the Graduate School of Northeastern University,
April, 2012
ABSTRACT

Syn5 is a marine cyanophage, which is propagated under laboratory conditions on the marine cyanobacterial strain *Synechococcus* sp. WH8109. Its genome has high homology with enteric bacteriophage T7 but it also contains several genes encoding protein sequences with no matches to phage proteins in the NCBI database. Cryo-EM images of Syn5 revealed a novel structure, a so-called horn, on the vertex opposite the tail.

Synchronized, exponentially growing host cultures were cultivated and the one-step growth curve of Syn5 was performed. Syn5 exhibited a short life cycle with an eclipse period of about 43 min and a latent phase of about 60 min with a burst size of 20–30 particles per cell at 28°C. SDS-PAGE and Western analysis of cell lysates at different times after infection showed the synthesis of the major virion proteins and their increase as the infection progressed.

Most of the genes of the novel Syn5 proteins as well as the coat, scaffolding and portal proteins were cloned, expressed in *E. coli* and purified, with the exception of the coat. The coat protein was insoluble when expressed on its own and its co-expression with the putative Syn5 scaffolding protein did not improve its solubility. The recombinant portal, coat and scaffolding proteins exhibit properties similar to those of homologs in other bacteriophages, especially to their homologs in T7. The characteristics of the recombinant novel proteins gp53 and gp54 suggest that they may be capable of polymerization.
Immunogold labeling of infectious Syn5 particles with polyclonal antibodies raised against novel proteins strongly suggests that gp58 is positioned on the capsid head, possibly building the decoration knobs, while gp53 and gp54 contribute to the horn structure.

Treatment of Syn5 infectious particles with polyclonal antibodies raised against novel proteins gp58, gp53, and gp54 did not affect the ability of Syn5 to infect host cells while anti-Syn5 antibodies raised against whole virus particles inactivated the phage.

Western blots of Syn5-infected cell lysates showed that the putative scaffolding protein (gp38) is expressed during infection. Particles lacking DNA, but containing the coat and scaffolding proteins, were purified from Syn5-infected cells, about 50 min after infection, using CsCl centrifugation followed by sucrose gradients. Electron microscopic images of the purified particles showed shells lacking condensed DNA, but filled with protein density, presumably scaffolding protein.

The findings described here strongly suggest that cyanophages such as Syn5 form infectious virions through initial assembly of scaffolding-containing procapsids, similar to the assembly pathways for the enteric dsDNA bacteriophages. Since cyanobacteria predate the enteric bacteria, this procapsid-mediated assembly pathway was probably inherited by enteric phages from more ancient bacteriophages including cyanophages.
This thesis is dedicated to my family
ACKNOWLEDGEMENTS

There are many people to whom I am deeply indebted in connection with this dissertation.

Prof. Jacqueline Piret, thank you for opening your lab and heart to me. I considered finding the right advisor the most important decision in my first year as a graduate student. I will be forever grateful for giving me a chance. You not only taught me the very important skills of how to design my experiments well and how to write papers but you were also considerate and understanding. Whether I was going through personal struggles or through hard times with my teaching obligations, you always had my best interest in mind and your focus was on what mattered the most. Finally, thank you for all the fun conversations about life in general.

Professor Jonathan King, most people are lucky if they find one good mentor during their PhD studies and I was blessed with two. Whether it was teaching me how not to fall in love with my hypothesis or how to articulate my data better your advice was always invaluable. Thank you for reminding us that as scientists we should also be responsible citizens of the world. You are also one of the few people in my life who encourage me to go back to Bulgaria. You are absolutely right that no matter how much I achieve here in the United States, it will never give me the satisfaction I will get from achieving even just a small portion of it back home.

I would also like to thank my committee members, Prof. Veronica Godoy-Carter, Prof. Kostia Bergman and Prof. Slava Epstein for listening, reading and giving
suggestions and ideas for the improvement of this work. Your help was much appreciated.

Cameron Haase-Pettingell, thank you for being always so tremendously helpful whether it was about ordering supplies, taking care of every minor detail around the lab for all of us or offering a helping hand on the everyday lab chores. Your assistance with the radioactive experiments was very important to me and I would be forever grateful that you offered that very special help.

Dr. Peter Weigele, your passion in science is contagious. I was lucky to have someone who knew so much on the topic and was so generous in sharing that knowledge. And you definitely helped cure my fear of sophisticated machines by teaching me how to use the ultracentrifuge and the electron microscope.

I am immensely indebted: to Althea Hill who washed countless bubblers; to Jeannie Chew who purified large amounts of Syn5 which my experiments devoured quickly; to Erika Erickson who started the use of artificial sea water in the King lab which tremendously facilitated the work on this project.

I would also like to thank all the past and current members of the King lab who I was very fortunate to meet. People and scientists like Dr. Yongting Wang, Dr. Ryan Simkovksy, Dr. Ishara Mills, Dr. Jiejin Chen, Dr. Ligia Acosta-Sampson, Dr. Kate Moreau, Dr. Takumi Takata, Dr. Fanrong Kong, Dan Goulet, Nathaniel Schafheimer, Oksana Sergeeva, Lisa Guisbond—thank you all for listening, providing technical assistance and ideas, and for the words of encouragements when I needed them the most. Special thanks to Dr. Kelly Knee for keeping an eye on that good old JEOL machine.
Many thanks to the Baker lab for allowing me to use their Typhoon scanner and to the Sinskey lab for letting me take over their ultracentrifuge when necessary.

I am also very grateful to have worked with Prof. Wah Chiu, Dr. Wei Dai and Dr. Preeti Gipson from Baylor College of Medicine. It was always inspirational to know that we can rely on the powerful cryo-EM techniques for answering some of the questions about our research. And growing WH8109 in a lab without any microbiological set-up turned out to be one of my favorite challenges as a PhD student. Special thanks to Prof. Bill Fowle for sharing his experience about immogold labeling.

Since most of the experiments were physically performed at MIT, there was important and complicated paperwork that had to be done. Cindy Woolley and Francine Chaput, thank you so much for working so hard to get it done.

Many special thanks to my NU family. Frauke Argyros and Aaron Roth did a great job supervising me as a TA in their lab sections. Thanks to Adrian Gilbert for allowing me to teach Microbiology lab as many semesters as I wanted. Thanks to Janeen Greene, Laura McGann and Petya Koeva from the Bio office for all the help with administrative work. Many thanks to all my TA colleagues, especially to Dr. Srikanth Subramanian, Ryan Benson and Alyssa Theodore. Being part of a team is always much more fun than working alone.

I would also like to thank the hundreds of students whom I met through the years as a TA. The experience as a teacher was one of the best in my days as a graduate student. It made every hour at the bench worth it and gave me the strength to continue through the everyday challenges of research.
Thanks to my old and new friends across the globe for their invaluable support. Thank you Maria, Iva, Gena and family, Evi and family, Mariusz, Stoichko and Vessela, Kameliya and Kalin for those long phone conversations, short but so sweet trips, all the holidays we celebrated together and all the fun we had in each other’s company. Special thanks to everyone who organized Bulgarian elections in Boston with me, definitely an amazing experience for all of us.

I dedicate this thesis to my family because without their support I would have never reached this point in my life. Thanks so much to my parents, grandparents, my aunt, for encouraging my curiosity for as long as I can remember and giving me the full freedom to develop my potential. Thank you for teaching me the values of hard work and to never give up easily when things get tough. I know you miss us and we miss you too but hopefully the experience and knowledge we gained here will be for the greater good of everyone. Many thanks to my brothers. Special thanks to my in-laws for their support.

And last but not least, I am especially grateful to my husband. The past years have been quite a journey, one that I am extremely grateful to have shared with you. Thank you for all your help, understanding, unconditional love and support and for always reminding me about the other point of view. Hopefully, our son (who helped in writing this thesis in his own way too) will be at least a little bit like you. And most importantly—thank you for believing in me every time I did not.

This work was partially supported by:

NIH grants GM17980 and AI075208 awarded to Prof. Jonathan King, MIT;
Department of Biology, Northeastern University;
Dissertation Completion Fellowship from the Provost office at Northeastern University
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................................. 2

ACKNOWLEDGEMENTS ........................................................................................................... 6

TABLE OF CONTENTS ................................................................................................................ 10

TABLE OF FIGURES .................................................................................................................... 13

TABLE OF TABLES ..................................................................................................................... 15

LIST OF ABBREVIATIONS ......................................................................................................... 16

CHAPTER I: INTRODUCTION .................................................................................................. 19

1.1 Cyanobacteria ...................................................................................................................... 20

1.2 The genus Synechococcus .................................................................................................. 21

1.3 Synechococcus sp. WH8109 ................................................................................................. 23

1.4 Marine viruses ...................................................................................................................... 23

1.5 Bacteriophages ..................................................................................................................... 27

1.6 Cyanophages of marine Synechococcus and Prochlorococcus ........................................... 30

1.7 Bacteriophage assembly ...................................................................................................... 33

1.8 Syn5 .................................................................................................................................... 36

1.9 Thesis objectives .................................................................................................................. 43

CHAPTER II: SYN5 INFECTION AND GROWTH .................................................................. 46

2.1 Introduction .......................................................................................................................... 47

2.2 Materials and Methods ........................................................................................................ 48

2.2.1 Growth of Synechococcus sp. WH8109 ........................................................................... 48

2.2.2 Phage growth ................................................................................................................... 50

2.2.3 Phage titering .................................................................................................................. 51

2.2.4 One-step growth curve of Syn5 in the Percival chambers ............................................. 54

2.2.5 Syn5 protein analysis during infection ........................................................................... 55

2.3 Results ................................................................................................................................ 56

2.3.1 Growth of Synechococcus sp. WH8109 ........................................................................... 56

2.3.2 One-step growth curve of Syn5 ...................................................................................... 57

2.3.3 Host and phage proteins during infection ....................................................................... 60

2.4 Discussion ............................................................................................................................ 65

CHAPTER III: STRUCTURAL PROTEINS OF SYN5 ......................................................... 67

3.1 Introduction .......................................................................................................................... 68

3.2 Materials and Methods ......................................................................................................... 70
CHAPTER IV: LOCALIZATION OF SYN5 NOVEL PROTEINS BY IMMUNOGOLD LABELING

4.1 Introduction................................................................................. 98
4.2 Materials and Methods................................................................. 100
  4.2.1 Immunogold labeling of Syn5 with anti-Syn5, anti-gp53, anti-gp54 and anti-gp58 antibodies .................................................. 100
  4.2.2 Syn5 inactivation with anti-Syn5 serum and antibodies to novel Syn5 proteins .................................................................. 101
  4.2.3 Sucrose gradients of the recombinant proteins gp53 and gp54 .................. 102
4.3 Results..................................................................................... 103
  4.3.1 Immunogold localization of Syn5 particles with anti-gp53, anti-gp54, anti-gp58 and anti-Syn5 antibodies ........................ 103
  4.3.2 Syn5 inactivation with anti-Syn5 and anti-novel protein antibodies 114
  4.3.3 Sucrose gradients of gp53 and gp54 ........................................ 117
4.4 Discussion.............................................................................. 121

CHAPTER V: STUDIES OF SYN5 IN VIVO PROCAPSIDS ...................... 124
5.1 Introduction.............................................................................. 125
5.2 Materials and Methods............................................................. 126
  5.2.1 Expression of the Syn5 scaffolding protein during infection ................... 126
  5.2.2 Purification of procapsids ...................................................... 127
  5.2.3 Electron microscopy ................................................................ 128
5.3 Results.................................................................................. 129
  5.3.1 Syn5 assembly involves a scaffolding protein ................................. 129
  5.3.2 Do Syn5 virions assemble through a procapsid intermediate? ........ 131
5.4 Discussion............................................................................. 138

CONCLUDING REMARKS.................................................................... 140
# TABLE OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–1</td>
<td>Dividing cell of <em>Synechococcus</em> sp. WH8102.</td>
<td>25</td>
</tr>
<tr>
<td>1–2</td>
<td>Phylogenetic tree of the <em>Synechococcus</em> genus.</td>
<td>26</td>
</tr>
<tr>
<td>1–3</td>
<td>The effects of viruses on the marine community.</td>
<td>28</td>
</tr>
<tr>
<td>1–4</td>
<td>Assembly of procapsids in dsDNA bacteriophages.</td>
<td>35</td>
</tr>
<tr>
<td>1–5</td>
<td>Cryo-EM images of Syn5.</td>
<td>39</td>
</tr>
<tr>
<td>1–6</td>
<td>Symmetric reconstruction of Syn5.</td>
<td>40</td>
</tr>
<tr>
<td>1–7</td>
<td>A map of the Syn5 genome.</td>
<td>41</td>
</tr>
<tr>
<td>1–8</td>
<td>Syn5 structural proteins.</td>
<td>42</td>
</tr>
<tr>
<td>2–1</td>
<td>Syn5 infecting WH8109.</td>
<td>52</td>
</tr>
<tr>
<td>2–2</td>
<td>Syn5 plaques.</td>
<td>53</td>
</tr>
<tr>
<td>2–3</td>
<td>Images of <em>Synechococcus</em> cells under epifluorescent microscope.</td>
<td>58</td>
</tr>
<tr>
<td>2–4</td>
<td>Growth curve of <em>Synechococcus</em> sp. WH8109.</td>
<td>59</td>
</tr>
<tr>
<td>2–5</td>
<td>One-step growth curve of Syn5.</td>
<td>61</td>
</tr>
<tr>
<td>2–6</td>
<td>Host and phage proteins during Syn5 infection.</td>
<td>64</td>
</tr>
<tr>
<td>3–1</td>
<td>Map of the pET-21 vector cloning region.</td>
<td>72</td>
</tr>
<tr>
<td>3–2</td>
<td>Map of the pET-15b vector cloning region.</td>
<td>74</td>
</tr>
<tr>
<td>3–3</td>
<td>Map of the pETDuet-1 vector cloning regions.</td>
<td>77</td>
</tr>
<tr>
<td>3–4</td>
<td>SDS gels of the overexpressed protein products of gp39 and gp46.</td>
<td>81</td>
</tr>
<tr>
<td>3–5</td>
<td>Purification of the recombinant portal protein (gp37).</td>
<td>82</td>
</tr>
<tr>
<td>3–6</td>
<td>Gel electrophoresis of the Syn5 recombinant scaffolding protein (gp38).</td>
<td>84</td>
</tr>
</tbody>
</table>
Figure 3–7. Purification of the recombinant protein gp53......................................................... 86
Figure 3–8. Purification of the recombinant protein gp54......................................................... 87
Figure 3–9. Purification of the recombinant protein gp58......................................................... 88
Figure 3–10. Gel electrophoresis and Western blot of the three novel Syn5 structural recombinant proteins................................................................. 90
Figure 3–11. Co-expression of the scaffolding (gp38) and coat proteins (gp39). ............ 92
Figure 3–12. Sucrose gradients of cell lysates with co-expressed scaffolding and coat proteins................................................................. 94
Figure 4–1. Electron micrographs of the negative control .................................................. 105
Figure 4–2. Labeling of Syn5 virions with anti-Syn5 antibodies........................................... 107
Figure 4–3. Schematic drawing of Syn5................................................................. 108
Figure 4–4. Labeling of Syn5 with anti-gp53 antibodies..................................................... 110
Figure 4–5. Labeling of Syn5 with anti-gp54 antibodies..................................................... 112
Figure 4–6. Labeling of Syn5 with anti-gp58 antibodies..................................................... 115
Figure 4–7. Inactivation of Syn5 infectivity ........................................................................ 118
Figure 4–8. SDS electrophoresis of sucrose gradients of gp53 and gp54. ...................... 119
Figure 4–9. Electron micrographs of gp53 and gp54. ..................................................... 120
Figure 5–1. The Syn5 scaffolding (gp38) in Syn5 infected cell lysates of WH8109. .... 130
Figure 5–2. Western blot of sucrose gradients of Syn5 procapsid particles............... 133
Figure 5–3. EM images of negatively stained procapsids .................................................... 135
TABLE OF TABLES

Table 3–1. Summary of *in silico* data obtained for gene products 53, 54 and 58............. 69

Table 4–1. Summary of the labeling results from the negative control.............................. 104

Table 4–2. Summary with the positions of anti-gp53 gold labels on the Syn5 surface.. 111

Table 4–3. Summary of the positions of anti-gp54 gold labels on the Syn5 surface. .... 113

Table 4–4. Summary of the positions of anti-gp58 gold labels on the Syn5 surface. .... 116

Table 5–1. Summary of the species of Syn5 procapsid particles. ................................. 136
LIST OF ABBREVIATIONS

AP  Alkaline phosphatase
ASW  Artificial seawater
BLAST  Basic Local Alignment Search Tool
CaCl$_2$$\cdot$2H$_2$O  Calcium chloride
CO$_2$  Carbon dioxide
CoCl$_2$$\cdot$6H$_2$O  Cobalt chloride
Cryo-EM  Cryoelectron microscopy
CsCl  Cesium chloride
DMSO  Dimethyl sulfoxide
DNA  Deoxyribonucleic acid
DNase  Deoxyribonuclease
ds  Double-stranded
ECF  Enhanced chemifluorescence
E. coli  Escherichia coli
EDTA  Ethylenediamine tetraacetic acid
EM  Electron microscope
FITC  Fluorescein isothiocyanate
gp  Gene product
HCl  Hydrochloric acid
<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hli</td>
<td>High-light inducible</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-thiogalactoside</td>
</tr>
<tr>
<td>KCl</td>
<td>Potassium chloride</td>
</tr>
<tr>
<td>K$_2$HPO$_4$·3H$_2$O</td>
<td>Potassium phosphate dibasic</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani broth</td>
</tr>
<tr>
<td>LC</td>
<td>Liquid chromatography</td>
</tr>
<tr>
<td>LMP</td>
<td>Low melting point</td>
</tr>
<tr>
<td>MgCl$_2$·6H$_2$O</td>
<td>Magnesium chloride</td>
</tr>
<tr>
<td>MgSO$_4$</td>
<td>Magnesium sulfate</td>
</tr>
<tr>
<td>MnCl$_2$·4H$_2$O</td>
<td>Manganese chloride</td>
</tr>
<tr>
<td>MOI</td>
<td>Multiplicity of infection</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>MSC</td>
<td>Multiple cloning site</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular weight</td>
</tr>
<tr>
<td>N$_2$</td>
<td>Nitrogen</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>Na$_2$CO$_3$</td>
<td>Sodium carbonate</td>
</tr>
<tr>
<td>NaHCO$_3$</td>
<td>Sodium bicarbonate</td>
</tr>
<tr>
<td>Na$_2$MoO$_4$·2H$_2$O</td>
<td>Sodium molybdate</td>
</tr>
<tr>
<td>NaNO$_3$</td>
<td>Sodium nitrate</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>Ni-NTA</td>
<td>Nickel nitrolothiocetic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>psb</td>
<td>Photosystem II protein</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>pI</td>
<td>Isoelectric point</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium-dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SN</td>
<td>Natural seawater</td>
</tr>
<tr>
<td>ss</td>
<td>Single-stranded</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission electron microscopy</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)-amino methane</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>ZnSO₄·7H₂O</td>
<td>Zinc sulfate</td>
</tr>
</tbody>
</table>
CHAPTER I: INTRODUCTION
1.1 Cyanobacteria

Cyanobacteria are ancient oxygenic photosynthetic organisms originating about 3 billion years ago (Summons et al., 1999). They were noted in the literature as early as 1753 by Linnaeus in his taxonomy (Linnaeus, 1753). Some scientists argue that they can be considered the most successful group of prokaryotes on Earth (Stewart and Falconer, 2008).

They are widely distributed in very different environments all over the globe—marine, freshwater environments, hot springs. Their tolerance to high salt makes them one of the most successful inhabitants of extreme saline environments (Bauld, 1981). The ability to tolerate desiccation facilitates their distribution in many terrestrial environments, including deserts (Palmer and Friedmann, 1990). Cyanobacteria can thrive in slightly higher temperatures than eukaryotic algae and thus occupy more niches (Castenholz and Waterbury, 1989). This prokaryotic group is also characterized by great genetic diversity (Beck et al., 2012). They are quite efficient in their reduction of CO$_2$ even in low inorganic carbon concentrations (Pierce and Omata, 1988) and many species have the ability to fix N$_2$ from the atmosphere if needed. For their photosynthetic activity, all cyanobacteria synthesize chlorophyll a as the major light-harvesting pigment. The presence of accessory pigments like phycocyanin and phycoerythrin (the genus *Synechococcus*) improve the ability of the bacteria to use light of different wavelengths. Some species lack accessory pigments but possess chlorophyll b instead (the genus *Prochlorococcus*).
Marine cyanobacteria of the genera *Prochlorococcus* and *Synechococcus* belong to the dominant picophytoplankton in the open oceans and on average contribute about half of the primary production in this environment (Goericke, 1993; Liu, 1998; Scanlan, 2003). They regulate the flow of carbon through the food webs (Scanlan, 2003).

*Synechococcus* was first described in the literature in 1979 when its orange fluorescence was observed under blue light with an epifluorescent microscope (Johnson and Sieburth, 1979; Waterbury et al., 1979), while *Prochlorococcus* was recognized about a decade later with the assistance of flow cytometry (Chisholm et al., 1988). Of the two genera, *Synechococcus* is more abundant in nutrient-rich zones, has a wider distribution in aquatic environments worldwide (from the poles to the tropics) and can thrive on its own, while *Prochlorococcus* is found only where *Synechococcus* is also present (Scanlan, 2002) but is concentrated deeper in the water column than *Synechococcus* (Partensky et al., 1999a).

### 1.2 The genus *Synechococcus*

*Synechococcus* cells are small cocci (0.5–1 µm diameter) which can inhabit marine and fresh water environments (Fig. 1–1). Motile *Synechococcus* strains utilize a novel non-flagellar type of motility, since no flagella or other motility organelles have been observed in most strains (Brahamsha, 1999). Only one strain, *Synechococcus* sp. WH8113, has been shown to possess a layer of possible motility structures called spicules. This layer surrounds the outer membrane and the spicules extend into the surrounding environment, but they also span through the cell wall and all the way to the inner membrane (Samuel et al., 2001).
The species of the genus *Synechococcus* exhibit circadian rhythms and light control of the cell cycle (Armbrust et al., 1989; Sweeney and Borgese, 1989). The lowest growth temperature for marine strains is about 5°C and most fail to grow at temperatures above 30°C under laboratory conditions (Waterbury et al., 1986). They grow well in light irradiation as low as 10 µmol m⁻² s⁻¹ and tolerate high irradiation at the water surface of 2,000 µmol m⁻² s⁻¹ (Kana and Gilbert, 1987). Their light-harvesting organelles are the phycobilisomes, which contain a core of allophycocyanin and rods with phycoerythrin and phycocyanin (Grossman et al., 1993b; Sidler, 1994). Each of these accessory pigments absorbs light of different maximum wavelengths. The phycobilisomes are attached to the thylakoid membranes (Fig. 1–1), which are bound to chlorophyll a as well. It has been shown that the number of phycobilisomes and their components can vary based on the light intensity (Grossman et al., 1993a; Mullineaux and Emlyn-Jones, 2005).

In terms of nutrition, organisms living in the oceans have to adapt to an environment of dilute and sometime scarce resources. Marine strains of *Synechococcus* can utilize a variety of N sources—ammonium (NH₄⁺), nitrate (NO₃⁻), nitrite (NO₂⁻), urea (Glibert et al., 1986; Glibert and Ray, 1990; Lindell et al., 1998) with most showing highest preference for ammonium. Some strains also possess aminopeptidase activity (Martinez and Azam, 1993) while others assimilate amino acids. It has also been shown that under conditions of nitrogen depletion the cell’s own phycoerythrin can be broken down and used as nitrogen source (Glibert et al., 1986). As for supplying their P needs, marine *Synechococcus* strains have been shown to use readily both organic and inorganic sources (Donald et al., 1997). Some of the important trace metals for growth are Mn, Fe,
Co, Cu, Ni and Zn (Geider and Roche, 1994; Moffett et al., 1997; Tortell et al., 1999; Sunda, 2000; Michel et al., 2003). Levels of Cu\(^{2+}\) higher than 10\(^{-11}\) M have been shown to have toxic effects on growth (Moffett et al., 1997).

### 1.3 Synechococcus sp. WH8109

*Synechococcus* sp. WH8109 (or WH8109) was isolated from the Sargasso Sea by John Waterbury in the late 1970’s. It belongs to marine cluster A, clade II, with its closest relative being *Synechococcus* sp. CC9605 (Fig. 1–2). WH8109 is an obligate autotroph that contains phycoerythrin and is non-motile. Like all strains in its clade it has an increased growth requirement for Na\(^+\), Cl\(^-\), Mg\(^{2+}\), Ca\(^{2+}\). Its genome, which has been sequenced, has a size of 2.1 Mb, close to that of *Synechococcus* sp. WH8102 (2.4 Mb). Strain WH8109 is very susceptible to lytic phage infection and is used as the laboratory host for many cyanophages (Sullivan et al., 2003). It is grown and maintained exclusively in liquid culture since it does not form colonies in pure culture but only in the presence of other marine bacteria. All experiments described in this thesis were performed with WH8109 as the host culture.

### 1.4 Marine viruses

Viruses are the most abundant life forms on Earth with an estimated number of 4 × 10\(^{30}\) in the oceans which is also considered about 10-fold higher than the population of marine prokaryotes (Suttle, 2005). Investigating their impact is an important part of understanding the biogeochemical processes in these environments (Suttle, 2007). Viruses infect both eukaryotic and prokaryotic photosynthetic organisms and these
infections decrease primary productivity (Suttle et al., 1990; Suttle, 1992; Danovaro et al., 2011).

Studies have shown that marine viruses can have substantial effect on the levels of CO$_2$ in the atmosphere and as such are major players in the life cycles in the oceans. Viral lysis also contributes to the circulation of nitrogen and phosphorus in the ecosystems, as well as trace elements including iron, which has been described as a limiting factor for the growth of marine organisms (Suttle, 2005).

Viruses can have a direct effect on the life cycle of their hosts (Fig. 1–3). Examples have been studied in some phytoplankton where the organism goes from diploid and non-motile form to haploid and motile to avoid infection (Rohwer and Thurber, 2009). Bacterial viruses contribute to microbial diversity. This is achieved by their ability to reduce the number of representatives of the dominant species. The higher the density of a given species, the higher the likelihood of a lytic infection and hence reduction in the population numbers of that species. This model has been described by “killing the winner” theory (Thingstad, 2000).

In the marine environment, gene transfer by viruses (horizontal gene transfer) is a key mechanism in generating and transmitting diversity (Waldor and Mekalanos, 1996; Zeidner et al., 2005). There are many examples of genes transferred from the phage to the host and vice versa. One remarkable feature of cyanophages is that some of them carry photosynthetic genes as part of their genomes. The products of these genes are expressed during infection and increase the energy capacity of the cell in order to facilitate phage production (Lindell et al., 2005; Clokie et al., 2006).
Figure 1–1. Dividing cell of *Synechococcus* sp. WH8102. Layered thylakoid membranes can be seen lining the cell envelope and three icosahedral carboxysomes are visible in the cell on the right (marked with black lines). Image courtesy of John Waterbury.
Figure 1–2. Phylogenetic tree of the *Synechococcus* genus. The tree is built in relation to *Escherichia coli* and some marine bacteria. (Scanlan, 2003).
1.5 Bacteriophages

Viruses that infect bacteria were first described in the literature in 1915 by F. W. Twort (Twort, 1915) and in 1917 by d’Herelle (d’Herelle, 1917). Independently, they both reported on agents that cleared bacterial cultures and d’Herelle named them bacteriophages. The development of electron microscopy techniques led to their recognition as viruses and established that they can have quite different structures (Luria et al., 1943; Anderson, 1975). It also confirmed a different mechanism for infection than animal viruses— injection of the genome of the infectious agent through the cell envelope rather than entry of the entire particle into the host cell cytoplasm (Hershey and Chase, 1952). Hershey and Chase used the same experiment to prove that DNA is the carrier of the genetic material. In the next 50 years bacteriophages were used as research tools and as model systems to develop the techniques needed to study animal viruses. Enzymes found in bacteriophages are also widely used as tools in molecular biology (Weiss and Richardson, 1967; Rittie and Perbal, 2008).

Bacteriophages are built of protein shells (heads) which contain, protect and deliver the nucleic acids. Sometimes the shells have appendages on one vertex, which are usually called “tails” and function as the host attachment apparatus. Tailed bacteriophages are divided into three classes—Myoviridae, long contractile tails (e.g. T4); Siphoviridae, long non-contractile tails (e.g. λ); and Podoviridae, short tails (e.g. T7, P22). Bacteriophages are also classified into different groups based on the shape of their protein shells (binary, cubic, helical, pleomorphic) and the nature of their nucleic acid (DNA or RNA, both either single-stranded or double-stranded conformations).
Figure 1–3. The effects of viruses on the marine community. GTAs, generalized transducing agents. (Rohwer and Thurber, 2009).
The bacteriophages studied in greatest detail are the double-stranded DNA (dsDNA) tailed bacteriophages. Tailed bacteriophage infection happens in several steps. The first one is the recognition of phage host receptors on the cell surface. These can be outer membrane proteins (Omps) or even sugar receptors (Rakhuba et al., 2010). However, the host receptors of few bacteriophages have been determined. Once the phage is attached to the cell surface the next step is injection of the nucleic acid. Some phages have long contractile tails that act like a tool to forcefully inject the DNA, others have short tails but contain internal virion proteins that are injected into the cell wall, building a channel to facilitate the passage of the DNA through the thick wall to the cytoplasm. Inside the cell, the infection can be followed by alternative pathways of phage replication, lytic or lysogenic. One scenario is the lysogenic pathway, where the injected phage DNA is integrated into the host DNA and remains inactive in the replicating host cell until stress conditions induce it to progress to the lytic pathway. The second scenario is the lytic cycle which can be entered directly or following the lysogenic state. In the lytic cycle, the phage take over the machinery of the host cell to produce new phage proteins and copies of the phage genome. The phage particles are assembled from these components into mature infectious virions. The time between infection and the assembly of the first new phage particles inside the cells is called the eclipse phase. This time can vary from a few minutes to a few hours. While the host cell is making new phages, two phage encoded proteins, endolysin and holin, are preparing the host membrane for lysis. By still unknown timing mechanisms the cell begins to lyse and the newly formed infectious particles are released into the environment and diffuse towards new hosts. The time frame between infection and the onset of cell lysis by the phage is called the latent
period. It is usually quite standard for each phage-host system but can vary slightly based on conditions such as temperature (Ellis and Delbruck, 1939). The average number of phage particles produced per infected cell (burst size) can vary substantially from a few to a few hundred (Delbruck, 1945). Changing the growth conditions of the phage can also cause variations in the burst size (Hadas et al., 1997).

Bacteriophages have found applications not only as molecular biology tools, but also in phage display (Pande et al., 2010) and more recently in the field of phage therapy. The current crisis involving widespread antibiotic resistant pathogens has brought phages back on the stage as potential therapeutic agents (Burrowes et al., 2011).

1.6 Cyanophages of marine Synechococcus and Prochlorococcus

Some of the most numerous prokaryotic marine viruses measured in the oceans are the cyanophages that infect Synechococcus sp. (Suttle, 2000). This was estimated by lysis assay with exponentially-growing Synechococcus cultures incubated with concentrated samples of natural sea water samples (Suttle and Chan, 1993).

The majority of the isolated and studied cyanophages infect the two dominant marine cyanobacteria genera—Synechococcus and Prochlorococcus. These bacteriophages are usually isolated from liquid dilution cultures since obtaining growth of the two genera on solid medium is often challenging. The reports on the rates of mortality that cyanophages have on their hosts vary substantially. For Synechococcus, the numbers can range between 0.005% to 28% of the host population per day (Waterbury and Valois, 1993; Suttle and Chan, 1994; Garza and Suttle, 1998; Ortmann et al., 2002;
Mann, 2003). Studies have indicated that the major reason for the low numbers is resistance of the hosts to infection rather than due to lysogeny (Stoddard et al., 2007).

The genomes of cyanophages published to date share remarkably similar features with enteric phages both in sequence homology and gene order (Chen and Lu, 2002; Mann et al., 2005; Sullivan et al., 2005; Liu et al., 2007; Pope et al., 2007; Weigele et al., 2007; Liu et al., 2008; Millard et al., 2009; Sullivan et al., 2009; Huang et al., 2012). Their sizes vary from about 40 Kb to about 200 Kb with siphoviruses and myoviruses having larger genomes while podovirus genomes are smaller. Since cyanobacteria evolved long before terrestrial animals and their enteric bacteria, cyanophages are likely to represent lineages that emerged earlier than phages of enteric bacteria.

Some cyanophages have been shown to carry and express genes specifying proteins involved in host cell photosynthesis and other aspects of host cell physiology (Millard et al., 2004; Sullivan et al., 2005; Thompson et al., 2011). That has been shown with phages that infect both *Synechococcus* and *Prochlorococcus*. The *Prochlorococcus* phage MED4 contains *psbA* and *hli* genes, which code for the D1 protein in photosystem II and high-light-inducible proteins, respectively. The two genes are transcribed during infection with an increasing amount of D1 in the course of the infection (Lindell et al., 2007). The *Synechococcus* S-PM2 phage also carries the genes for *psbA* and *psbD*. The S-PM2 *psbA* phage protein product is similar to the *psbA* product of *Synechococcus* sp. WH8102 and the two genes are homologous as well (Mann et al., 2003). This suggests that the phage acquired the gene horizontally from its cyanobacterial hosts (Lindell et al., 2005; Clokie et al., 2006; Lindell et al., 2007).
An interesting characteristic of cyanophages is that many of them possess a holin gene, whose product is needed to permeabilize the host membrane at the end of the infectious cycle, but lack the endolysin gene needed to digest the cell wall (Chen and Lu, 2002; Sullivan et al., 2005; Pope et al., 2007). This may reflect a mechanism of host lysis which only requires holin or that a different, unidentified gene and hence protein is involved in the process. Cyanophages also carry many genes for as yet unidentified functions.

The life cycle of cyanophages is usually quite extended with the latent phase lasting for a number of hours. The burst size is usually quite low. This is in sharp contrast with the enteric phages, in which latent phases are shorter and burst size is typically a few hundred new phage particles per cell. Part of the difference may be due to the lower optimal growth temperatures for cyanobacteria, usually room temperature or close to it, vs. 37°C for enteric hosts. The lower growth rates of cyanobacterial hosts could also be a contributing factor.

Most studies of cyanophages have focused on their ecology and genomics. The structure and assembly of cyanophages have not been investigated in any detail. One of the few exceptions is a study of the Prochlorococcus bacteriophage P-SSP7 (Liu et al., 2010). It revealed very similar structures of this podovirus to phage T7 and epsilon 15. P-SSP7 has a short tail and dsDNA packaged in an icosahedral head. There is a core of internal proteins, which probably has similar function to the core of T7—building a channel through the host cell wall and cell membrane needed to transport the DNA all the way to the cytoplasm.
1.7 Bacteriophage assembly

Viral assembly was first studied in detail in bacteriophages. The model systems were mainly bacteriophages of enteric bacteria. Some of the best-studied phages are T4, T7, P22, λ, HK97. These studies facilitated assembly studies of animal viruses not only through the acquired knowledge but also through the development of some of the crucial techniques.

As soon as new phage protein production commences in the infected host cell the assembly of new particles begins. There are variations from phage to phage but there is a common assembly pathway shared by many of the well-studied dsDNA phages. Since Syn5 is a short-tailed phage the model described here as an example will be the assembly of the best-known Salmonella podovirus—P22 (Fig. 1–4) (King et al., 1973; Casjens and King, 1974; King et al., 1976; Botstein et al., 1973).

One of the key proteins in building phage P22 particles is the coat or capsid protein (gp5), which builds the icosahedral shell of the phage head. It is the highest copy number protein in any phage (415 copies in P22) and forms the most intense protein band when purified mature virions are separated by SDS-PAGE. Coat protein has the remarkable property of assembling two structures with different conformations from the same polypeptide chain—pentameric and hexameric arrangements. The mechanism making this possible is still unclear (Caspar, 1962). However, in P22 and many other phages the coat protein is not capable of assembling shells on its own. Instead it needs the assistance of a chaperone or scaffolding protein (gp8). The roles of the scaffold and portal in procapsid formation are critical, as they participate in the initiation of the process and ensure the proper organization of the coat molecules in the correct conformation (Bazinet, 1985; Fane and Prevelige, 2003).
In a number of phage systems, the coat protein expressed on its own produces only aberrant structures and no shells (or very few) are assembled (Kemp et al., 1968; Casjens, 1974; Kellenberger, 1990; Cerritelli and Studier, 1996a; Huet et al., 2010). Sometimes the presence of scaffolding only is not sufficient since other host chaperones (GroEL and GroES) are also needed to complete the process successfully (van Duijn et al., 2006). By themselves, the coat and scaffolding proteins can form spherical shells, but they are not capable of packaging DNA (Casjens, 1974; Serwer and Watson, 1982; Fu and Prevelige, 2009). In order for these structures to be functional a few additional proteins are needed. These are usually the portal (gp 1, forms the DNA translocation channel, Fig. 1-4) and three core proteins, connected to the portal inside the capsid (gp16, gp20 and gp7; build the phage DNA entry channel during infection).

The complex of the portal and the internal proteins occupies a unique vertex of the protein shells, which are then called procapsids or proheads (the term procapsids will be used exclusively hereafter). These particles are transitory but their assembly completes the first step towards forming mature infectious phage particles. The next step is packaging the DNA inside the procapsids. That happens after the scaffolding is released and the DNA is pumped inside the shell through the portal channel. The spherical head expands into an icosahedron. However, it is not known whether the expansion occurs prior to the start of DNA packaging or is triggered by its commencement. As the DNA is pumped into the shell, with the assistance of accessory proteins (gp2 and gp3 for P22), the scaffolding leaves the procapsid through the openings in the five-fold and six-fold units. The fate of the scaffolding protein is different in different phages.
Figure 1–4. Assembly of procapsids in dsDNA bacteriophages. The model is based on the assembly pathway of P22. Image adapted from Lander et al., 2006.
In P22 it is recycled and continues to build new procapsids while in others it is digested (Dokland, 1999). In P22, before the addition of the tail components, the needle (gp26), the tailspike (gp9), and two accessory factors are added to the portal ring (gp4 and gp10). The construction and attachment of the tail apparatus also has a role in stabilizing the packaged genome (Steven et al., 2005; Johnson and Chiu, 2007). As seen from the described model, DNA packaging, DNA ejection, the assembly of the host recognition and the attachment of the tail complex, are centered on one vertex—the portal vertex.

Capsid assembly through a procapsid intermediate is not limited to bacteriophages. Eukaryotic viruses, e.g., adenoviruses (Edvardsson et al., 1976; D'Halluin et al., 1978), and herpes viruses (Rixon, 1993; Newcomb et al., 2000) follow the same model. The latter have also been shown to contain a portal-occupied vertex with a role in DNA packaging (Lamberti and Weller, 1996; Newcomb et al., 2001).

1.8 Syn5

Syn5 is a dsDNA, short-tailed cyanophage isolated from the Sargasso Sea (Waterbury and Valois, 1993). Its laboratory host is the cyanobacterial strain *Synechococcus* sp. WH8109 (Waterbury and Valois, 1993; Sullivan et al., 2003), which belongs to marine cluster A of *Synechococcus*, clade II, one of the most widely distributed clades in the world oceans (Scanlan, 2003; Zvirglimaier et al., 2008). Electron microscopic (EM) studies of Syn5 show an icosahedral virion with a head diameter of 60 nm and a short tail (Pope et al., 2007). Cryo-electron microscope (cryo-EM) images reveal a mysterious long flexible fibrous horn-like protrusion always positioned on the tail-opposite vertex of the capsid. The horn is about
30–35 nm in length, 10 nm wide at its base, 2–3 nm wide at its tip and only one is observed per particle (Fig. 1–5). It is flexible and can tilt at the base to about 90°.

Bacteriophages such as φX174, φ29, PRD1, PM2 carry fibers or spike structures on their capsids, but these are usually multiple molecules attached symmetrically on a number of vertices (Bamford and Bamford, 1990; McKenna et al., 1992; Tao et al., 1998; Huiskonen et al., 2004). There are just a few bacteriophages with structures similar to the single horn of Syn5. Two Caulobacter crescentus bacteriophages—φCb13 and φCbK—have long head filaments on the vertex opposite the portal (Guerrero-Ferreira et al., 2011). Also, another Myoviridae cyanophage, Bellamy, with a tail-like appendage on the vertex opposite the tail has recently been isolated (personal communication, Welkin Pope and Roger Hendrix).

Symmetrical cryoelectron microscopic (cryo-EM) reconstruction of the Syn5 capsid reveals an assembly of hexameric and pentameric units. The capsid is decorated with three knob-like proteins associated with each hexamer, as well as with distinctive curved ridges radiating from the pentamer edges (Fig. 1–6). Decoration proteins are found among other bacteriophages, e.g. SPP1, λ, ε15 and T5, but the knobs arrangement of Syn5 is unusual (Casjens and Hendrix, 1974; Becker et al., 1997; Effantin et al., 2006; Jiang et al., 2008).

Syn5 has a 46,214 bp genome similar to T7-like phages, with an RNA polymerase gene and short terminal repeats, Fig. 1–7 (Pope et al., 2007). The sequence encodes 61 predicted open reading frames (ORFs) with an organization typical of dsDNA phages, i.e. DNA replication genes (e.g., primase/helicase and DNA polymerase genes) clustered on the left, and the structural proteins (e.g., coat, portal, tail apparatus) on the right. The gene sequences share high synteny with T7 and with several cyanophages especially Synchococcus bacteriophage P60 and Prochlorococcus phage P-SSP7. Three genes—for
ribonucleotide reductase, thymidylate synthase, and thioredoxin—encode proteins that may be involved in breaking down the host genomic DNA in order for it to be a source of building blocks for the phage DNA. The Syn5 genome also contains a putative holin sequence at a similar genome position to the T7 holin, but no endolysin gene. Photosynthetic genes are absent in Syn5, in contrast with some other cyanophages (Mann et al., 2003; Sullivan et al., 2005; Weigele et al., 2007; Sharon et al., 2009).

The proteins of purified Syn5 infectious particles were separated through sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). As seen in Fig. 1–8, mass spectrometry (MS) of gel-excised bands initially identified thirteen proteins closely resembling structural proteins found in other dsDNA phages, including the coat, portal, three internal core proteins and the tail apparatus components (Pope et al., 2007). However, at least four Syn5 structural proteins—gp53, gp54, gp55, gp57—are novel in that their sequences do not match other proteins sequences in the NCBI database (BLASTp, E value <0.001) (Altschul et al., 1990). The C-terminal domain of gp58 showed a good match to the C-terminal domain of a structural protein in the newly sequenced genome of the *Synechococcus* siphovirus S-CBS2.
Figure 1–5. Cryo-EM images of Syn5.
The horn is indicated on a magnified particle in the left corner (black arrow). Image courtesy of Juan Chang and Wen Jiang.
Figure 1–6. Symmetric reconstruction of Syn5. Pentamers and their distinct ridges as well as hexamers and their knobs are labeled. Image modified from Pope et al., 2007.
Figure 1–7. A map of the Syn5 genome. Image from Pope et al., 2007.
Figure 1–8. Syn5 structural proteins.
The protein profile of CsCl purified Syn5 infectious virions electrophoresed through SDS-PAGE. Table of gene product number (gp), assigned name or designation, and calculated molecular mass. A couple of host proteins band appear in the purified particles if the protein is not pure enough (bands labeled host). The lower host band was identified as WH8109 porins. Scaffolding protein is included for completeness (bottom row), but is not present in the mature virion. ORF, open reading frame; MW, molecular mass in kDa. Image modified from Pope et al., 2007.

<table>
<thead>
<tr>
<th>Gene product #</th>
<th>Predicted protein (T7)</th>
<th>MW, kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>Internal</td>
<td>152</td>
</tr>
<tr>
<td>46</td>
<td>Tail fiber</td>
<td>139</td>
</tr>
<tr>
<td>41</td>
<td>Tail tube</td>
<td>99</td>
</tr>
<tr>
<td>44</td>
<td>Internal</td>
<td>90</td>
</tr>
<tr>
<td>54</td>
<td>???</td>
<td>65</td>
</tr>
<tr>
<td>37</td>
<td>Portal</td>
<td>60</td>
</tr>
<tr>
<td>53</td>
<td>???</td>
<td>47</td>
</tr>
<tr>
<td>39</td>
<td>Coat (capsid)</td>
<td>35</td>
</tr>
<tr>
<td>43</td>
<td>Internal</td>
<td>23</td>
</tr>
<tr>
<td>40</td>
<td>Tail tube</td>
<td>22</td>
</tr>
<tr>
<td>55</td>
<td>???</td>
<td>16</td>
</tr>
<tr>
<td>57</td>
<td>???</td>
<td>14</td>
</tr>
<tr>
<td>58</td>
<td>???</td>
<td>16</td>
</tr>
<tr>
<td>38</td>
<td>Scaffolding</td>
<td>30</td>
</tr>
</tbody>
</table>
1.9 Thesis objectives

Studies of cyanophages from both fresh and marine waters have focused primarily on their ecology and population genetics. Examining the assembly processes of bacteriophages infecting cyanobacteria should reveal aspects of DNA packaging into virions which may be important for understanding gene transfer processes in the oceans. Since cyanobacteria branch early in the phylogenetic tree, these studies will enable us to better understand the evolution of phage structure, assembly, and DNA packaging.

Our ability to grow high titer phage preparations of Syn5 provides the opportunity to use it as a laboratory model for studying intracellular growth and assembly of dsDNA cyanophages.

The first objective of this thesis was to establish the life cycle of Syn5 under the current laboratory conditions (chapter II). Once a synchronized culture of WH8109 was obtained, the one-step growth curve experiment of Syn5 was performed in order to determine the eclipse and latent periods as well as the burst size of this cyanophage. It was also important to confirm that techniques like SDS-PAGE and Western blots of infected cell lysates would allow the efficient monitoring of Syn5 proteins in vivo.

Syn5 exhibits some novel features such as the horn and the knob arrangements. To define which proteins build these structures as well as their function, in vitro studies were performed with three of the novel proteins—gp53, gp54, gp58—and described in chapter III. To address the question of whether Syn5 can assemble procapsids in vitro, the key proteins in a putative procapsid structure—the coat, scaffolding and portal proteins—were expressed in
*E. coli* and purified where possible. The coat and scaffolding proteins were co-expressed as well.

The presence of the novel proteins raises the question as to which parts of the Syn5 structure do they contribute. To address this question, antibodies were raised against three of the novel proteins and immunogold labeling was used to localize their positions on the phage surface (chapter IV). The serum was also used to test whether the novel structures play any role in phage infectivity.

Since this work represents the first reported studies of cyanophage assembly, one of the most interesting questions was whether Syn5 maturation includes a procapsid step. Since the putative Syn5 scaffolding protein is not part of the final virion, it first had to be confirmed that it was expressed late during infection. The evident presence of the scaffolding protein in wild-type infected cell lysates raised the next questions: do Syn5 mature virions assemble through a procapsid structure and if yes, can they be purified from infected cells? If such structures do exist which structural proteins build them? Are the horns and the knobs part of the procapsids? These questions are addressed in chapter V and in the results included in the Appendix.

Answering the questions posed in this thesis marks the beginning of a deeper understanding of how the cyanophage world functions on the molecular level. Such studies are needed to improve our knowledge of the large impact of cyanophages on life in the oceans and hence our planet.
Parts of this thesis have been published in:

CHAPTER II: SYN5 INFECTION AND GROWTH
2.1 Introduction

Marine cyanophages are relatively new additions to our understanding of the bacteriophage world with the first representatives isolated in 1981 from the Black Sea (Moisa et al., 1981). The first challenge in investigating these bacteriophages often arises in growing their hosts to sufficiently high densities in synchronized cultures. This, in turn, limits the preparation of highly concentrated and pure phage stocks, needed for detailed studies. The majority of work reported to date on marine cyanophages has focused on their isolation, propagation and purification and lately on their genome sequences (Mann, 2003; Clokie et al., 2010). There are very few reports on protein structure, molecular biology and genetics of cyanophages, studies where high titre phage stocks are mandatory (Pope et al., 2007; Weigele et al., 2007; Liu et al., 2010; Kuznetsov et al., 2012). However, there is increasing evidence for the importance of marine viruses in the oceanic cycles and life on our planet (Suttle, 2007; Danovaro et al., 2011). This was one of the major motivations for the research described in this thesis.

The phage-host system of cyanophage Syn5 and Synechococcus sp. WH8109 was chosen for this study due to its robustness. Laboratory conditions for the growth and infection of host cultures as well as a Syn5 purification protocol had been described earlier (Pope et al., 2007). The existing protocols were optimized here (See Materials and Methods, this chapter).

To undertake studies of the infection process and particle assembly of phage Syn5, it was important to first define a detailed growth curve for Synechococcus sp. WH8109 in order to determine the optimal growth stage and culture conditions for host infection. It was also essential to optimize the plaque assay for Syn5 in order to monitor
the development of infectious phage particles. This permitted defining the phases of the life cycle of the phage and the optimal times to harvest mature phage particles as well as possible phage assembly intermediates.

2.2 Materials and Methods

2.2.1 Growth of Synechococcus sp. WH8109

The Syn5 host used in all experiments was *Synechococcus* sp. WH8109 (kindly provided by John Waterbury), grown in artificial seawater (ASW) or SN medium. ASW was prepared as described by Wyman *et al.* and modified by Lindell *et al.* with some further adjustments (Wyman *et al.*, 1985; Lindell *et al.*, 1998). The components dissolved in double-distilled water (MilliQ-purified) were: 428 mM NaCl; 9.8 mM MgCl$_2$$\cdot$6H$_2$O; 6.7 mM KCl; 17.8 mM NaNO$_3$; 14.2 mM MgSO$_4$ (anhydrous); 3.4 mM CaCl$_2$$\cdot$2H$_2$O; 0.22 mM of K$_2$HPO$_4$$\cdot$3H$_2$O, 5.9 mM NaHCO$_3$; 9.1 mM Tris. The pH was adjusted to 8.0 with HCl. After autoclaving (20 min per 1 L) and cooling to room temperature, Trace Metals mix (0.77 µM ZnSO$_4$$\cdot$7H$_2$O; 7.0 µM MnCl$_2$$\cdot$4H$_2$O; 0.14 µM CoCl$_2$$\cdot$6H$_2$O; 30 µM Na$_2$MoO$_4$$\cdot$2H$_2$O; 30 µM citrate; 5 µM ferric citrate) was added, as well as 15 µM EDTA (disodium salt) and 100 µM Na$_2$CO$_3$.

SN liquid medium was prepared as described by Waterbury and Willey (without the addition of vitamins) (Waterbury and Willey, 1988). The seawater used was from Woods Hole Oceanographic Institution, Woods Hole, MA or the Marine Science Center of Northeastern University, Nahant, MA. The seawater was filtered through 0.2 µm Supor filters (Pall Corporation) to remove debris, mixed with double-distilled water (MilliQ-purified) in a ratio of 3:1 (seawater: water) and autoclaved under the same
conditions as ASW. After cooling to room temperature the medium was supplemented with 9 mM NaNO₃, 90 µM K₂HPO₄, 100 µM Na₂CO₃, 15 µM EDTA (disodium salt), and 1 ml Trace Metals mix per L medium (the same used for ASW).

The cells were grown in controlled environmental chambers (Percival Scientific) at 28°C under continuous cool white fluorescent light at irradiance of about 20–50 µmol of photons m⁻² s⁻¹. The culture vessels were the same as described previously (Pope et al., 2007). These were cleaned by soaking for a few hours to a few days in about 0.01-0.1% Micro-90 solution (Cole-Parmer), rinsed with MilliQ water followed by a second soak step in 10% HCl (a few hours to overnight). In the final step the vessels were rinsed and soaked overnight in MilliQ water. Cells were counted in a Petroff-Hauser chamber with an epifluorescence microscope (Zeiss Axiostar plus) fitted with a fluorescein isothiocyanate (FITC) filter. To store the cells, an exponentially growing culture was centrifuged at 9,500 × g for 10 min and the cell pellet was resuspended and concentrated 10-fold in fresh culture medium. Dimethyl sulfoxide (DMSO, 7% final concentration) or sterilized glycerol (15%) was added as a cryoprotectant, equilibrated with the cells for 10 min at room temperature, and the suspensions aliquoted in cryogenic sterile tubes (Corning) and stored at -80°C. To recover cells from frozen stocks, the suspensions were thawed on ice, transferred to ASW or SN liquid medium and, for the first two days of incubation, shielded with aluminum window screens to reduce the light irradiance to about 10 µmol of photons m⁻² s⁻¹.
2.2.2 Phage growth

Syn5 was purified as described in Pope et al. with some modifications (Pope et al., 2007). In brief, Synechococcus cells in mid-exponential phase (2–8 × 10^8 cells/ml) were infected with Syn5 with a multiplicity of infection (MOI) of 0.001. At 4–5 h after infection the first signs of lysis appeared—green cell debris on the culture vessel walls and transition from the deep red color of the culture to bright pink. For convenience the infected cultures were usually left in the incubator overnight; the cultures were typically completely cleared at 16–21 h post-infection (Fig. 2–1).

To obtain complete lysis, 0.1–1% Triton X-100 and 0.01 mg/ml lysozyme (optional) were added and the lysate stirred at room temperature for about 1 h. When phage samples were prepared for cryoelectron microscopy, chloroform (0.1–1%) was also added at this step. The cell debris was then pelleted at 9,500 × g for 15 min at 4°C and the supernatant stored at 4°C. If debris was still present in the supernatant it was filtered through a Glass Microfibre 47 mm filter (Whatman). To improve phage precipitation, 500 mM NaCl (925 mM final concentration) were dissolved in the supernatant followed by addition of 10% (w/v) polyethylene glycol (PEG) 8000 and stirred overnight to five days at 4°C.

The phages were pelleted at 9,500 × g for 30 min at 4°C and resuspended in phage buffer (50 mM Tris pH 8.0 or 7.5, 100 mM NaCl, 100 mM MgCl_2 ). The suspension was loaded onto a CsCl step gradient with ρ (density in g/cm^3) of 1.2, 1.25, 1.3, 1.4, 1.5 and 1.6 layers; each prepared in phage buffer, 50 mM Tris (pH 7.5), 100 mM NaCl, 100 mM MgCl_2) followed by ultracentrifugation at 28,000 rpm for 3 h at 8°C in a Beckman SW28 rotor. The Syn5 scattering band typically sedimented at the interface of the 1.4 and 1.5 layers. The phage band was withdrawn with a needle syringe, and dialyzed stepwise.
against 50 mM Tris pH 7.5, 100 mM MgCl$_2$ buffer containing decreasing salt concentrations as follows: 2 M (1 h), 1 M (1 h) both at room temperature, and 100 mM NaCl (2 h to overnight) at 4°C. If needed the phage were concentrated through 100,000 MWCO centrifugal regenerated cellulose filters (Amicon Ultra, Millipore) at 1,500 × g for 10–30 min (depending on the volume). Phage stocks were stored at 4°C.

2.2.3 Phage titering

Serial dilutions of phage suspensions were prepared in ASW. ASW agar (1% LMP Agarose, Invitrogen) was melted and held in a 32°C water bath until use. A sample of each dilution was mixed with 1 ml of exponential phase *Synechococcus* cells (4–7 × 10$^8$ cells/ml) and, after 2–3 min at room temperature, 3.5–5 ml of ASW agar were added, the sample mixed and poured into a 50 mm Petri plate. Plaque-forming units were counted following overnight incubation (Fig. 2–2) under the same conditions used to grow cells (see “Host growth”). If the culture used for plating was in the very early exponential phase the lawn color was very pale pink and the plaques could be difficult to see. Looking at the plaques by positioning the plates against different light intensities would usually allow their visualization and counting.
Figure 2–1. Syn5 infecting WH8109. Uninfected (left) and Syn5-infected (right) cell cultures of *Synechococcus* sp. WH8109. The lysed culture is at about 14 h post-infection with MOI of about 0.0001.
Figure 2–2. Syn5 plaques.
Syn5 plaques (black arrows) developing on a lawn of *Synechococcus* sp. WH8109 after about 16–20 h of incubation at 28°C.
2.2.4 One-step growth curve of Syn5 in the Percival chambers

*Synechococcus* sp. WH8109 was cultured in ASW medium (as described in “Host growth”) and grown to mid-exponential phase. The culture was then diluted two-fold in fresh medium the night before the experiment and the incubation continued. Cultures with cell densities of $3–7 \times 10^8$ cells/ml were used. Only cells with bright yellow fluorescence were counted. A 20 ml aliquot was transferred to a 125 ml Erlenmeyer flask, infected with Syn5 at MOI=0.01 and incubated shaking (400 rpm; VWR Mini Shaker, cat. # 12620–938, orbital diameter 3 mm) in a Percival light chamber. Further phage attachment was interrupted at 14 min after infection by diluting the culture ($2 \times 10^4$ and $2 \times 10^5$-fold) in pre-warmed ASW medium in two 125 ml flasks, followed by continued incubation under the same conditions. Samples were collected from the $2 \times 10^4$ dilution flask at times from 30–70 minutes and from the $2 \times 10^5$ dilution flask at times from 80–120 min. At time 0 and 14 min of infection, samples were collected from the original undiluted flask. Two sets of samples were collected at each time point. To determine the eclipse period, the first set of samples was treated with 1% (v/v) chloroform and incubated for at least 1 h before dilution and plating. To define the latent period, the second set of samples was left untreated, held on ice and plated within a few minutes of collection. All samples were titered as described above. The experiment was repeated three times and the mean of each time point is plotted in Fig. 2–5.
2.2.5 Syn5 protein analysis during infection

To analyze host and phage proteins during infection, 50 ml of cells were collected at different times after infection, concentrated 100-fold by centrifugation at 9,000 × g for 10 min, and resuspended in 0.5 ml of lysis buffer (50 mM Tris, pH 8, 100 mM NaCl) and stored at -20°C. The cells were thawed and treated with lysozyme (2 mg/ml), sonicated for 5 × 15 sec, and treated with DNase I (1 U/ml; Worthington) all at room temperature. The cell debris was pelleted at 10,000 × g for 10 min, the supernatants were decanted and retained, and the pellets were resuspended in the same volume of resuspension buffer. Pellets and supernatants were diluted 1:2 and the proteins separated by SDS-PAGE (12%) at 100 V for 2 h after boiling in reducing buffer (60 mM Tris, pH 6.8, 2% SDS, 5% β-mercaptoethanol, 10% glycerol, bromophenol blue stain) at 80–85°C for 10 min. Gels were Krypton-stained (Pierce Biotechnology) and visualized on a Typhoon 9400 scanner (GE Life Sciences). For Western blots, the gels were equilibrated in transfer buffer (10% ethanol, 25 mM Tris base, and 192 mM glycine) and the proteins transferred to a polyvinylidene fluoride (PVDF) membrane (0.45 µm, Millipore) overnight in the same transfer buffer at 15 V constant voltage and 4°C (Bio-Rad Criterion transfer apparatus). The membranes were washed in PBS-Tween 20 (0.1% v/v) and blocked with 5% Carnation nonfat dry milk in PBS (phosphate buffered saline, pH 7.5) overnight at 4°C. The blot probing steps were as described in the enhanced chemifluorescence (ECF) Western blot manual (Amersham). A mixture of primary rabbit antibodies—anti-Syn5 (Covance), anti-gp53 (Strategic Diagnostics), anti-gp54 (Pacific Immunology), and anti-gp58 (Strategic Diagnostics)—all diluted 1:1,000 was used for the membrane probing in Fig. 2–6B. The secondary antibodies used were alkaline phosphatase (AP) conjugated goat anti-rabbit antibodies (Bio-Rad) at a dilution of 1:3,000. The membranes were
incubated for 1 h at 4°C with the primary antibodies, washed and incubated for 45 min at room temperature with the secondary antibodies. For visualization they were soaked for 10 min in ECF substrate (Amersham) and scanned on a Typhoon 9400.

2.3 Results

2.3.1 Growth of *Synechococcus* sp. WH8109

Growing and maintaining a synchronized culture of the host is the first prerequisite for successful host-phage experiments. The host growth results were obtained with the cultures grown in SN medium while most of the host-phage experiments were performed in ASW. However, the growth trend was very similar in both media.

For the growth curve experiment the WH8109 cultures were started from a seed culture that was diluted about 1:60. The light irradiation was about 20 µmol of photons per m² s⁻¹. The experiment was performed twice in triplicate but only one of the trials is presented here—the one performed for 360 h. The cell count was determined via counting on a Petroff–Hausser chamber on an epifluorescence microscope with a FITC filter. In the exponential phase, especially mid-exponential, the culture was homogenous with majority of the cells (>95%) being round with bright yellow fluorescence, which could also be seen as a halo surrounding the cells, and any slightly elongated cells were usually couples of round dividing cells that had not completed their separation (Fig. 2–3). In the stationary phase, the culture was turning heterogeneous and although the majority of the cells were still the ones with yellow fluorescence (about 60–70%) the rest of the
cells were smaller with green fluorescence and elongated (bacilli-like) yellow cells sometimes grouped in chains of 2–3 or more.

The mean generation time (doubling time) under the tested conditions was about 26.6 h (Fig. 2–4). The generation time could be shortened by increasing the light irradiation to 40–50 μmol of photons per m² s⁻¹. It is worth noting that growing the cells in ASW also shortened their doubling time.

The WH8109 cultures used in all experiments were adapted to grow only under continuous light from the beginning of their recovery from frozen stock. If the light irradiation was interrupted for 10–12 h or longer, the cultures stopped growing and did not resume growth even if they were transferred to fresh medium and illumination was restored.

2.3.2 One-step growth curve of Syn5

To characterize the growth of Syn5 in *Synechococcus* sp. WH8109, a one-step growth experiment was performed with cells in mid-exponential growth phase in ASW, at 28°C and irradiated under continuous light, at MOI of 0.01 phages per cell (Fig. 2–5). The presence of infectious particles was assayed by a plaque assay on WH8109 lawns incubated overnight at 28°C. Under the experimental conditions (Fig. 2–5), the first infectious particles appeared about 43 min post infection (eclipse period, chloroform treated samples, dashed line) and cell lysis began at around 60 min (latent period, no chloroform, solid line). The phage growth cycle was considerably shorter than the doubling time of the host, which was about 15 h under the same conditions.
Figure 2–3. Images of *Synechococcus* cells under epifluorescent microscope. Cells are imaged with a FITC filter. Images A–C represent the majority of the cells in an exponentially growing and actively dividing culture with bright fluorescent single cells (A); cells with halos (B) visible when the cells are slightly out of focus; and cells close to completing the separation (C). Older and dying cells are usually smaller, have green fluorescence and do not exhibit halos (D). Magnification 400 ×.
Figure 2–4. Growth curve of *Synechococcus* sp. WH8109. Growth was performed in SN medium at 28°C and light intensity of 20 μmol of photons per m² s⁻¹.
The growth cycle of Syn5 is the most rapid among cyanophages reported to date (Safferman *et al.*, 1972; Padhy and Singh, 1977; Wilson *et al.*, 1996; Liu *et al.*, 2007; Liu *et al.*, 2008). The average burst size in the triplicate data shown in Fig. 2–5 was about 24 infectious particles per cell, five to six times lower compared with the T7 group (Wang *et al.*, 1999), but comparable to other marine cyanophages (see Discussion).

### 2.3.3 Host and phage proteins during infection

To monitor the dynamics of the major phage structural proteins and major host proteins during infection, *Synechococcus* sp. WH8109 cells were collected by centrifugation at various times after Syn5 infection. The infected cells were lysed, fractionated into supernatant and low speed pellet fractions, and the resulting proteins analyzed by SDS-PAGE and Western blotting (Fig. 2–6).

Comparison of the protein patterns from uninfected and infected cells on SDS gels, revealed that at least some host proteins remained relatively constant after infection, in particular the strong bands at 45, 53, 70 kDa were still present at 60 minutes after infection (Fig. 2–6A). In addition, the two lower molecular mass groups at about 16 and 22 kDa persisted through the infection. The 16-kDa protein band contained the antenna pigments of the host identified by their color in the gel. In comparing protein levels before and after infection it should be noted that infected cells lysed more efficiently than uninfected under the same conditions.
Figure 2–5. One-step growth curve of Syn5.
The two data sets represent samples treated with chloroform (dashed line) and samples without chloroform (solid line) after collection. The cultures (before and after infection) were incubated at 28°C under 45–50 µmol m⁻² s⁻¹ of light irradiation. Each curve represents the average results from three experiments. The lengths of the eclipse and latent periods are marked.
In the analysis of Syn5 proteins present in the mature virions (Fig. 2–6A) the coat protein and at least eight minor bands were visible. In the infected samples, the coat protein was present at 0.5 min after infection in the supernatant fraction. Although the 0.5 min sample was collected immediately following infection, some phage attachment to cells occurred during the centrifugation step. Thus the phage proteins in the supernatants at early times (0.5–20 min after infection) probably reflect the presence of input phage released from the cell membranes during lysis. The increase in coat protein at later times presumably represents new protein synthesized in the infected cells. The remaining viral proteins could not be unambiguously identified on the gel, partly due to interference by host proteins.

To more clearly follow the expression of the phage proteins in the infected cells Western blots were carried out using a mixture of anti-Syn5 serum and polyclonal antibodies raised against the recombinant protein products of novel genes 53, 54, and 58 (Fig. 2–6B). The anti-Syn5 antibodies, raised against whole virion particles, would detect epitopes of proteins exposed at the surface of virions. In the Syn5 particles sample (Fig. 2–6B) the coat, tail fiber, and portal were very well recognized, as were several minor proteins—the two tail tubes (gp40 and gp41), the five novel proteins gp53, gp54, gp55, gp57 and gp58, and interestingly, one of the internal proteins – gp44. It may be that gp44 lies in close proximity to the portal in the internal core and hence was more exposed to the immune system of the rabbit than the other two internal proteins, gp43 and gp45. All of the virion protein bands were absent from the supernatant of uninfected WH8109 cells,
while the pellet contained a protein of unknown function that migrated at a position similar to gp53.

As in the SDS-PAGE, the phage coat produced the strongest band in the Western blot (Fig. 2–6B). It was present most abundantly in the early supernatant fraction; however at 60 min and thereafter, it was recovered in the pellets as well. This may represent aggregated or misfolded coat protein, newly assembled phage adsorbed to cell debris and particles in unlysed cells. In addition to the coat, most of the other proteins were resolved as of 40 min and clearly visible at 60 min. Other proteins with similar intensity patterns to the coat were the portal (gp37) and gp53. Gp54 and the three small novel proteins, gp55, gp57, gp58, were predominantly in the supernatant, clearly present as of 60 min and increased over time. The tail fiber (gp46) and one of the tail tubes (gp41) were present in the pellets (40 min and later).

In summary, it is evident that as the host proteins were fading, the phage proteins were increasing in intensity.
Figure 2–6. Host and phage proteins during Syn5 infection. (A) SDS polyacrylamide gel (12%, Krypton-stained) of supernatants (sup) and pellets (pel) of infected cells collected at different time points from 0.5 to 120 min after infection. 8109, uninfected Synechococcus sp. WH8109 host cells collected 10–20 min before infection. Syn5, infectious particles used in the experiment and run as control for phage structural proteins. The gp (gene product) samples (right panel) contain purified recombinant phage proteins gp37, gp53, gp54, and gp58. Protein annotation: gp45, internal virion protein; gp46, tail fiber protein; gp41, tail tube B; gp44, internal protein; gp54, novel structural protein; gp37, portal; gp53, novel structural protein; coat, capsid protein; gp43, internal virion protein; gp40, tail tube A; gp55, gp57, and gp58, novel structural proteins; M, protein standards (Bio-Rad, Precision plus). (B) Western blot of a gel run in parallel to the one in 3a, probed with a mixture of anti-Syn5 polyclonal antibodies and rabbit serum raised against the recombinant forms of gp53, gp54, and gp58.
2.4 Discussion

*Synechococcus* sp. WH8109 is a member of marine cluster A, clade II, (Scanlan, 2003) of the *Synechococcus* photosynthetic family, which is widely distributed in the world’s oceans (Zwirglmaier *et al.*, 2008). The ancestors of these cyanobacteria are likely to be among the earliest photosynthetic prokaryotes to have evolved in the marine environment.

This marine species is also very robust and cultures with synchronous growth were readily obtained both in SN and ASW. The fluorescence of the cells under FITC filter was a very efficient tool to monitor the health and the growth of the culture and assess the optimal conditions under which to start an infection with Syn5. The plaque assay was the crucial method that made possible the one-step growth curve experiment for the cyanophage.

Syn5 had a remarkably short life cycle, with a latent period of 60 min, in comparison with other cyanophages studied under relatively similar experimental conditions (continuous light irradiation in the range of 15–50 µmol m\(^{-2}\) s\(^{-1}\), temperature 25–30°C). The latent periods of other cyanophages published to date are 9 h for S-PM2 infecting *Synechococcus* sp. WH7803 (Wilson *et al.*, 1996), 7 h for the *Nostoc* virus N-1 (Padhy and Singh, 1977), 8.5 h for the AS-1 virus of *Synechococcus elongates* (Safferman *et al.*, 1972), and 5 and 4 h for the two freshwater cyanophages Pf-WMP3 and Pf-WMP4, respectively (Liu *et al.*, 2007; Liu *et al.*, 2008). The burst size of Syn5 (20–30 pfu/cell) was comparable to those reported for marine cyanophage S-PM2.

The ability of Syn5 to form plaques on WH8109 lawn after an overnight incubation is a reflection on its fast life cycle. If the plates were left in the incubator the
plaques kept growing until the whole lawn was cleared which is a feature very similar to the plaque growth of T7 (Yin and McCaskill, 1992). Another important property of our system was the ability to monitor the newly synthesized phage proteins during infection via SDS-PAGE and Western blots. Despite the input of phages to start the infection there was a clear indication about the increase in the protein levels because of the newly synthesized phage proteins as a result of the infection.

The increase in protein expression correlated perfectly with the different phage developmental stages in the one-step growth curve. All these features make Syn5 a very promising model system to study in depth the physiology and molecular biology of the infectious process in cyanophages.
CHAPTER III: STRUCTURAL PROTEINS OF SYN5
3.1 Introduction

The first step towards investigating the biochemical properties and functions of the Syn5 phage and possible precursor structures, such as procapsids, was to express and study the key structural proteins in vitro. The major proteins of interest were the portal, scaffold, coat, tail fiber and the novel proteins. The portal, coat and scaffolding proteins were of interest because it has been shown in all dsDNA bacteriophages that these proteins are the major building blocks for procapsids, phage precursor shells (Johnson and Chiu, 2007; Cole et al., 2008; Aksyuk and Rossmann, 2011). Since the Syn5 studies described here are the first assembly studies in cyanophages it was important to know if they follow similar assembly pathways to those of other phages. The tail fiber protein was important because a recombinant version of it can be used to search for the Syn5 host receptors on the surface of *Synechococcus* sp. WH8109 cells.

Originally, the profile of the phage virion structural proteins on SDS-PAGE and Western blots revealed three proteins, which had no match in the PDB (see Fig. 1–8 in Chapter 1) and whose functions were unknown. N-terminal sequencing of these proteins and mass-spectrometry revealed they were encoded by genes 53, 54 and 58 of the Syn5 genome (Pope et al., 2007). In subsequent mass-spectrometry of the gp58 protein band two additional proteins from the Syn5 genome were localized—gp55 and gp57. The five novel structural proteins are the primary candidates for the horn structure and the knobs on the capsid head. With the exception of gp58, the other novel proteins had no matches in the NCBI protein database (as of 03/30/2012) (Table 3–1).
<table>
<thead>
<tr>
<th>gene product</th>
<th>homologous proteins with $E$ value $&lt; 0.001^*$</th>
<th>Secondary structure prediction (jnet)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>gp53</td>
<td>-</td>
<td>80-85% β-sheets and about 15-20% α-helices</td>
</tr>
<tr>
<td>gp54</td>
<td>-</td>
<td>70% β-sheets, with alpha domains in 220-260 aa region and in the C-terminal region</td>
</tr>
<tr>
<td>gp55</td>
<td>-</td>
<td>mostly α-helices in the N and C-terminal domains and predominantly extended β-sheets in the middle</td>
</tr>
<tr>
<td>gp57</td>
<td>-</td>
<td>exclusively α-helical domains were predicted</td>
</tr>
<tr>
<td>gp58</td>
<td>gp20, cyanophage S-CBS2</td>
<td>3 α-helical domains in the first 110 aa region; β-sheets towards the C-terminus</td>
</tr>
</tbody>
</table>

Table 3–1. Summary of *in silico* data obtained for gene products 53, 54 and 58.

* Data obtained on Mar 30th, 2012 using Blastp.

**Each prediction was made solely on the single protein sequence because no matches were found in the database used to generate alignments. (Cole *et al.*, 2008).
3.2 Materials and Methods

3.2.1 Cloning, expression, and purification of the Syn5 novel proteins (gp53, gp54 and gp58)

Reverse and forward primers were designed and used to amplify the genes by PCR from Syn5 genomic DNA. The primers included restriction sites for subsequent cloning in pET21a as follows: g53—NdeI and SacI; g54—NdeI and BamHI; g58—NdeI and BamHI (see cloning region in Fig. 3–1).

As cloned, each protein carried a C-terminal 6-His tag. The clones were verified by complete sequencing of the inserted genes (Harvard MGH DNA Sequencing Core Facility). The constructs were individually transferred into Escherichia coli (E. coli) BL21-Gold (DE3) (Stratagene) for overexpression. The proteins were expressed for 3 h at 26°C following induction with 1 mM IPTG. The cells were pelleted and resuspended in lysis buffer (25 mM Tris pH 8, 300 mM NaCl, 30 mM imidazole and 0.1% Triton-X) and frozen at -20°C. On the day of purification the cells were thawed at room temperature in the presence of protease inhibitors (Roche, Complete Mini EDTA-free, 1 tablet per 10 ml suspension). Lysozyme (2 mg/ml) was added and mixed for 20 min at room temperature, the cells sonicated (5 × 1 min at 9 Watts) to complete the lysis, and DNase (2 U/ml, Worthington) added to degrade the genomic DNA with mixing for 1 h at room temperature. The lysate was centrifuged at 10,000 × g for 20 min and the supernatant incubated with Ni-NTA (nitrolothriatic acid, Qiagen) agarose (10 ml matrix per cell pellets from 1 L culture) for 1 h at 4°C to ensure complete protein binding. The matrix-protein mixture was poured in an empty glass column (80 ml volume, BioRad) and the matrix left to settle for 1 h at 4°C. The flow-through of the unbound proteins was
collected. The column was washed with 6 volumes of the lysis buffer (without Triton-X 100) to remove nonspecifically bound proteins. To elute the affinity bound proteins of interest, 6 volumes of elution buffer (25 mM Tris pH 8, 300 mM NaCl, 150 mM imidazole) were run through the column. The final wash step was performed with 3 volumes of wash buffer (25 mM Tris pH 8, 300 mM NaCl, 300 mM imidazole). The collected fractions were separated by 12% SDS-PAGE for 1.5 h at 150 V. Protein bands of the expected sizes—49.5 kDa for gp53, 67.6 kDa for gp54, and 18.9 kDa for gp58—were present in the elution lanes. The predicted pI values for the expressed products were 4.37 for gp53, 5.39 for gp54 and 5.19 for gp58 as predicted by Compute pI/MW tool from the Expasy server.
Figure 3–1. Map of the pET-21 vector cloning region.
3.2.2 Cloning and expression of coat (gp39), scaffolding (gp38), portal (gp37) and tail fiber (gp46) proteins and purification of the portal (gp37) and scaffolding (gp38) proteins

The cloning vector pET-15b (Novagen), which provides a thrombin cleavable N-terminal 6xHis-tag (Fig. 3–2), was chosen. Syn5 genes—g39, g38, g37 and g46—were amplified by PCR with the addition of NdeI and BamHI restriction sites at the 3’ and 5’ ends, respectively (Fig. 3–2). Following verification of the wild-type gene sequence by full-length DNA sequencing, the constructs were transferred into Escherichia coli (E. coli) BL21-Gold (DE3) (Stratagene) for overexpression. The proteins expression and purification was under the same conditions as for gp53, gp54 and gp58.

Overexpressed protein bands with the expected sizes for the cloned products—gp39 (37 kDa) and gp46 (141 kDa)—were present in the pellet fractions of the cell lysates.

Liquid chromatography-electrospray mass spectrometry (MS) analysis was performed for molecular mass confirmation of the purified products (Proteomics Core Facility, MIT). The proteins were diluted to an appropriate concentration, desalted with a C4 ZipTip and analyzed on a QSTAR Elite mass spectrometer by isocratic elution from a reverse phase trap with acidified aqueous solvent, containing acetonitrile. The spectra generated from the elution were averaged and the molecular mass spectrum was generated with QSTAR BioTools software. The expected mass accuracy is 50 ppm or better.
Figure 3–2. Map of the pET-15b vector cloning region.
The scaffolding protein (gp38) migrated by SDS-PAGE as a 40-kDa band instead of the expected 30-kDa band. The purified product gave a molecular mass of 32,360 Da, in agreement with the theoretical value of 32,360.8 Da for the expressed fusion product of scaffolding protein, linker peptide and His-tag sequence. This product (2 mg/ml) was used to raise polyclonal antibodies in two New Zealand white rabbits by a standard protocol (Strategic Diagnostics, Inc.).

The purified protein product of gp37 has a theoretical molecular mass of 61,129 Da (excluding the N-terminal Met); the mass-spectrometry analysis identified a major species of 61,130 Da and a few minor products. The product of gp54 has a theoretical mass of 67,529 Da (excluding the N-terminal Met) and the analysis showed a single peak with a mass of 67,531 Da. The calculated mass of the expected protein product of gp58 is 18,796 Da. There were two products present in the purified gp58 sample, one with a molecular mass of 18,797 Da, the other with a mass of 18,829 Da. It is possible that the mass difference between the two species is due to an oxidation event since the difference equals the mass of two oxygen atoms. The gp53 sample did not generate protein spectrum after the elution from the column.

The theoretical pI values of the protein products as computed by the Compute pI/MW tool from Expasy were 4.40 for gp38, 6.53 for gp37, 5.98 for gp39 and 4.48 for gp46.
3.2.3 Co-cloning and expression of Syn5 scaffolding (gp38) and coat (gp39) proteins

Syn5 genes—g38 and g39—were amplified by PCR with the addition of NcoI and HindIII restriction sites at the 3’ and 5’ ends, respectively, for g38, and NdeI and AvrII restriction sites at the 3’ and 5’ ends, respectively, for g39. The two products were inserted into cloning vector pETDuet-1 (Novagen) with g38 cloned in multiple cloning site 1 (MSC1) and g39 cloned in MSC2 (Fig. 3–3). Following verification of the wild-type gene sequence by full-length DNA sequencing (Genewiz), the construct was transferred into Escherichia coli (E. coli) BL21-Gold (DE3) (Stratagene) for small-scale expression (15 ml culture). Following induction with 1 mM IPTG, the proteins were expressed for 6 h at three different temperatures—23ºC, 30ºC and 37ºC, and overnight at 18ºC. One ml aliquots of cells were harvested and pelleted (10 min at 5000 × g) at different time points after induction, resuspended in 100 µl lysis buffer (50 mM Tris pH 7.3, 100 mM NaCl) and frozen at -20ºC. To check the protein expression levels, thawed cells were treated with lysozyme at room temperature (2–2.5 mg/ml; time) in the presence of protease inhibitors (Roche Complete Mini EDTA-free) and the extract treated with DNase I (2 U/ml, Worthington) for 30 min at room temperature. The cell debris was separated from the supernatant by centrifugation at 10,000 × g for 10 min at 4ºC. Pellets and supernatants were diluted as needed and the proteins separated by SDS-PAGE (12%) at 150 V for 1.5 h after boiling in reducing buffer (60 mM Tris, pH 6.8, 2% SDS, 5% β-mercaptoethanol, 10% glycerol, bromophenol blue for color) at 80–85ºC for 10 min. Gels were stained with Coomassie Brilliant Blue R-250 or Krypton protein stain (Pierce) and scanned.
pETDuet-1 cloning/expression regions

Figure 3–3. Map of the pETDuet-1 vector cloning regions.
To perform Western blots, freshly run SDS gels were equilibrated in transfer buffer (10% ethanol, 25 mM Tris base, unbuffered, and 192 mM glycine) and the proteins transferred to a polyvinylidene fluoride (PVDF) membrane (0.45 µm, Millipore) overnight in the same transfer buffer at 15 V constant voltage or for 1.5 h at 300 mA constant amperage (Bio-Rad Criterion transfer apparatus) in both cases at 4ºC. The membranes were washed in PBS-Tween 20 (0.1% v/v) and blocked with 5% Carnation nonfat dry milk in PBS (phosphate buffered saline, pH 7.5) from 1 h to overnight at 4ºC. The blot probing steps were as described in the enhanced chemifluorescence (ECF) Western blot manual (GE Healthcare). The primary rabbit antibodies used for membrane probing were anti-Syn5 (Covance) at a dilution of 1:1,000 to detect the coat (gp39) expression and anti-gp38 antibodies (Strategic Diagnostics) diluted 1:2,600 to detect the scaffolding protein. Following washes in PBS-T (1x PBS with 0.05% Tween-20), the membranes were incubated with secondary alkaline phosphatase (AP) conjugated goat anti-rabbit antibodies (GE Healthcare, diluted 1:10,000) or AP-conjugated goat anti-rabbit antibodies (BioRad, diluted 1:3,000). The membranes were incubated for 1 h at room temperature with the primary antibodies, washed in the same wash buffer as above and incubated for 45 min at room temperature with the secondary antibodies. For visualization they were soaked for 10 min in ECF substrate (Amersham) and scanned on a Typhoon 9400 (GE Life Sciences).
3.2.4 Transmission electron microscopy

For observation by electron microscopy the proteins were diluted to 0.1–0.5 mg/ml and stained negatively by the following protocol. The protein sample (4 µl drop) was applied to a glow-discharged Formvar/Carbon coated 200 or 400-mesh copper grid and allowed to adhere for about 1 min. The grid was edge-blotted on filter paper and stained by floating on the surface of a 1% uranyl acetate drop or 4 µl of the stain were applied straight to the grid surface instead. Stained for about 30 sec and edge-blotted on filter paper to remove the stain and air-dried for at least 30 min. Observed under the electron microscope at 60 kV at magnifications ranging from 100,000–250, 000 ×.

3.3 Results

3.3.1 Expression and purification of the recombinant portal (gp37), coat (gp39), tail fiber (gp46) and putative scaffolding proteins (gp38)

Good expression of all four proteins was obtained under the tested conditions but the overexpressed proteins identified as coat (gp39) and tail fiber (gp46) based on molecular mass, were insoluble and were not purified (Fig. 3–4). The tail fiber protein was expressed at lower levels than the other proteins with this induction protocol (3 h of expression at 26°C after induction with 1 mM IPTG). The tail fiber is a much longer polypeptide chain than the other three (146 kDa). The insolubility of coat and tail fiber is not surprising since such results have been reported with other bacteriophages (see Introduction). The tail fiber protein is most likely fibrous in nature, due to its function,
while many coat proteins have been reported to be insoluble due to their inability to fold properly in the absence of their chaperone—the scaffolding protein.

The portal (gp37) and the scaffolding (gp38) proteins were soluble, with the majority of the products appearing in the supernatant fraction of the cell lysate and very little in the pellet. Both were successfully purified via Ni column affinity purification (the results from the purification of gp37 are shown in Fig. 3–5). The molecular masses of the purified products were confirmed via mass spectrometry (see Materials and Methods for details).

The portal protein of Syn5 (gp37) precipitated under low salt conditions in the storage buffer (25 mM NaCl, 50 mM Tris pH 7.6, 2 mM EDTA). The same property was observed for the portal of bacteriophage T7 (Cerritelli and Studier, 1996b). For both proteins increasing the salt concentration to 100 mM dissolved the precipitate.
Figure 3–4. SDS gels of the overexpressed protein products of gp39 and gp46. The data for gp39 is in panel A) and for gp46 in panel B). Overexpression was induced with 1 mM IPTG for 3 h at 26°C. The gene products are marked by arrows. M, protein marker lane (BioRad); cr., crude lysate; pel., pellet fraction; sup., supernatant fraction. The gels were stained with Coomassie Brilliant Blue.
Figure 3–5. Purification of the recombinant portal protein (gp37).
SDS gel of the Ni column affinity purification fractions. The gene product (arrow) was expressed for 3 h at 26°C after induction with 1 mM IPTG. M, protein marker lane (BioRad); cr., crude lysate; pel., pellet fraction; s., supernatant fraction; fl. fr., flow through fractions; w. fr, wash fractions. The gels were stained with Coomassie Brilliant Blue.
Scaffolding proteins, while required for assembly in dsDNA enteric bacteriophages, are absent from the mature virion. Gene 38 from the Syn5 genome was tentatively designated as the scaffolding protein gene based on sequence homology with genes of T7 and other bacteriophages. If Syn5 assembly is similar to the majority of the commonly studied dsDNA bacteriophages, it is to be expected that the scaffolding protein would not be present in the mature virion. As seen in Fig. 2–6B (Chapter 2), the anti-Syn5 antibodies raised against whole virions did not recognize any proteins of the expected size for the scaffolding protein. Of course it is possible that the antibodies do not recognize the scaffolding because it might be an internal protein. However, the anti-Syn5 antibodies do recognize most of the virion proteins, which are predicted to be internal based on their sequence and homology to other phage proteins. To examine whether gp38 is the scaffolding protein, its open reading frame was cloned, expressed (Fig. 3–6), and the resulting protein purified.

The product of gene 38 has a theoretical mass of 30.3 kDa and was predicted by Jpred 3 (Cole et al., 2008) to be a α-helical protein with a single β-sheet in the N-terminal region. The recombinant gp38 His-tagged fusion had a total calculated molecular mass of 32,360.8 Da (excluding the N-terminal Met), which was confirmed by LC-electrospray MS analysis. Unexpectedly, it migrated with an apparent molecular mass of approximately 40 kDa (about 30% larger than expected) on 12% SDS-PAGE (Fig. 3–6). Unboiled and boiled preparations of the protein were separated by electrophoresis to examine whether there would be any change in migration.
Figure 3–6. Gel electrophoresis of the Syn5 recombinant scaffolding protein (gp38). The overexpressed scaffold (black arrow, ~40 kDa) is shown in the supernatant fractions of cell lysates of the expression host *E. coli* BL21. Unboil, electrophoresed without boiling; boiled, boiled at 95°C before electrophoresis; M, protein standards (BioRad). The expressed product (arrow) was predicted to migrate at about the position of the 31-kDa protein marker.
However, as seen in Fig. 3–6 its mobility was not affected. Interestingly, this aberrant migration resembles that of the T7 scaffold (Studier and Maizel, 1969; Cerritelli and Studier, 1996a; Larkin et al., 2007) although the two proteins share only about 17\% amino acid identity, in the central and especially in the C-terminal region, when aligned using Clustal W2 (Larkin et al., 2007).

### 3.3.2 Expression of the recombinant novel proteins—gp53, gp54 and gp58

All three novel proteins were overexpressed as soluble products and were purified. As assessed by Bradford assay the yields were 4 mg/L of culture for gp53, 22 mg/L for gp54, and 30 mg/L for gp58. The collected samples were dialysed and stored in 50 mM Tris pH 7.6, 25 mM NaCl and 2 mM EDTA at 4°C. Similar to gp37, gp54 protein precipitated in these conditions, but when the NaCl was raised to 100 mM no precipitate was observed.

As seen in Fig. 3–7, the protein yield of gp53 was low under the expression conditions used. Comparison of the band intensities in the pellet and supernatant lysate fractions shows that more than half of the protein is insoluble. A future modification in the expression procedure, to improve both yield and solubility, could be to carry out the expression at lower temperature, such as 18°C. Recombinant gp54 (Fig. 3–8) was soluble under these conditions and most of it was recovered in the elution fractions. The gp58 product was also soluble (Fig. 3–9). In the elution fractions, a smaller protein was present, which may be a breakdown product of gp58 still carrying the C-terminal His-tag.
Figure 3–7. Purification of the recombinant protein gp53. Gel electrophoresis (10%) of the Ni-NTA purification fractions. The protein elutes as a clean peak in the elution fractions (arrow). The protein was overexpressed for 3 h at 26°C after induction with 1 mM IPTG. M, protein marker lane (BioRad); cr., crude lysate; pel., pellet fraction; sup., supernatant fraction; flow fr., flow through fractions; wash fr, wash fractions. The gel was stained with Coomassie Brilliant Blue.
Figure 3–8. Purification of the recombinant protein gp54. Gel electrophoresis (12%) of the Ni-NTA purification fractions. The protein appears as a clean peak in the elution fractions (arrow). The protein was overexpressed for 3 h at 26°C after induction with 1 mM IPTG. M, protein marker lane (BioRad); cr., crude lysate; pel., pellet fraction; flow fr, flow through fractions; wash fr., wash fractions. The gel was stained with Coomassie Brilliant Blue.
Figure 3–9. Purification of the recombinant protein gp58.
Gel electrophoresis (12%) of the Ni-NTA purification fractions. The protein appears as a clean peak in the elution fractions (arrow). The protein was overexpressed for 3 h at 26°C after induction with 1 mM IPTG. M, protein marker lane (BioRad); cr., crude lysate; pel., pellet fraction; fl. fr, flow through fractions. The gel was stained with Coomassie Brilliant Blue.
3.3.3 The three recombinant proteins were recognized by anti-Syn5 polyclonal rabbit serum

Before sending the recombinant proteins for antibody production, their antigenicity to anti-Syn5 polyclonal antibodies was determined by Western blotting. The *E. coli* lysate supernatants were run on SDS-PAGE and transferred to a PVDF membrane. The gel was silver-stained after transfer to check transfer efficiency. The gp53 and gp58 products transferred completely to the membrane, while gp54 did not transfer as well. The membrane was incubated consecutively in blocking agent, anti-Syn5 polyclonal rabbit serum (Covance) and in anti-rabbit antibodies (Amersham). ECF (enhanced immunofluorescence) was used as labeling agent and the membrane was visualized under fluorescence. The results are shown on Fig. 3–10.

All three recombinant proteins produced fluorescence bands on Western blot (Fig. 3–10). Two paler bands, one below (band 2) and one above (band 1) the gp58 major band, were also present. Both bands are probably related to gp58, since no such bands were present in the gp53 and gp54 lanes, although they contain proteins from the same *E. coli* strain. The dark band present across all lanes at the bottom of the Western is the lysozyme used for cell lysis.
Figure 3–10. Gel electrophoresis and Western blot of the three novel Syn5 structural recombinant proteins. (A) SDS-PAGE (12%) of *E. coli* lysate supernatants of the expressed gp53, 54 and 58. The gel was silver stained. (B) Western blot of a gel run in parallel to the one in (A). Each protein is boxed. M, molecular marker; 1 and 2 mark a couple of protein bands with different molecular mass than gp58 but recognized by the anti-Syn5 antibodies.
3.3.1 Co-expression of gp38 (scaffold) and gp39 (coat) proteins

Since the coat protein (gp39) was insoluble when expressed on its own, one of the plausible explanations was that the coat could not fold properly due to the lack of its chaperone, the scaffolding protein (gp38). One way to test this hypothesis was to co-express the two proteins simultaneously in the same cells.

The expression system of choice was pETDuet-1 (Novagen). In this vector each gene can be cloned in its own multiple cloning site under the control of a separate promoter (T7 promoter). This should guarantee the expression of both proteins albeit at possibly different levels. Different temperatures of growth after induction were tested as well as two different concentrations of IPTG. The control in the expression experiments was a culture transformed with the original pETDuet-1 vector (without inserts) and grown under the same conditions without being induced with IPTG.

Overall, the scaffolding protein was expressed well and at higher levels compared with the coat (Fig. 3–11). This may be due to its position in the construct in the first multiple cloning site. The results from both the SDS-PAGE and the Western blot indicate good expression of the scaffolding protein at both temperatures. The expression of the scaffolding protein was better at the induction temperature of 23°C and the yields dropped as the temperature was increased to 37°C. At all of the expression conditions the scaffolding protein was present in both the supernatant and pellet fractions of the cell lysates.
Figure 3–11. Co-expression of the scaffolding (gp38) and coat proteins (gp39).
Gel electrophoresis (12%) of the cell lysates with co-expressed scaffolding and coat proteins at different temperatures—23°C and 37°C for 6 h (left panel) and Western blot (right panel) of the same samples. A sample of uninduced culture transformed with the original pETDuet-1 vector is included for comparison (unind.) in each case. The gels were stained with Coomassie dye. M, protein standards (BioRad); s., supernatant; p., pellet; scaff., scaffolding.
The expression of the coat protein was low and was best at 37°C where a distinct coat band could be identified on the gel alone (Fig. 3–11). For the other temperatures the coat protein band was clearly identified on a Western blot at both 23°C and 37°C. However, under all tested conditions the coat protein was identified only in the pellet fractions of the cell lysates.

Since both proteins were present in the pellet fraction it was important to determine whether they formed some type of complex. The pellet samples where both proteins were present were loaded on sucrose gradients (5–20%), the gradients fractionated and the fractions were separated through electrophoresis on SDS protein gel and transferred onto a Western membrane (for simplicity only the results from the 37°C expression are shown here in Fig. 3–12). The results showed that most of the scaffolding protein sedimented in the top fractions of the gradient (5%) with a small proportion of what might be aggregates sedimenting in the bottom fraction. All of the coat protein was in the bottom fraction of the gradient. The bottom fraction was collected, the sucrose dialyzed out and the sample was observed under the electron microscope. No procapsids-like structures were detected.
Figure 3–12. Sucrose gradients of cell lysates with co-expressed scaffolding and coat proteins.
Gel electrophoresis of the sucrose gradient (5-20%) fractions of the pellet of the sample expressed at 37°C. The gradient was run for 45 min at 42,000 rpm in the SW 55 Ti rotor.
3.4 Discussion

The Syn5 cloned proteins, in the pET expression system under the control of the T7 promoter were expressed well under the tested conditions. Most of them were soluble, full length, free of major breakdown products and could be purified at good yields. All of the purified proteins (with the exception of gp53) were analyzed through mass-spectrometry and their molecular masses confirmed. Although the functions of the Syn5 structural proteins were assigned based only on homology to other phage proteins, the expressed and purified proteins exhibited features similar, in regard to solubility and other characteristics, to their homologous counterparts, which have been shown experimentally to function in the phage machinery. Remarkable similarities were observed between the recombinant proteins of Syn5 and the homologous proteins of the enteric bacteriophage T7 (Ceritelli and Studier, 1996b, a).

The Syn5 scaffolding protein (gp38) exhibited anomalous electrophoretic migration. Such behavior has also been described for the scaffolds of bacteriophages λ (Ziegelhoffer et al., 1992), the external scaffold of P4 (Dokland, 1999), and, as mentioned earlier, the scaffolding protein of T7. These three proteins possess abnormal gel filtration patterns also, eluting at higher than expected apparent molecular weights, which may be due to elongated shape. The Syn5 and T7 scaffolding proteins both have highly-negatively charged sequences with theoretical pI values of 4.16 and 4.31, respectively. Excessive charge has been shown to alter gel migration (Hu CC, 1995), and replacing acidic amino acids with basic residues in highly negatively-charged proteins can result in the predicted SDS-PAGE migration behavior (Armstrong and Roman,
Another explanation for anomalous migration is the presence of intrinsically disordered regions (Iakoucheva et al., 2001), and a few such sequences were predicted to be present in the Syn5 and T7 scaffolding proteins by PrDOS analysis (Ishida and Kinoshita, 2007).

The precipitation of the portal protein (gp37) in low salt concentrations may be due to ring formation by the protein subunits. As mentioned earlier, this was observed in many other portal proteins including T7 and P22. Since portal protein is building a multi-ring channel needed for packaging and ejection of the DNA, this high molecular weight complex may be formed under low salt concentrations and precipitate out of solution. A similar property was observed for gp54, one of the novel Syn5 proteins. This raises the possibility that gp54 builds a structure that has similar organization to the portal protein in the mature virion.

The inability of the coat protein (gp37) to fold correctly by itself is similar to that of other phage coat proteins (Laemmli et al., 1970; Gordon et al., 1994; van Duijn et al., 2006). In some cases the expression of the coat in the presence of the scaffolding is sufficient to restore that ability and the coat and scaffolding alone can form procapsids. However, this does not seem to be the case with Syn5. Although the coat was not expressed at very high levels in the experiment where it was co-expressed with the scaffolding, all of the expressed protein was insoluble. The abundant scaffolding protein produced failed to rescue any of the expressed coat. It may be that the putative Syn5 scaffolding protein does not assist in the coat folding or that it is not the only chaperone needed to fold the coat. Some bacteriophages need the chaperone system of the host cell
to complete their folding correctly (van Duijn et al., 2006). More experiments are needed to test this hypothesis.
CHAPTER IV: LOCALIZATION OF SYN5 NOVEL PROTEINS BY IMMUNOGOLD LABELING
4.1 Introduction

During the latter half of the 20th century structural studies of bacteriophages were undertaken, focused on a few enteric phages. In the past decade, the search for and discovery of new bacteriophages from a broad diversity of environments intensified. The advances in the development of genomic and proteomic tools in combination with the decrease in their cost allowed a wider variety of bacteriophages from different environments to become subjects of structural and genetic studies. These developments will broaden our knowledge of bacteriophage properties and roles in the natural world, leading to a deeper understanding of viral and host evolution.

With the discovery of novel viruses come novel features. Some of the most remarkable properties have been discovered in marine and freshwater viruses: cyanophages carry host photosynthetic genes that are expressed during phage infection (Lindell et al., 2005; Shan et al., 2008); the largest viruses are of marine origin and mimic microbes in their size (Yamada, 2011); many marine bacteriophages possess extraordinary structures (Huiskonen et al., 2004; Huiskonen et al., 2007).

One of the most interesting structures of cyanophage Syn5 is the horn appendage. Only one is present per particle and it is always attached on the vertex opposite the tail. It is long, flexible and can bend. Its length is about 30–35 nm. Prior to the cryoelectron microscopy of Syn5 only two bacteriophages were reported with appendages attached opposite the tail (Lake and Leonard, 1974). The presence of the horn raises many questions about its structure, function, and assembly. The purpose of the experiments described in this chapter was to answer some of these questions.
There are five novel proteins among the structural proteins of Syn5—gp53, gp54, gp55, gp57, and gp58. Two of these proteins (gp53 and gp54) are larger than the other three and are the primary candidates for the horn. Immunogold labeling of phage structures were performed for gp53 and gp54. Antibodies to gp58 were available and also tested to find out which phage structures they build. The antibodies raised against these proteins were used in neutralization assays to test the effects, if any, on the phage infectivity.

4.2 Materials and Methods

4.2.1 Immunogold labeling of Syn5 with anti-Syn5, anti-gp53, anti-gp54 and anti-gp58 antibodies

Ni grids (Formvar/Carbon coated, 200 mesh, Ted Pella) were washed 2 × 30 min in 100% acetone and then washed 2 × 20 min in double-distilled water (MilliQ) and rinsed additionally with drops of water before left to dry on blotting paper overnight. Syn5 at a titer of 2–3 × 10^{11} pfu/ml in phage buffer (50 mM Tris, pH 7.5, 100 mM NaCl and 100 mM MgCl_2) was used for all samples and 6 µl were applied per grid. The phages were allowed to bind for 10–15 min. The grids were edge blotted briefly on filter paper and incubated on the surface of a convex drop (98 µl) of 3% cold fish gelatin in phage buffer for 30 min at room temperature with gentle shaking on a Labnet shaker 20 (speed 200) to block non-specific binding. The solutions for each step were added to the wells of 96-well plates sealer, enclosed in a plastic box with a sheet of wet blotting paper at the bottom to minimize evaporation of the liquid during incubation. Grids were edge blotted briefly on filter paper and incubated on a floating drop of 0.3% gelatin in phage buffer
mixed with the primary antibody serum with the optimal dilution for each antibody (determined empirically)—anti-gp53 (1:2), anti-gp54 (1:2), anti-gp58 (1:2) and anti-Syn5 (1:100). As a negative control, grids carrying Syn5 were incubated with the pre-immunization serum of each antibody at a dilution of 1:2. Incubation with the primary antibodies was for 30 min at 37°C with gentle shaking (same as the blocking step). For the wash step, the grids were edge blotted briefly on filter paper and floated 3 × 10 min on a drop of phage buffer with 0.3% of gelatin, with no blotting between the wash steps but with blotting before the incubation with the secondary antibodies. The latter incubation was on a drop of 0.3% gelatin in phage buffer containing goat anti-rabbit IgG F(ab')2 (H+L) (AH) (10 nm gold, Ted Pella) at 1:100 dilution for 30 min at 37°C with gentle shaking as above. The wash step was repeated in buffer followed by a final wash step in MilliQ water for 15 min.

Each grid was floated on top of a drop of 1% of uranyl acetate for 30 sec, blotted and air-dried for about 30 min. The grids were observed under a JEOL 1200 transmission electron microscope (TEM) at 60 kV. Images were recorded with an Advanced Microscopy Techniques (AMT) XR41S side-mounted charge-coupled device (CCD) camera and saved as TIFF files.

4.2.2 Syn5 inactivation with anti-Syn5 serum and antibodies to novel Syn5 proteins

CsCl-purified Syn5 particles with a titer of 3 × 10^6 pfu/ml in phage buffer (50 mM Tris, 100 mM NaCl and MgCl₂) were used. The host cells used for plating were diluted with fresh medium the day before, incubation was continued overnight and used at density of about 3–4 × 10^8 cells/ml.
The anti-Syn5, anti-gp53, anti-gp54 and anti-gp58 sera were diluted in phage buffer at dilutions of 1:200, 1:2,000, 1:20,000 and 1:200,000, respectively.

Phage solution (0.1 ml) was added to 0.9 ml of each serum solution and incubated for 1 h at 37°C in a water bath. To stop the reactions 0.1 ml of the phage-serum mixture was added to 9.9 ml of room temperature phage buffer. Additional serial dilutions were prepared if necessary and appropriate volumes of each dilution were mixed with 1 ml of host cells and held for 1–2 min to allow for adsorption. About 3.5 ml of 32°C melted agar (1% LMP agarose in ASW) were added to each host-phage dilution and the contents of the entire tube poured into a Petri plate (51 mm). The plates were incubated overnight at 28°C in the light incubator at irradiation of about 40 µmol of photons per m² s⁻¹ and the plaque forming units were counted.

4.2.3 Sucrose gradients of the recombinant proteins gp53 and gp54

The purified proteins (100 µl of 0.3–0.4 mg/ml stocks) were loaded in a 5 ml tube with a 5–20% continuous sucrose gradient prepared in the protein storage buffer (25 mM Tris, pH 7.5, 100 mM NaCl and 2 mM EDTA). The gradient was run at 48,000 rpm for 6 h at 4°C in a SW 55 Ti Beckman rotor. Each gradient was fractionated into 20 fractions with a gradient fractionator (BioComp Instruments). Aliquots of each fraction were mixed with SDS loading dye and loaded on 12% SDS-PAGE. The gel was run for 2 h at 110 V, stained with Krypton (Pierce) and scanned on a Typhoon 9400. Gradient fractions of interest were collected and dialyzed to reduce the sucrose concentration. The samples were concentrated in Vivaspin concentrators with a 10,000 Da molecular weight cut-off (GE Healthcare) for gp53 and in Amicon concentrators with molecular weight cut-off of 30,000 Da (Millipore) for gp54.
For electron microscopy examination the proteins were negatively stained and observed using the protocol described in Chapter III.

4.3 Results

4.3.1 Immunogold localization of Syn5 particles with anti-gp53, anti-gp54, anti-gp58 and anti-Syn5 antibodies

Primary antibodies raised against intact infectious particles of Syn5 and individually against each one of the three novel proteins—gp53, gp54, gp58—were used in labeling experiments. The secondary antibodies were labeled with gold particles with size of 10 nm. To define any experimental background and false-positive labeling, the pre-immunization rabbit serum for each primary antibody was used as a primary antibody in labeling Syn5 virions (negative control). The data is summarized in Table 4–1. Only phage particles present in full in the field of the micrograph, and gold labels that were less than about 20 nm away from the phage surface, were included in the counts. This distance was chosen because it is the estimated maximum distance that a gold bead should be positioned when the lengths of connected primary and secondary antibodies are combined if the antigen is at the phage surface. When counting labeled viral particles for all experiments, only grid areas with low background labeling were chosen (not more than 5 beads attached to the grid surface on a 120,000 × image, Fig. 4–1). The results indicated background labeling of false positives at about 3.5%.
Table 4–1. Summary of the labeling results from the negative control.
The data represent a summary of the pre-immunization sera for all proteins.

<table>
<thead>
<tr>
<th>Primary antibody</th>
<th># of counted particles</th>
<th># of total gold labels</th>
<th>% of labeled particles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-immunization serum (negative control)</td>
<td>398</td>
<td>14</td>
<td>(14/398) 3.5%</td>
</tr>
</tbody>
</table>
Figure 4–1. Electron micrographs of the negative control. Syn5 labeling with pre-immunization serum of a rabbit before it was injected with gp58 (negative control). The results were similar to the other pre-immunization sera tested. Negatively stained with 1% uranyl acetate and observed under 120,000 x.
The novel proteins are not abundant in the phage particles (based on estimates from SDS protein gels) and the antibodies were raised against the recombinant proteins. These factors suggest that their labeling might not be very pronounced. Since the protocol used for the experiment was adapted from protocols used in other systems it was desirable to have a positive control to test the efficiency of the labeling procedure. The primary antibodies chosen for the positive control were anti-Syn5 antibodies raised in rabbits against the whole virion particles. They should allow for sufficient labeling of the phage particles and be a good test of the adapted protocol. As seen on Fig. 4-2 the labeling of the phage particles gave clearcut results. Most of the particles were associated with numerous gold beads distributed over the entire particle surface. The level of background binding of beads to the grid surface was low. The same protocol used for the positive and negative labeling was used for the experiments with primary antibodies against the novel proteins. The only variation was in the concentrations of the primary antibodies (1:100 for anti-Syn5 and 1:2 for the anti-novel proteins and for the negative control).

In the interpretation of the labeling data for each novel protein it was important to consider the distribution of gold particles on the surface of the virion. Only particles where the tail vertex could be identified were included in the results to eliminate ambiguity about the positions of the labels and to determine the horn position. To facilitate the summary of the data the phage particle vertices were designated (Fig. 4-3) and the labels on the counted particles were described based on where they were positioned.
Figure 4–2. Labeling of Syn5 virions with anti-Syn5 antibodies. The secondary antibodies are gold-labeled goat anti-rabbit antibodies. Grids were stained with 1% uranyl acetate and observed under 120,000 ×.
Figure 4–3. Schematic drawing of Syn5. To facilitate the summary of the gold beads positions, each vertex was assigned a number. Image courtesy of Cameron Haase-Pettingell.
The first of the novel proteins to be examined was gp53. Representative labeled particles are shown on Fig. 4–4 and the data is summarized in Table 4–2. The horn is invisible under negative staining but when observed by cryo-EM it appears flexible and at slightly varying positions—pointing straight away from the phage or tilted at the base and pointing towards the sides of the phage particle (vertex 3). Its length is about 30–35 nm as estimated by cryoelectron observations of Syn5 particles. If gp53 is a horn protein, the gold labels should be located at and around vertices 2 and 3 or pointing away from the horn. As seen in Table 4–2, about 90% of the labels are at vertices 2, 3 and between the two. About 7% of the labels are between vertices 3 and 4. About 4% of the labels are at other positions which is in good agreement with the percentages of false positive labels. In many particles, strings of gold beads were observed associated with vertex 2 (Fig. 4–4A), where the horn is presumed to be (in the cases of strings of beads only the ones close enough to the capsid were included in the counts). The data suggests that gp53 is forming the shaft of the horn or at least present along the length of the horn.

The labeling of phage particles with anti-gp54 primary antibodies was most often observed at and around vertex 3 (Fig. 4–5 and Table 4–3). About 77% of the labels were at vertex 3 and between vertices 3 and 4. About 11% of the gold beads were at vertex 2 and 7% between vertices 2 and 3. And, about 5% were at other positions, a proportion close to the results for false-positive labels. Strings of gold beads in the area of the presumed horn position were rarely observed and even in that case there were usually no more than 2 particles present. The labeling distribution for gp54 suggests that this is also a horn protein, perhaps forming the tip of the horn.
Figure 4–4. Labeling of Syn5 with anti-gp53 antibodies. A) Image of a typical labeling pattern with anti-gp53 antibodies, 120,000 x. B) Panel of selected labeled single particles. Black arrows point at the tails.
<table>
<thead>
<tr>
<th>Antibody</th>
<th># of counted particles</th>
<th># of total gold labels</th>
<th>% of labels on vertex 2</th>
<th>% of labels on vertex 3</th>
<th>% of labels between vertices 2 and 3</th>
<th>% of labels between vertices 3 and 4</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-gp53</td>
<td>126</td>
<td>146</td>
<td>~12</td>
<td>~38</td>
<td>~38</td>
<td>~7</td>
<td>~4</td>
</tr>
<tr>
<td></td>
<td>(18/146)</td>
<td>(56/146)</td>
<td>(56/146)</td>
<td>(10/146)</td>
<td>(6/146)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4–2. Summary with the positions of anti-gp53 gold labels on the Syn5 surface.
Figure 4–5. Labeling of Syn5 with anti-gp54 antibodies. 
A) Image of a typical labeling pattern with anti-gp54 antibodies, 120,000 ×. B) Panel of selected labeled single particles. Black arrows point at the tails.
Table 4–3. Summary of the positions of anti-gp54 gold labels on the Syn5 surface.

<table>
<thead>
<tr>
<th>Antibody</th>
<th># of counted particles</th>
<th># of total gold labels</th>
<th>% of labels on vertex 2</th>
<th>% of labels on vertex 3</th>
<th>% of labels between vertices 2 and 3</th>
<th>% of labels between vertices 3 and 4</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-gp54</td>
<td>126</td>
<td>152</td>
<td>~11</td>
<td>~44</td>
<td>~7</td>
<td>~33</td>
<td>~5</td>
</tr>
</tbody>
</table>

(16/152)  (67/152)  (11/152)  (50/152)  (8/152)
The labeling data for gp58 (Fig. 4–6 and table 4–4) suggests a different attachment pattern compared with gp53 and gp54. About half of the beads are distributed between all the pentamer vertices (where the hexamers are found) while the majority of the other half present at vertices 3 and 4. Since anti-gp58 antibodies do not recognize the coat protein (on Western blots) and the capsid knobs are not made of coat protein (cryo-EM data, personal communication Preeti Gipson), it is reasonable to conclude that the virion components labeled by the antibodies are the decoration knobs; hence gp58 is a knob protein.

4.3.2 Syn5 inactivation with anti-Syn5 and anti-novel protein antibodies

Once it was determined that gp53 and gp54 are horn proteins it was possible to begin to address the question of the horn function. The availability of antibodies against the horn proteins provided the opportunity to test their ability to reduce or block entirely the infectivity of Syn5. As a positive control, anti-Syn5 antibodies raised against the whole virion particles were used. The anti-Syn5 serum contains antibodies against the tail fiber protein (as evidenced by Western blots), which is predicted based on sequence homology with other phage proteins to participate in the host attachment phase in the infection process. It has been shown in other bacteriophages that serum raised against the whole virion has the ability to eliminate phage infectivity (Edgar, 1964; King and Wood, 1969).

As seen from the results on Fig. 4–7, the anti-Syn5 antibodies inactivated the Syn5 infectious particles, partially at a serum dilution of 1:20,000 and completely at dilutions of 1:2,000 and 1:200. Dilution of the antibodies to 1:200,000 was sufficiently high to no longer affect Syn5 infectivity.
Figure 4–6. Labeling of Syn5 with anti-gp58 antibodies.
A) Image of a typical labeling pattern with anti-gp58 antibodies, 120,000 ×. B) Panel of selected labeled single particles. Black arrows point at the tails.
Table 4–4. Summary of the positions of anti-gp58 gold labels on the Syn5 surface.

<table>
<thead>
<tr>
<th>Antibody</th>
<th># of counted particles</th>
<th># of total gold labels</th>
<th>% of labels on vertex 1</th>
<th>% of labels on vertex 2</th>
<th>% of labels on vertex 3</th>
<th>% of labels on vertex 4</th>
<th>% of labels between vertices</th>
</tr>
</thead>
</table>
Sera raised against each of the novel proteins (gp53, gp54, and gp58) did not lead to loss of phage infectivity at any of the tested dilutions—1:200, 1:2,000, 1:20,000 and 1:200,000. The titers of antibody-treated phage suspensions did not drop. This suggests that the horn structure is not needed for infection of WH8109.

4.3.3 Sucrose gradients of gp53 and gp54

Once designated as putative horn proteins, how gp53 and gp54 contribute to the assembly of the horn appendage could be approached. Their potential to form structural complexes was studied. In sucrose gradients, the two proteins sediment with quite different patterns (Fig. 4–8). The sedimentation coefficients of the proteins were estimated to be about 43 S for gp53 and about 6 S for gp54. Both proteins have some breakdown products which remain in the top fractions, but most of each protein sediments as a single peak.

Fractions of the gradients where protein was present were collected, the sucrose dialyzed, the protein concentrated, negatively stained on EM grids and observed under the electron microscope. Fractions 19–20 of gp53 and fractions 7–8 of gp54 showed protein structures (Fig. 4–9). It remains possible that the protein concentrations in the other samples were below detection. Very long and numerous fibrous structures were present in the gp53 sample. The fiber-like chains were significantly longer than those visible in the gp53 stock which was also visualized prior to the sucrose gradient step (not shown). In the case of gp54, the observed structures were small, short and rod-like, and quite similar to the structures observed in the gp54 stock protein preparation.
Figure 4–7. Inactivation of Syn5 infectivity.
Syn5 inactivation by anti-Syn5, anti-gp53, anti-gp54, and anti-gp58 antibodies.
Figure 4–8. SDS electrophoresis of sucrose gradients of gp53 and gp54. Gradients were 5–20%, run for 6 h at 48,000 rpm and separated into 20 fractions (f.1, fraction 1; f.20, fraction 20). The gels were stained with Krypton. M, molecular marker; c-l, the stock protein run as a control. 150,000 x.
Figure 4–9. Electron micrographs of gp53 and gp54. The images are of sucrose gradient fractions of gp53 (top) and gp54 (bottom). Both samples were negatively stained with 1% uranyl acetate and observed at 150,000 x. Black arrows point at the protein structures.
4.4 Discussion

The horn structure is one of the most interesting morphological features of cyanophage Syn5. The results from the labeling experiments described in this chapter strongly suggest that gp53 and gp54 are both components of the horn structure. Estimates of the intensities of the two proteins upon SDS-PAGE fractionation of Syn5 infectious particles, show that there are about 12 copies of gp53 and 6 copies of gp54 per particle. With anti-gp53 antibodies, it quite often appeared that the gold beads were tracing the invisible horn structure seen in cryo-EM images. The long fibers formed by recombinant gp53, both in purified stocks and sucrose gradients, suggest that this protein may have the ability to polymerize and form elongated structures. The longer fiber-like structures in the sucrose gradient samples may be due to separation of the short chains from the longer ones. It is possible that the sucrose present in the buffer intensified the ability of gp53 to form such structures.

The labeling with anti-gp54 antibodies was more pronounced in areas of phage particles where the tip of the horn is expected to be positioned. This suggests that gp54 may be building the tip of the horn. It is also possible that the horn is a double-layered structure with gp53 on the outside and gp54 on the inside and at the tip.

The distribution of the gold beads in the labeling with anti-gp58 antibodies fits well the profile of the distribution of the knob structures. Gp58 is most likely a knob protein. Since there are two additional proteins (gp55 and gp57) present in the protein band of gp58 in Syn5 particles run on SDS-PAGE, it is quite possible they too contribute to building the knobs. Each Syn5 capsid hexamer carries three knob-like protrusions.
One interesting observation from the immuno-labeling experiments was that very good levels of labeling were achieved with primary anti-Syn5 antibody concentrations as low as 1:1,000, while for the novel proteins the concentration had to be much higher, at least about 1:5. That is most likely because each novel protein is present in just a few copies per particle. The anti-Syn5 serum contains antibodies raised against the exposed native proteins of phage particles while the antibodies to the novel phage proteins were raised against the recombinant versions of these proteins. Potentially, the conformations of the native and recombinant proteins may differ. It is also possible that some of the antigenic epitopes are not exposed when the proteins are part of the phage structure.

The phage binding experiment did not reveal any inhibition of phage infectivity by the anti-novel proteins antibodies. One explanation is that none of the structures the novel proteins build are involved in host recognition or attachment phases of the infection process. This is not unexpected for gp58 since it is positioned on the phage head and is more likely involved in building the hexamer knobs. Bacteriophages infect their hosts via tail appendages. It is more striking that the binding of the horn proteins (gp53 and gp54) with antibodies did not have an effect on the ability of Syn5 to infect its host.

One hypothesis for horn function is that it is involved in host recognition. This is the case for the *Caulobacter* phages Cb13 and CbK, which have a long head appendage on the vertex opposite the tail and wrap it tightly around the host flagellum (Guerrero-Ferreira *et al.*, 2011). Other phages that use their tails to bind to the host flagella or pili have also been described in the literature (Schade *et al.*, 1967; Lotz *et al.*, 1977; Wilson and Takahashi, 1978; Scholl and Jollick, 1980).
The horn of Syn5 may have a different function. Alternatively, it may be involved in host attachment but to a host other than *Synechococcus* sp. WH8109. On the other hand, WH8109 is the only known lytic host identified to date for Syn5; about 18 cyanobacterial species have failed to exhibit lysis triggered by Syn5. *Synechococcus* sp. WH8109 is non-motile and in general, motile *Synechococcus* cyanobacteria do not possess flagella or pili (Brahamsha, 1999). Novel putative swimming structures were described in WH8113 (Samuel et al., 2001).

It is also possible that the horn is used for attachment, not to the bacterial host, but to other particulates in the water, which might keep the phage in nutrient-rich waters and hence closer to its host(s). Finally, it remains possible that the horn does assist Syn5 in the attachment steps but since the antibodies utilized in the phage binding experiments were raised against recombinant proteins, and not against the horn appendages in their native form, the antibodies may not be as efficient in binding to the horn epitopes of the mature phage responsible for attaching to the bacterial host.

The results from the analysis of gp53 and gp54 in sucrose gradients and under EM clearly indicate that the proteins are not in their monomer form. Their sedimentation coefficients and the presence of visible structures under EM, strongly suggest that both proteins form multimers or complexes likely needed to build and shape the horn appendage of Syn5. Such structures are microtubules, which are made of polymerized dimers of tubulin, and actin, which has two forms—globular monomer and filamentous polymer. They can also build extended structures like collagen fibers or the tail fibers of bacteriophage T4.
CHAPTER V: STUDIES OF SYN5 IN VIVO PROCAPSIDS
5.1 Introduction

The study of viral assembly is crucial to understanding how viruses function and to devising approaches to control their propagation. In general these studies were first undertaken in bacteriophages, which were then used as model systems for animal viruses. The majority of this work has been performed with enteric bacteriophages since they are readily isolated, widely produced under laboratory conditions and already extensively studied. The common assembly pathway which has been elucidated consists first of the assembly of an intermediate protein shell (procapsid or prohead). The nucleic acid is pumped through a portal vertex into the protein shell and then the host attachment proteins are added (tail apparatus).

Prior to the present work, no assembly studies had been conducted in the cyanophages. Among the well-studied enteric phages, phage T7 appears to most closely resemble cyanophage Syn5. In E. coli, T7 first assembles first into procapsid consisting of coat, portal, scaffolding and the three internal virion proteins. Protein shells with no DNA and containing all of the above proteins except the scaffolding have been isolated as well (Studier, 1972; Serwer, 1976; Roeder and Sadowski, 1977). When co-expressed together, the coat and scaffolding proteins of T7 assemble shells. The portal was is not incorporated into these shells when it is co-expressed with the coat and scaffolding proteins (Cerritelli and Studier, 1996a).

The recombinant putative scaffolding protein of Syn5 has biochemical properties very similar to the scaffolding of T7. To define its role in Syn5 assembly, we first had to
establish that the Syn5 scaffolding protein is expressed during wild-type infection. If so, does Syn5 assemble through a procapsid step?

5.2 Materials and Methods

5.2.1 Expression of the Syn5 scaffolding protein during infection

To test whether the scaffolding protein (gp38) of Syn5 is expressed during infection cell samples were processed during infection. About 50 ml of cells were collected at different time points after infection, concentrated 100-fold by resuspension in 0.5 ml of lysis buffer (50 mM Tris, pH 8, 100 mM NaCl) at room temperature and frozen at -20°C. The cells were thawed and treated with lysozyme (2 mg/ml) for 30 min at room temperature, sonicated for 5 × 15 sec, and treated with DNase I (1 U/ml) for 1 h at room temperature. The cell debris was pelleted at 10,000 × g for 10 min, the supernatants were decanted and retained, and the pellets were resuspended in the same volume of resuspension buffer. Pellets and supernatants were diluted 1:2 and the proteins separated by SDS-PAGE (12%) at 100 V for 2 h after boiling in reducing buffer (60 mM Tris, pH 6.8, 2% SDS, 5% β-mercaptoethanol, 10% glycerol, bromophenol blue for color) at 80–85°C for 10 min. Gels were Krypton-stained (Pierce Biotechnology) and visualized on a Typhoon 9400 scanner. For Western blots, the gels were equilibrated in transfer buffer (10% ethanol, 25 mM Tris base, and 192 mM glycine) and the proteins transferred to a polyvinylidene fluoride (PVDF) membrane (0.45 μm, Millipore) overnight in the same transfer buffer at 15 V constant voltage and 4°C (Bio-Rad Criterion transfer apparatus). The membranes were washed in PBS-Tween 20 (0.1% v/v) and blocked with 5% Carnation nonfat dry milk in PBS (phosphate buffered saline, pH 7.5) overnight at 4°C.
The blot probing steps were as described in the enhanced chemifluorescence (ECF) Western blot manual (Amersham). A 1:1,666 dilution of the anti-scaffolding polyclonal antibodies (Strategic Diagnostics) was used to probe the membranes. The secondary antibodies used were alkaline phosphatase (AP) conjugated goat anti-rabbit antibodies (Bio-Rad) at dilution of 1:3,000. The membranes were incubated for 1 h at 4°C with the primary antibodies, washed and incubated for 45 min at room temperature with the secondary antibodies. For visualization they were soaked for 10 min in ECF substrate (Amersham) and scanned on a Typhoon 9400 (GE Life Sciences).

5.2.2 Purification of procapsids

A *Synechococcus* sp. WH8109 culture (1.4 L) at $6 \times 10^8$ cells/ml in ASW was infected with Syn5 phages at MOI = 2–3. Samples were harvested at different times from 20 to 90 min after infection. Optimal procapsid yield was determined to be 40–50 min after infection. Samples were collected in pre-chilled bottles and held on ice for 10 min. The cells from 1 L of culture were pelleted at $7,200 \times g$ for 10 min, and the pellet was resuspended in 4 ml of 50 mM Tris, 100 mM NaCl. The lysis buffer also included protease inhibitors (1 tablet per 10 ml Complete Mini tablets, EDTA-free, Roche). Before infection a sample of uninfected cells was collected and pelleted at $9,500 \times g$. All samples were frozen at -20°C.

The cell pellets were thawed and lysed for 40 min with lysozyme (2 mg/ml) at room temperature, treated with DNase I (Worthington, 2 U/ml) with gentle mixing on a rocker for 1 h at room temperature. The cell debris was pelleted at $20,000 \times g$ for 10 min at 4°C and the supernatants were loaded onto a CsCl step gradient. To separate procapsids from mature phages, the following CsCl layers were found to be optimal for 1.5 ml of
250-fold concentrated samples in 17.5 ml centrifuge tubes (Beckman): 2 ml of $\rho = 1.2$, 3.5 ml of $\rho = 1.25$, 4 ml of $\rho = 1.3$, 4 ml of $\rho = 1.4$, 1 ml of $\rho = 1.5$, and 1 ml of $\rho = 1.6$. The gradient was spun at 28,000 rpm (Beckman SW28 rotor) for 3 h at 8°C. Procapsids migrated as an opalescent band at the interface of the 1.25 and 1.3 layers whereas the Syn5 DNA-filled particles migrated to the interface of the 1.4 and 1.5 layers.

The initial CsCl-purified samples were contaminated with outer membrane vesicles. These contained host porins as determined by mass spectrometry analysis. In early experiments the vesicles did not separate well from the procapsids on sucrose gradients, and their amounts varied substantially. Mild treatments to lyse infected cells, such as using lysozyme alone, and most importantly, eliminating the sonication step, decreased the concentration of vesicles.

The procapsid band was harvested with a needle and dialyzed against phage buffer (50 mM Tris, pH 7.5, 100 mM MgCl$_2$) with stepwise decrease of NaCl concentration—2 M NaCl (1 h), 1 M NaCl (1 h), and 100 mM NaCl (overnight at 4°C). The procapsids were further purified on a 5–20% sucrose gradient at 42,000 rpm (Beckman, SW55 Ti rotor) for 1 h at 8°C. Twenty fractions were collected using a gradient fractionator (BioComp Instruments).

5.2.3 Electron microscopy

Phage and procapsid samples were blot dried on glow-discharged Formvar/Carbon coated copper grids (Ted Pella) and negatively stained with 1% uranyl acetate. The grids were observed under a JEOL 1200 transmission electron microscope (TEM) at 60 kV. Images were recorded with an Advanced Microscopy Techniques
(AMT) XR41S side-mounted charge-coupled device (CCD) camera and saved as TIFF files.

5.3 Results

5.3.1 Syn5 assembly involves a scaffolding protein

To track whether Syn5 scaffolding protein is expressed during the infection process, polyclonal antibodies were raised against the recombinant gp38 and used to probe the time course of Syn5 infection (Fig. 5–1). Controls for the specificity of the antibodies were uninfected WH8109 cells, the recombinant protein itself and Syn5 infectious particles. No bands corresponding to the scaffolding protein were detected in the uninfected WH8109 supernatants, at 0.5 minutes after Syn5 infection nor in the mature Syn5 virions (Fig. 5–1). The polyclonal serum recognized a few host proteins weakly, but they were predominantly in the pellet fraction, and of higher molecular mass than the scaffolding protein. In the recombinant protein control lane, there were two protein products—one with the size of 40 kDa and another at about 24 kDa, possibly a degradation product due to prolonged storage.
Figure 5–1. The Syn5 scaffolding (gp38) in Syn5 infected cell lysates of WH8109. Western blot of time course of Syn5 infected host cells, probed with anti-scaffolding antibodies. gp38, recombinant scaffolding protein, 40-kDa and 24-kDa bands; 8109, uninfected *Synechococcus* cells; host prot., host proteins recognized by the anti-scaffolding antibodies; sup, supernatant fraction; pel, pellet fraction.
The anti-gp38 antibodies recognized five protein bands predominantly in the supernatant of the infected cells (Fig. 5–1). Their molecular masses range from 40 to 13 kDa. The 40-kDa band together with the 29 and the 15-kDa products were detected as early as 20 min after infection, peaking at 60 min and decreasing at later times. In addition, the other two bands, the 23 and the 13-kDa, appeared later (40 min) and persisted longer. The lower molecular mass bands are likely to be degradation products of the 40-kDa scaffolding protein. This may be caused by the scaffold sensitivity to proteases resistant to inhibitors used in the experiment. Well-studied scaffolding proteins in other bacteriophages have been found to be highly protease sensitive (Dokland, 1999).

5.3.2 Do Syn5 virions assemble through a procapsid intermediate?

The evidence for scaffolding protein expression during Syn5 infection implies that the virus proceeds through an assembly pathway involving procapsids empty of DNA, similar to the dsDNA enteric bacteriophages such as P22, T7, and T4. If scaffolding-associated procapsids are precursors of Syn5 virions, they should form early in the infectious cycle.

In an effort to isolate procapsids, infected host cells were sampled at 30, 40, 50, 60 and 70 min after infection and lysed. To distinguish capsids lacking DNA from DNA-filled virions the lysate supernatants were fractionated by CsCl density centrifugation. In these gradients, a band representing DNA-containing particles was visible in the $\rho = 1.4–1.5$ layers (same as Syn5 infectious particles purified after host lysis), and a second, bluish-white scattering band was present in the $\rho = 1.25$ layer. The two bands were clearly resolved. The intensities of both bands increased in later time samples. The time chosen as optimal for procapsid purification was 50 min after infection at 28°C. While the
intensity of the $\rho = 1.25$ band was stronger at later times (60–70 min), SDS gels of these particles indicated significant contamination with ghosts (phages which have ejected or lost their DNA). Earlier times (30–40 min) yielded very faint bands in the $\rho = 1.25$ layer.

To further separate capsids from cell structures, the $\rho = 1.25$ band was recovered and fractionated in a sucrose density gradient (5–20%). The fractions were separated by electrophoresis through SDS-PAGE and analyzed by Western blots to assess the presence of mature virion proteins and the scaffolding protein (Fig. 5–2).

The blot probed with anti-Syn5 antibodies (Fig. 5–2) revealed two rapidly sedimenting populations of coat protein structures at 240 S and 340 S. The results also suggest that the portal might be present in both peaks, since a few faint bands are seen closer to the expected size of the portal. However, none of them aligns perfectly with it, so further investigation is needed. The presence of the portal is to be expected if the Syn5 assembly pathway is similar to that of other dsDNA phages. The Western analysis showed that the structures in both peaks lacked the tail apparatus proteins, as expected. It was not possible to reliably identify whether the other minor internal proteins (gp45, gp44, and gp43) were present in the procapsids.
Figure 5–2. Western blot of sucrose gradients of Syn5 procapsid particles. The CsCl purified procapsid band was run on sucrose gradients (5–20%, fractions 1–20), electrophoreosed on SDS-PAGE and transferred to a Western membrane. A) Membrane probed with anti-Syn5 antibodies. Syn5, purified infectious particles (control). Fractions 10–12 contain the 240 S peak, fractions 15–17 the 340 S peak. B) Membrane probed with anti-scaffolding serum. Scaff. 1, 40-kDa scaffolding protein band; scaff. 2, faint 29-kDa scaffolding band; M, protein standards.
Figure 5–2B shows the blot probed with anti-scaffolding serum. The 40-kDa scaffolding species was clearly present in both peaks. It co-sedimented with the coat protein in both fractions (240 S and 340 S). In addition, there was a very faint 29-kDa scaffolding species. The data indicated that the scaffolding protein was associated with the coat in rapidly sedimenting large complexes. The fractions containing both types of particles (240 S and 340 S) were concentrated after sucrose dialysis, negatively stained with uranyl acetate and examined under transmission electron microscopy (TEM). In both peaks, the images revealed spherical particles about 50 nm in diameter without tail-like appendages (Fig. 5–3). Most were internally electron-dense due, presumably, to the presence of the scaffolding protein and will be referred to as procapsids (Fig. 5–3D). In some of them the density filled the shells completely, while in others it was lining the inner capsid surface leaving a small empty area in the middle of the shell. Empty capsids lacking internal density and appearing more collapsed, instead of spherical, were present as well (Fig. 5–3E).

The particles in both peaks were very similar in appearance. The 240 S fraction (Fig. 5–3B) contained about 70% filled procapsids with regular spherical shape containing internal protein densities, presumably representing the scaffolding protein. About 30% were stain-permeable empty capsids (Table 5–1). These might be procapsids, which have lost the scaffold within the cells or during purification (King and Casjens, 1974). The 240 S fraction also contained a background of stain-excluding 5–10 nm structures. We have not yet been able to identify the composition of this material. It may be derived from the extensive cyanobacterial host membrane system.
Figure 5–3. EM images of negatively stained procapsids.
A) Micrograph of particles in the 240 S sucrose gradient fraction showing procapsid particles (pc, solid arrows) filled with protein density (most likely the scaffolding protein, making them impermeable to stain), and empty stain-permeable capsids. B) Procapsid particles (pc, solid arrows) in the 340 S fraction were similar to those in the 240 S peak with diameters of about 50 nm. Empty shells were more rarely observed. C) Syn5 mature virions for size and shape comparison. D) Panel of enlarged procapsids-like particles found in both peaks. The arrows indicate areas of higher densities at the coat wall. E) Enlarged empty capsids present in both fractions. F) Icosahedral capsids found in the 340 S peak. All samples were negatively stained with 1% uranyl acetate and observed at magnification of 120,000 ×.
Table 5–1. Summary of the species of Syn5 procapsid particles.
Particles were counted from the electron micrographs of each peak and grouped based on their appearance. Fig. 5–3 contains images of these species.

<table>
<thead>
<tr>
<th>Fractions</th>
<th>Particles species</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>240 S</td>
<td>Procapsids</td>
<td>104</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Empty capsids</td>
<td>46</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Icosahedral capsids</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>340 S</td>
<td>Procapsids</td>
<td>127</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>Empty capsids</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Icosahedral capsids</td>
<td>8</td>
<td>5</td>
</tr>
</tbody>
</table>

Total counted 150
About 85% of the capsids in fraction 340 S (Fig. 5–3B) appeared to be identical to the procapsids in the 240 S peak (Table 5–1). About 10% resembled empty capsids (Fig. 5–3E). Particles with more angular shape (less than 5%) were also present (Fig. 5–3F). Although somewhat icosahedral, they lacked the tightly packed structure of the mature virion and did not exhibit tails.

In some of the procapsid particles, a localized feature of stain-excluding density in the coat shell was visible (Fig. 5–3D). This might be the core, composed of the portal and possibly the internal virion proteins. Since Syn5 has a horn structure, this density could also be the base of the horn. It could also be the scaffolding leaving the procapsids.

Critical factors for the success of the procapsid purification protocol were the elimination of Triton X-100 in the lysis buffer and lowered NaCl concentration. In the presence of detergent, empty shells of the coat protein were abundant as seen by EM and no scaffolding protein was present in sucrose gradient fractions with these particles. Thus in the presence of 0.1% Triton X with 300 mM NaCl the scaffold was apparently released from the procapsids. When Triton X was eliminated from the lysis buffer and the NaCl concentration was lowered to 100 mM, the particles described above were obtained.

As noted above, the isolated procapsid shells were smaller in size to the 60 nm infectious mature Syn5 virions (Fig. 5–3C). This suggests that Syn5 virions expand after DNA packaging as in T7, P22, and λ (Earnshaw and Casjens, 1980).
5.4 Discussion

A distinctive feature of the intracellular assembly of dsDNA phages and viruses is the formation of a procapsid shell with the assistance of a scaffolding protein and containing a portal vertex, for DNA packaging and ejection (Chang et al., 2006; Fu and Prevelige, 2009). Upon DNA packaging into the procapsids, scaffold exits from the procapsids, or is proteolysed, and replaced by tightly packed genomic DNA in an expanded mature capsid (Casjens, 1987; Tuma et al., 1998).

For dsDNA phages and viruses the exit of the scaffolding subunits and the entry of the DNA is a complex process. This process is generally disrupted by cell lysis resulting in particles, which have lost the scaffolding protein but have not yet been stably filled with DNA. These species often have different sedimentation coefficients from the precursor procapsids. This was shown in the analysis of P22 where 240 S precursor procapsids are resolved from 170 S empty capsids, which are derived from intermediates in DNA packaging (King et al., 1973). We presume that the two peaks of Syn5 particles (240 S and 340 S) observed here might be the precursor and a stage further along within the intracellular DNA packaging process. By analogy with P22 it seems that the 240 S peak is the break down product and the 340 S fraction is the precursor.

The classes of particles shown in Fig. 5–3E, lacking scaffolding and tightly packed DNA, closely resemble particles observed in lysates of P22 infected cells. Using pulse chase experiments, together with mutants blocking late steps in DNA packaging and particle maturation, the P22 particles were shown to be derived from procapsids, which have initiated DNA packaging but have not completed it at the time of lysis (King et al., 1973; Strauss and King, 1984). The Syn5 icosahedral particles (Fig. 5–3F) were probably derived from capsids that had packaged DNA in vivo, with the assistance of the
Syn5 terminase (gp60), and gone through the procapsid to capsid lattice transition, but had not been stabilized by the addition of the tail proteins that close the portal.

The early appearance of scaffolding-containing procapsids in infected cells and the absence of the scaffolding protein in mature virions indicate that procapsids are precursors to the mature virion in the dsDNA cyanophage Syn5 and that it therefore possesses a procapsid assembly pathway similar to the one described for the dsDNA bacteriophages (most of them enteric) (Casjens, 1987). Given the high frequency of lateral gene transfer among dsDNA phages (Canchaya et al., 2003; Hambly and Suttle, 2005), it is possible that although the evolution of the host cyanobacterium precedes the emergence of enteric bacteria, the phages may have evolved in enterics and then spread in more recent times to cyanobacteria. However, since cyanobacteria precede enteric bacteria in evolution, it is reasonable to propose that the enteric bacteriophages inherited and preserved the assembly pathway originally evolved in the cyanophages.
CONCLUDING REMARKS
Studies on cyanophages infecting single-celled cyanobacteria have been expanding rapidly in the past 20 years. However, in the majority of these reports the focus is exclusively on the ecology of these phages and the impact they have on marine environments. Detailed information on the processes occurring within host cells during cyanophage infection is critical for expanding and deepening our knowledge about these important phage-host interactions.

This thesis marks the beginning of studies on cyanophage assembly. Such studies are more challenging in cyanophages than in enteric virus systems. This is mainly due to the long life cycles of the host organisms and to the low burst sizes of viruses that infect cyanobacteria. Another obstacle is the lack of readily available techniques to perform genetic manipulations on the hosts and phages.

A striking result of the detailed molecular investigation of Syn5 and its assembly is the remarkable similarity between Syn5 and dsDNA enteric phages such as T7 and P22. The genome sequence of Syn5 shows high homology in gene sequence as well as gene order with the T7 genome; the proteins of the two phages have similar characteristics as well. The two scaffolding proteins behave similarly in their recombinant forms. The same applies to the coat and portal proteins. The only discerned difference is the inability of the Syn5 coat and scaffolding to assemble procapsid-like particles when the two proteins are co-expressed in E. coli. However, the ability of bacteriophage T7 to assemble such particles might be possible because these proteins were expressed in T7’s natural host, E. coli, while this is not true for Syn5. An unrecognized host factor may be critical for this assembly to occur. T7 also possesses two
types of procapsid particles—one with and one without the scaffolding protein. Both
types can be purified from cell lysates after an infection.

These significant similarities between Syn5 and T7 lead to questions about the
evolution of cyanophages. There are at least two hypotheses. One states that the
cyanophages are ancestors of the enteric phages since the cyanobacteria evolved earlier
than the bacteria of the mammalian gut. The other idea suggests that both groups may
share the same ancestor.

The major molecular differences between Syn5 and T7 arise from the presence of
a cluster of five genes at the “righthand” end of the Syn5 genome. All of these genes
encode novel proteins. The two largest of these (gp53 and gp54) are apparently involved
in building the interesting Syn5 horn structure. Possibly similar structures have been
found in only a few other bacteriophages. For two of them (φCb13 and φCbK) it has been
recently shown that they use their head filaments to attach to their host. No data is
available about the function of the tassle of cyanophage Bellamy (personal
communication with Welkin Pope and Roger Hendrix). Although the major hypothesis
for the Syn5 horn is that it facilitates attachment and infection of the host, the data from
the infectivity experiment reported here gives no support to this hypothesis.

*Synechococcus* sp. WH8109 is a non-motile strain and no flagella, pili or spicules
have been identified via EM tomography (Wei Dai, personal communication). It is
possible that the Syn5 horn has an unrelated function or that, although the current data
strongly suggest that it is not needed for the infection of WH8109, the horn may be
required to infect other hosts. Experiments by others aimed at identifying other lytic hosts
of Syn5 have not identified any. However, these tests did not look for possible lysogenic
hosts. Syn5 possesses an integrase gene, which raises the possibility that it may have a lysogenic life cycle in other hosts.

All phages, described so far, with structures located on the vertex opposite the tail complex are phages that infect aquatic prokaryotes. This definitely raises the possibility that such structures may not be an exception in hosts inhabiting dilute water environments. Encounters between phages and their hosts probably happen relatively rarely in these ecosystems and the phages may have evolved structures to facilitate the process. It is also plausible that the phages use these structures to attach to other organisms or to debris to reside closer to nutrient- and host-rich areas.

There is a considerable interest in cyanobacteria in the field of green energy technologies but a critical requirement is to use genetically engineered strains. The detailed studies reported here about the Syn5 life cycle and assembly can be used as a basis to develop this phage as a tool to genetically manipulate the host. It may be possible to develop an in vitro DNA packaging system in which modified or exogenous DNA is packaged into preformed procapsids, and then converted to mature infectious phage, as was originally done with phage lambda. The existence of such technique may not only answer questions about the biology and impact of Syn5 but may also allow metabolic engineering of the host.
APPENDIX
Objectives

The experiments in the appendix were performed in an attempt to answer the question whether these procapsid-like particles are indeed part of the common assembly pathway or defective particles that were insufficient in the assembly process. If these particles were indeed procapsids it was also important to find out which are the proteins built them.

Materials and Methods

Radioactive labeling of phage proteins during wild-type infection

Host cells with density between $2 \times 10^8$ cells/ml were diluted with fresh medium 24–48 h before the experiment. For each experiment, a culture was split into two flasks and each flask infected. One culture was incubated without labeling. The other was radioactively labeled with $[^{35}\text{S}]$ Met–Cys (20 µCi/ml Perkin Elmer /NEN) or a mixture of $[^{35}\text{S}]$ Met–Cys and $[^{14}\text{C}]$-labeled amino acids (Perkin Elmer /NEN, 10 µCi/ml of the former and 100 µCi/ml of the latter). The radioactive amino acids were added at 25 min after infection, labeled for 10 min and chased at 35 min. The sample labeled only with $[^{35}\text{S}]$ Met–Cys was chased with 10 mM Met and 1 mM Cys, while the sample with mixed labeling was chased with Amino Acid Mixture (Sigma, RPMI-1640 Amino Acid mix supplemented with 20mM L-Met, 20 mM L-Cys and 47 mM L-Ala; 40 ml per 600 ml culture). When casamino acids were used instead of the the Sigma Amino Acid mix, the scaffolding protein was not present in the samples, both labeled with radioactive amino acids and the unlabeled control. The amount of culture used depended on the experiment.
and ranged from 5 to 1000 ml infected with MOI = 3–5. The experiments were conducted in a 28°C shaking (approximately 200 rpm) water bath under the illumination of a cool white fluorescent lamp with irradiation of about 30–40 µmol m⁻² s⁻¹ at the water surface. Since the experiment was performed in a water bath and at lower irradiation compared with the one-step growth curve experiment (see Chapter 2) the life cycle of Syn5 was altered, with an eclipse period of about 53 min and a latent period of about 80 min. In the course of infection aliquots were harvested at different time points with the final one collected at 90 min.

After collection each aliquot was kept on ice for 10 min to slow the assembly process and then centrifuged at 9,000 × g for 10 min at 4°C. The cell pellet was resuspended in 100 mM NaCl and 50 mM Tris pH 7.5 containing Roche Protease Inhibitor (Mini tablets, EDTA free, 1 tablet per 7 ml buffer) with concentration factor between 500–1000-fold. Pellets were frozen at -20°C.

The samples were thawed and lysed in the presence of 2 mg/ml lysozyme for 30 min at room temperature. DNaseI (Worthington) was added at a final concentration of 1 µg/µl and the sample followed by mixing for 30 min at room temperature. The supernatant was separated from the debris by spinning at 10,000 × g for 10 min at 4°C.

Sucrose gradients (5–20%) were prepared, run at 42,000 rpm (SW55 rotor, Beckman) for 45 min and collected under the same conditions described above for procapsid purification. In some cases 200 µl of a 60% sucrose shelf was added to the bottom of the tube to prevent the mature virions from sticking to the bottom. From each fraction, 30 µl were mixed with 15 µl of 3-fold concentrated loading dye and boiled for 9 min at 95°C. For each time point of the unlabeled samples, two 12% SDS gels were run
in parallel—one for Krypton staining and one for Western blotting. For each sample, 12 µl were loaded per well. The SDS gels were run for 2 h at 105–110 V and the Westerns were performed as described earlier in this section (see Purification of Procapsids).

Results

Labeling of Syn5 infected cells with radioactive mixture of Met and Cys amino acids

In the process of purification of Syn5 particles from infected host cells the population of particles obtained is quite often heterogenous. During the cell lysis step any infecting particles attached to the cells are released from the cell debris, resuspended in lysate and remain in the supernatant fraction after separation of the debris. Some of these particles still contain some or a whole complement of DNA while others are empty. In the density gradients purification steps, particles that are not completely emptied of DNA will co-purify with the procapsids due to their similar density. This makes the task of determining the protein composition of the intermediate particles solely by purification methods almost impossible because purified particles will always include some contaminating Syn5 particles used to infect the cells.

One way to separate newly synthesized particles from the ones used in the infection is to preform a pulse-chase experiment adding radioactively labeled amino acids during the infection.

First, a test experiment was performed to determine the labeling efficiency (Fig. A–1). The results indicate that the labeled proteins match the profile of the structural proteins of Syn5. It is relatively easy to identify the coat protein and the proteins with higher molecular mass than the coat. The identification of the proteins smaller than the
coat based on molecular mass alone is somewhat inconclusive. There are more protein bands present than in mature virions. As seen in the panel on the right in Fig. A–1, the scaffolding protein breaks down into many small products. It is possible that some or all of the smaller bands in the radioactive gel are scaffolding break down products. Alternatively these could include break down products of the rest of the Syn5 stuctural proteins.

**The coat protein migrates gradually from a free soluble form to the mature virion over the time course of a Syn5 infection.**

To confirm that the particles identified as procapsids are indeed structures in the assembly pathway, and not defective particles accumulating off pathway, radioactive labeling was used to monitor the transition of the coat from free soluble protein through procapsid intermediate to mature Syn5 virions.

The time course of labeled amino acid incorporation during phage infection was determined. As seen in Fig. A–2, the coat protein is present to some degree at all time points: in the top fractions in its free soluble form, in the middle of the gradient where the procapsids sediment under the chosen conditions, as well as in the bottom of the gradient where the Syn5 mature virions sediment. However, the coat band in the top fractions (free form) is most pronounced at 36 min after infection and very faint at 65 and 90 min. In the middle (procapsid) fractions of the gradient coat protein increases as the infection progresses with time.
Figure A–1. Comparison (SDS-PAGE) of the proteins present in the mature purified Syn5 virion (left panel) and Syn5 infected cells (middle panel) after pulsing with a mixture of $[^{35}\text{S}]\text{Met–Cys}$ (pulsed for 10 min between 25 and 35 min after infection). Western blots (right panel) of infected Syn5 cells with anti-scaffolding antibodies for scaffolding visualization comparison.
A similar pattern is observed for the coat distribution in the bottom fraction where the mature virions sediment. The coat is present in the bottom fractions as early as 36 min after infection, but at very low intensity, signaling the presence of the first newly synthesized infectious particles. But at 90 min the coat band at the bottom is very intense and most of the other Syn5 structural proteins can be identified as well. Overall these results reflect the path of coat protein from soluble form to procapsid to mature virion.

**Groups of Syn5 transitory structures and the proteins that build them**

Radioactive labeling was applied to identify the proteins present in newly formed Syn5 intermediate particles or procapsids. As described previously (Chapter V), when procapsids were isolated from infected cells two major groups of particles were present. One containing protein densities, most likely the scaffolding protein, the other group empty collapsed shells (Fig. 5–3).

Syn5-infected WH8109 cells were labeled with a mixture of $[^{14}\text{C}]$-labeled amino acids. The cells were harvested at 70 min after infection, run on sucrose gradients (5-20%), the gradients fractionated and separated by SDS-PAGE (Fig. A–3). The results indicate clearly that the coat and scaffolding proteins did not peak in the same fractions in the middle of the gradient where the procapsids are expected to be present. A similar peak offset was present in procapsids purified from an infection performed under the same conditions without the addition of labeled amino acids (Fig. A–4).
Figure A–2. SDS-PAGE of fractionated sucrose gradients of supernatants of lysed cells labeled with [\(^{35}\)S] Met–Cys during Syn5 infection.
Figure A–3. SDS-PAGE of fractionated sucrose gradients of supernatants of lysed cells labeled with a mixture of $[^{14}\text{C}]$ amino acids during Syn5 infection (25–35 min after infection). The cells were concentrated 1,000-fold before lysis. The gel was exposed for imaging for 22 days. Lys, whole cell lysate; int., internal protein; scaff., scaffolding protein.
For the coat protein, the peak in intensity is in fractions 8 through 11 while the scaffolding is present with higher intensity in fractions 10–14, with the peak in fraction 12. The intensity of the scaffolding protein signal is low in the radioactive gels. The other phage proteins clearly present in the same fractions as coat and scaffolding are the two major internal proteins gp44 and gp45, the portal and the three small proteins gp55, gp57 and gp58. All of these proteins peak in fraction 10. The broad peak of the three small proteins coincides very well with the coat—fractions 10 through 14. These data suggest that the three small proteins may not be present in the scaffolding-containing particles.

To perform Western blots and EM analysis of the samples, an infection was performed under the same conditions as the labeling experiment but without radioactive labeling. The cells were harvested at 70 min after infection, lysed and the supernatant run in a 5–20% sucrose gradient. The gradient was separated into twenty fractions, which were separated by SDS-PAGE, transferred to a membrane and probed with anti-Syn5 and anti-scaffolding antibodies. The results (Fig. A–4) show the same lack of overlap in the peaks of the coat and scaffolding proteins seen in the radioactive experiment. The coat protein peaks in fractions 10–12, while the scaffolding in fractions 13–15. The same proteins present in the radioactively labeled procapsids are also seen here. The two horn proteins, gp53 and gp54, are recognized by the antibodies as present in the middle fractions as well. The bands of the scaffolding, portal and the three small proteins (gp55, gp57 and gp58) are more intense on the Western blot compared with the radioactive gels. There is a very distinct band of gp53 on the Western blot, which spans the entire gradient. A protein band corresponding to gp54 is also present in the middle fractions; its peak
overlaps the coat peak with much lower intensity overlapping the scaffolding protein peak.

To observe the group of structures present in these samples, the peak fractions of coat protein (10–12) were combined, the sucrose dialyzed, the samples concentrated, negatively stained with uranyl acetate and observed under the EM. The same was done for the peak fractions (13–16) of the scaffolding protein. As shown in Fig. A–5, both samples contain a variety of particles with different morphologies, which is expected since the samples had not been purified in multiple steps. Three major groups of structures are observed—structures that are spherical or lack angularity with internal protein density (group 1); structures also with protein density but having a defined angular shape (group 2), and empty shells lacking internal density (group 3). Structures of group 1 are most likely procapsids where there has not been any expansion yet and the internal protein density is probably due to the scaffolding protein. Shells with angular properties (group 2) are likely to be particles that have expanded before or after DNA packaging where the internal density is due to the presence of the nucleic acid which is not completely packaged; alternatively they may be infecting particles where the DNA has not entirely left the phage head. Group 3 particles are empty shells, which can have different origin. They may be maturing particles that have lost the scaffolding but have not yet packaged the DNA; alternatively, they may be particles from which the nucleic acid has been completely ejected during infection.
Figure A–4. Western blot of fractionated sucrose gradients of supernatants of lysed cells collected at 70 min after Syn5 infection. The cells were concentrated 1,000-fold before lysis. The membrane was probed with a mixture of anti-Syn5 and anti-scaffolding antibodies. int., internal protein; scaff., scaffolding protein.
Figure A–5. EM images of particles isolated from the middle fractions of the sucrose gradient in Fig. A–4. A), fractions 10–12; B), fractions 13–16. Group 1, spherical particles with protein densities; group 2, angular particles with internal density; group 3, empty shells. Stained negatively with 1% uranyl acetate and visualized at 100,000 × magnification.
Discussion

Labeling of Syn5 infected cells with radioactive amino acids between 25–35 min after infection (Fig. A–1) produces distinctive profiles of labeled proteins that match the protein profile of mature Syn5 virions. The data suggests that Syn5 directs protein production in host cells exclusively towards phage particle production. All the labeled protein bands of higher molecular weight than the coat exhibit the same profile as the structural proteins in Syn5 particles, with the addition of the scaffolding band present slightly above the coat. The proteins smaller than the coat are more difficult to identify based on migration only. The scaffolding breaks down into smaller products so it is possible that some of the smaller products belong to the scaffolding. It is also possible that there is a Syn5 protein that is not a structural protein but is either expressed in enough copies to give an intense gel band or is a Met/Cys rich protein. Based on molecular size estimations, a likely candidate is ssDNA binding protein (gp21) since it has a close theoretical molecular mass to the intense bands below the coat (21 kDa) and may be present in high numbers in the cells. Syn5 proteins involved in phage DNA synthesis e.g. DNA polymerase, were not identified on the gels. It is possible that their copy numbers per cell are not sufficiently high to be visible on the gel.

The distinct sedimentation of the coat from free protein in the top fractions to the middle fractions of the gradient and then to the bottom of the gradient strongly suggests that the isolated procapsid-like particles are in fact truly procapsids and a mandatory step of the assembly pathway.
The SDS gels of the sucrose gradients of the radioactively labeled particles (Fig. A–3) show that the peaks of scaffolding proteins do not coincide with the peak of the coat in the middle fractions where procapsids are present. This suggests that there are two groups of newly made particles in the middle of the gradient with densities different from the mature virions and free soluble proteins. The first group consists of procapsid particles (fractions 10–14) that contain the coat, scaffolding, portal, the two major internal proteins and possibly the small internal protein (gp43). The second group of particle (fractions 8–11) contains the tail of the scaffolding peak, all of the other proteins present in the procapsid peak as well as the three smaller proteins—gp55, gp57 and gp58. This second group of particle is most likely procapsids which have released the scaffolding protein but have not yet packaged the DNA. This has been observed in other phages (Aksyuk and Rossmann, 2011). The presence of the three small proteins (gp58 was identified as a head protein, most likely one of the knob proteins) as part of this group of particle suggests that these proteins may play roles in stabilizing the capsid shell once the scaffolding leaves. The decoration proteins of phage lambda have such a function (Yang et al., 2008).

The scaffolding protein band signal intensity is very low in the radioactive labeling experiment but Western blots of unlabeled cells show more distinctly the presence of the scaffolding protein (Fig. A–4). These blots reveal very similar protein profiles for the structures in the middle of the gradient with the addition of the two horn proteins—gp53 and gp54. This raises the question whether the horn proteins are indeed absent from the procapsid particles or are just not present in high enough copy numbers in the procapsids to be detected via radiography. The presence of the horn proteins on the
Western could also be due to the presence of particles which had injected or released their DNA and sedimented in the middle of the gradient, but which are not newly synthesized particles. The EM images of the two peaks showed the presence of a mixture of different particles in each peak (Fig. A–5).


