Probing the structure, self-assembly, and nanomechanics of bacterial microcompartments

Matthew Faulkner

Department of Functional and Comparative Genomics

Institute of Integrative Biology

University of Liverpool

This Thesis is submitted for the degree of Doctor of Philosophy

May 2018

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Liverpool or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Liverpool or any other University or similar institution except as declared in the Preface and specified in the text. This Thesis does not exceed the prescribed word limit of 100,000 words.

Matthew Faulkner

The experiments in chapter three, visualising carboxysome shell protein assembly dynamics, were included as part of the following publication on which I was a joint first author (as indicated in the author contributions of the publication)

Makus Sutter, Matthew Faulkner, Clement Aussignargues, Bradley C. Paasch, Steve Barrett, Cheryl A. Kerfeld and Lu-Ning Liu. (2016). Visualization of Bacterial Microcompartment Facet Assembly Using High-Speed Atomic Force Microscopy. Nano Lett 2016, 16 (3), 1590-1595, **DOI** 10.1021/acs.nanolett.5b04259 (http //pubs.acs.org/doi/full/10.1021/acs.nanolett.5b04259)

Reproduced with the express permission of the American Chemical Society.

The experiments in chapter four, direct characterisation of β-carboxysome structure and mechanics, were included as part of the following publication on which I was the first author

Matthew Faulkner, Jorge Rodriguez-Ramos, Gregory F. Dykes, Sian V, Owen, Selene Casella, Deborah M. Simpson, Robert J. Beynon and Lu-Ning Liu. Direct characterisation of the native structure and mechanics of cyanobacterial carboxysomes. (2017). Nanoscale 9, 10662-10673. DOI 10.1039/C7NR02524F (http //pubs.rsc.org/en/content/articlelanding/2017/nr/c7nr02524f#!divAbstract)

Reproduced with the express permission of the Royal Society of Chemistry.

`

Some of the figures and text in these publications have been reproduced exactly in this thesis, these were my own original work and the written permission of the journals was obtained for their reuse.

Acknowledgements

I would like to thank Dr Lu-Ning Liu for his advice, support and supervision, and everyone in the Liu Lab for their support and help over the past three years, notably Greg Dykes for the EM imaging and Jorge Rodrigues-Ramos for all the AFM guidance, image processing tips and physics discussions. I would like to thank everyone in lab G for making it an enjoyable place to work, especially Jean Wood and Paul Loughane for their advice and great technical help. I would like to thank all the CCI staff for their help with the use of the BioAFM. Thanks to the Kerfeld lab and their collaboration on the BMC-H project, especially Dr M.Sutter for the X-ray crystallography work. Special thanks to Dr Steve Barrett from the UoL Physics department for his image analysis course, countless hours developing ImageSxM and patience whilst teaching me how to write macros. I would like to thank Prof Rob Beynon and the Centre for Proteome Research for their collaboration with the Top3 MS. Outside of the lab, I would like to thank my friends and family for their support and I would like you to know I would not be where I am today without you all.

Abstract

Bacterial carbon fixation activity accounts for a significant proportion of global carbon capture. Bacteria do not have the complex carbon concentrating mechanisms and membrane bound organelles of plants and yet they have incredibly high photosynthetic efficiency. This is, in part, because of the alternative carbon concentrating mechanism bacteria employ involving the carboxysome. Carboxysomes are a type of bacterial microcompartment, a proteinaceous structure comprised of an outer semi-permeable shell encapsulating an internal enzyme filled lumen. By encapsulating the enzyme Rubisco, that catalyses the often rate limiting step of carbon fixation, carboxysomes are able to boost catalytic efficiency. Carboxysomes therefore increase overall carbon fixation activity. Bacterial microcompartments are widespread amongst bacterial phyla and apply the same approach to many enzymatic reactions. Our knowledge of the structure of carboxysomes and other such compartments and how they self-assemble is limited. This is thereby preventing the successful use of recombinant compartments to boost enzymatic efficiency in synthetic systems, plants, and other organisms. In this work, a multifactorial approach is employed to study the carboxysome, from the initial self-assembly of monomeric protein building blocks to the structural and mechanical properties of the entire functional compartment. Novel observations include the first application of near-native high-speed atomic force microscopy to the study of a purified micro-compartment shell protein and the first purification and characterisation of intact carboxysomes, from the widely studied model organism *Synechococcus elongatus* PCC 7942. Also the first direct functional and structural comparison of two different kinds of carboxysome, α and β, from different model-bacteria. Previously unobserved protein dynamic events within tiled arrays of shell proteins were recorded. These data shed light on the mechanisms of the initial stages of micro-compartment self-assembly. These data also gave rise to the first clear evidence that manipulating the interactions between individual amino acids in bacterial microcompartment shell proteins can control self-assembly. The mechanical properties of carboxysomes were identified as potential biomarkers to evaluate future synthetic carboxysome constructs through atomic force microscopy and nanoindentation.

Contents

Tables

Figures

Appendices

Alphabetical list of Abbreviations and Acronyms

3PGA – the molecule 3-phosphoglycerate

AFM – Atomic Force microscopy

BMC(s) - bacterial microcompartment(s)

CBB cycle – the Calvin-Benson-Bassham cycle

CCM – carbon concentrating mechanism

CA – the enzyme Carbonic anhydrase

CB(s) – Carboxysome(s)

cryoEM – cryogenic electron microscopy

EM – electron microscopy

E. coli – Escherichia coli

Eut - Ethanolamine Utilisation compartment

HS-AFM – High speed Atomic Force Microscopy

HOCH – Haliangium Ochraceum

H. neap – Halothiobacillus neapolitanus

MS – Mass spectroscopy

NI - Nanoindentation

Pdu – Propandiol utilisation compartment

PDB – Protein Data Bank (https //www.rcsb.org/)

RMSD – Root mean square deviation

RPM – rotations per minute

RuBisCO (Rubisco) – the enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase

RuBP – the molecule Ribulose-1,5-bisphosphate

SLPM –standard litres per minute

Syn7942 – *Synechococcus elongatus PCC 7942*

TEM – transmission electron microscopy

VLP – Virus like particle

WT – wild type

This page is intentionally blank and marks the end of the preface

*"**Imagination is more important than knowledge.** For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution. It is, strictly speaking, a real factor in scientific research"* **– *Ablert Einstein,*** *1929.*

Chapter 1, General introduction

This thesis is concerned with the fundamental understanding of carboxysome (CB) structural organisation, nano-mechanics, and mechanisms of self-assembly. Carboxysomes (CBs) were discovered over 40 years ago and have been extensively studied since (Gantt and Conti, 1969; Shively et al., 1973a). Despite this, the underlying mechanisms behind the self-assembly of CBs, the molecular level detail of their structural organisation and their physical characteristics are yet to be described in detail. The aim of this introductory chapter is to describe the current knowledge already outlined in scientific literature and to highlight the gaps in our understanding.

1.1 Bacterial microcompartments

1.1.1 Bacterial microcompartment shell structure

Bacterial microcompartments (BMCs) are widespread throughout bacterial phyla and BMC gene clusters are found in a relatively high proportion of publically available bacterial genomes (Axen et al., 2014; Sommer et al., 2017). They are proteinaceous structures that partition the cytosol of prokaryotes, akin to organelles in Eukaryotes. Cytosolic partitioning, via membrane bound organelles, is a key step in the evolution of Eukaryotes from Prokaryotes (Alberts et al., 2002; Martin, 2010). These Prokaryotic methods of partition are very interesting because they are entirely proteinaceous and self-assemble spontaneously, making them much simpler and more versatile than Eukaryotic membrane bound organelles (Diekmann and Pereira-Leal, 2013). These proteinaceous structures are polyhedral in shape (Shively et al., 1973a, 1973b) (Figure 1-1), closely resembling an icosahedron and at least approximately obeying 60 fold symmetry (Prasad and Schmid, 2012), and closely resemble a viral capsid (Yeates et al., 2007) (Figure 1-1). Hexameric and pseudo-hexameric major shell proteins form the facets and Pentameric minor shell proteins form the vertices (Kerfeld et al., 2005; Yeates et al., 2007).

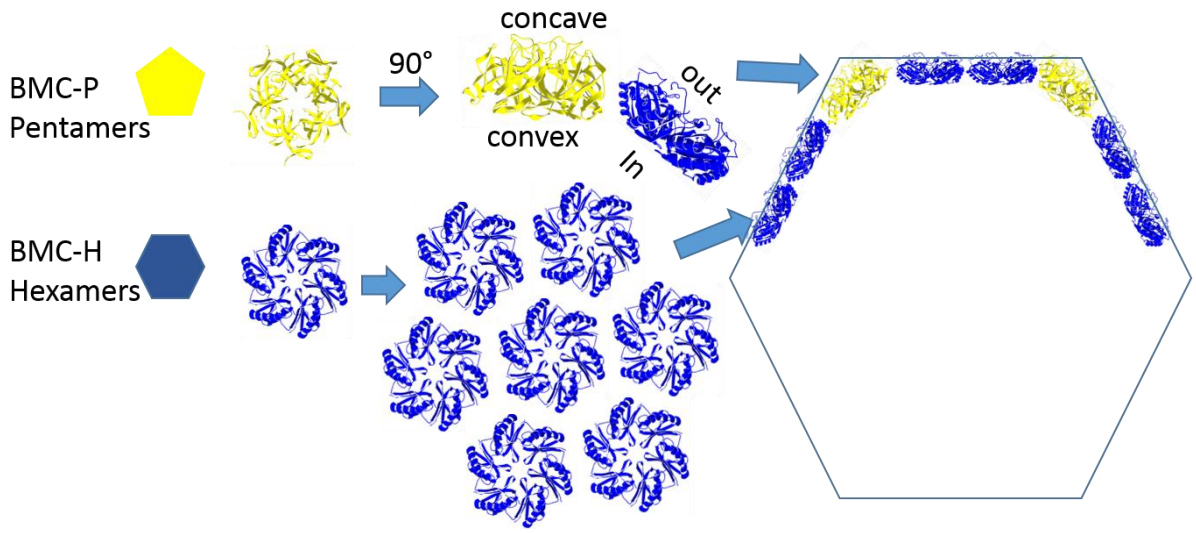


Figure 1-1 A schematic cartoon representation of BMC shell structure. The BMC shell (right) is formed of hexameric (blue), pseudo-hexameric (blue) and pentameric (yellow) shell proteins. The Pentameric proteins form the vertices and introduce curvature whilst the hexameric proteins form the facets (Kerfeld et al., 2005; Yeates et al., 2007).

1.1.2 Conserved features and homology between BMC

BMC shell proteins are highly conserved, at least structurally. All BMC studied in depth to date, and various bioinformatics investigations (Sommer et al., 2017), have shown that there are three conserved domains. These domains have been observed in all BMC shell proteins crystallised to date (Larsson et al., 2017; Pang et al., 2011; Sutter et al., 2013; Thompson and Yeates, 2014). Due to these domains, highly similar tertiary and quaternary structures are reported amongst all known BMC shell proteins (Tanaka et al., 2009). There are three distinct known conserved types of BMC shell proteins, those with either the BMC-H (pfam0936) domain, the BMC-T domain a tandem fused copy of (pfam0936), or the BMC-P (pfam03319) domain (Heldt et al., 2009; Kerfeld et al., 2005; Pang et al., 2011; Young et al., 2017). However, the contribution of these conserved features to the self-assembly and structural properties of the CB remains to be seen. The composition of the shell, in terms of the

ratio of BMC-H to BMC-T to BMC-P, and the effects changing this ratio may have are also yet to be investigated.

1.1.3 BMC encapsulate enzymes

BMC shell proteins are selectively permeable by allowing only specific small molecules to pass through their central "pore" (Chowdhury et al., 2015; Crowley et al., 2010; Park et al., 2017). BMC shells encapsulate enzymes in the BMC lumen (Fan et al., 2012). These encapsulated enzymes are therefore localised in a distinct environment from the cytosol (Perlmutter et al., 2016). The function of BMCs relies on this lumen, as they are used in metabolic processes to improve the efficiency of enzymes (Hopkinson et al., 2014), by raising substrate concentration or to prevent the escape of toxic intermediates from causing damage in the cytosol (Chowdhury et al., 2014 and Figure 1-2). The luminal environment is controlled by molecular transit through the semi-permeable shell protein pores (Chowdhury et al., 2015). This means the flux of reversible luminal enzymatic reactions can be modulated by controlling the concentration of substrate and product molecules (Cai et al., 2015b) (Figure 1-1 & Figure 1-2).

There are many diverse enzymatic processes encapsulated within diverse BMCs (Cheng et al., 2008). Three of the most widely studied BMCs with the most elucidated detail are the Ethanolamine utilisation compartment (Eut), the Propandiol utilisation compartment (Pdu) and the carboxysome (CB) (Bobik et al., 2015; Yeates et al., 2010) (Figure 1-2).

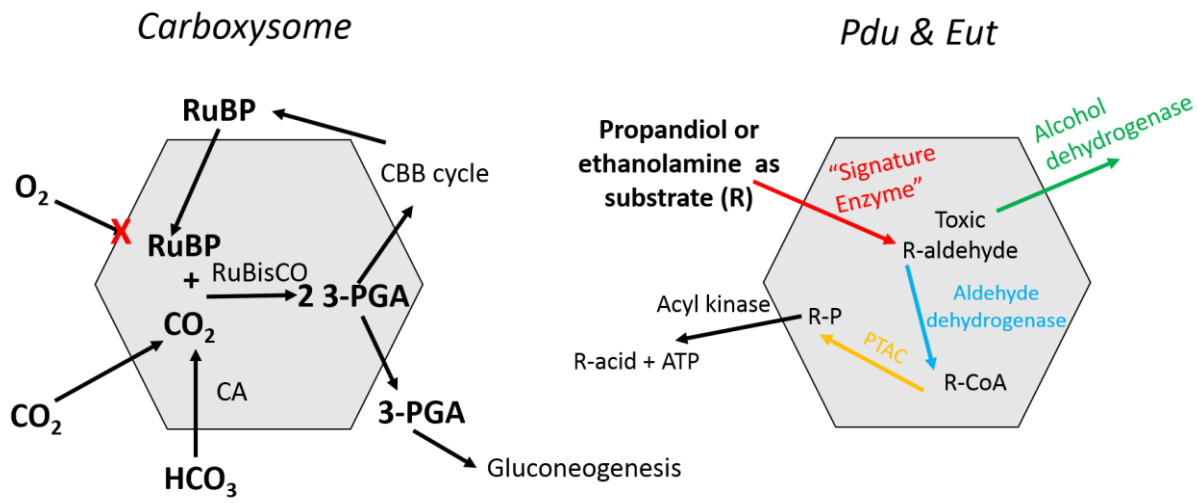


Figure 1-2 Generalised depictions of the reaction processes inside BMCs. The anabolic process of carbon fixation in the CB to the left, the catabolic process of propandiol/ethanolamine metabolism in the Pdu/Eut compartment to the right, adapted from (Axen et al., 2014). The exact details of molecular transit, such as if the CB shell excludes oxygen or presents a barrier to CO_2 escape, remain to be investigated but this model is widely accepted (Rae et al., 2013b).

1.2 Carboxysomes, the carbon fixation microcompartments

1.2.1 The encapsulation of Rubisco inside CBs

In the anabolic case of CBs, the shell proteins encapsulate two enzymes, carbonic anhydrase (CA) and Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco, commonly RuBisCO) (Shively et al., 1973a; Wildman and Bonner, 1947) (Figure 1-2 & Figure 1-3).

CA reversibly catalyses the interconversion of bicarbonate to carbon dioxide and water

$$H_2CO_3 \longrightarrow CO_2 + H_2O$$

Rubisco catalyses the first step in the Calvin-Benson-Bassam (CBB) cycle (Figure 1-3 and Bassham et al., 1950), otherwise known as the dark reaction of photosynthesis, fixing carbon dioxide into 3-phosphoglycerate (3PGA) (Wildman and Bonner, 1947)

$$CO_2 + \text{Ribulose-bisphosphate } (C_5H_{12}O_{11}P_2) \longrightarrow 2\ \text{3-phosphoglycerate } (C_3H_7O_7P)$$

However, Rubisco has both carboxylase and oxygenase activity. Therefore, as an oxygenase Rubisco can catalyse the oxygenation of Ribulose 1,5-bisphosphate (RuBP) in the first step of photorespiration (Peterhansel et al., 2010)

$$RuBP\ (C_5H_{12}O_{11}P_2) + O_2 \longrightarrow \text{3-phosphoglycerate } (C_3H_7O_7P) + \text{2-Phosphoglycolate } (C_2H_2O_6P^{-3}) + 2\ H^+$$

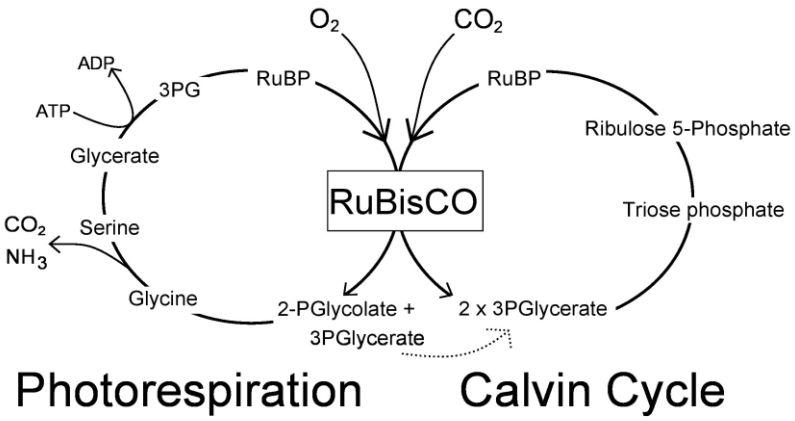


Figure 1-3 The Calvin-Benson-Bassam cycle. An overview of the CBB cycle showing the key steps including the reactions catalysed by CA and Rubisco inside the CB (Bassham et al., 1950).

CBs are found in most photosynthetic bacteria and some chemoautotrophic bacteria (Cai et al., 2012; Sommer et al., 2017). CB's enhance the efficiency of the often rate limiting step of the CBB cycle (Bassham et al., 1950). The carboxylase reaction is more desirable to the cell, under most

physiologically relevant conditions, than the oxygenase reaction (Peterhansel et al., 2010). The carboxylase reaction provides more 3PGA to the Calvin cycle, thus it results in the production of more ATP per photon of captured light; it is therefore more energetically favourable (Peterhansel et al., 2010). 3PGA also results in the formation of glucose for long term chemical energy (Michelet et al., 2013). Despite this, the enzyme kinetics of Rubisco favour the oxygenase activity, given that Rubisco has a higher affinity for O_2 than CO_2. This lower K_m for O_2 and higher K_m for CO_2 means at physiological, cytosolic, conditions in photosynthetic organisms Rubisco is very inefficient (Galmés et al., 2014; Kent and Tomany, 1984; McNevin et al., 2006; Yamori et al., 2006) (Figure 1-2 & Figure 1-3).

The encapsulation of Rubisco and CA by the CB functions as part of a carbon concentrating mechanism (CCM) to increase the concentration of CO_2 available to Rubisco, thereby increasing its efficiency (Figure 1-2). The CB shell is also thought to exclude O_2 and therefore prevent photorespiration by luminal Rubisco (Rae et al., 2013b), but this has yet to be proven. The CB increases the proportion of captured solar energy used for carboxylation, and therefore increases the photosynthetic efficiency of photosynthetic-bacteria (Cameron et al., 2013; Frey et al., 2016; Mangan and Brenner, 2013.; Rae et al., 2013; Shih et al., 2016).

Bacterial photosynthetic efficiency is higher than plants (Hanson et al., 2016; Price et al., 2013; Slattery and Ort, 2015; Zhu et al., 2010), and cyanobacterial photosynthesis is being optimised on an industrial scale (Kirst et al., 2014; Mangan et al., 2016; Ooms et al., 2016). Plants do not encapsulate Rubisco and it is localised freely inside the chloroplast (Harris and Königer, 1997). Some plants (called C4) do have a form of CCM (Gowik and Westhoff, 2011), but most plants (called C3 plants) do not, and lose ~30% of their carbon fixation potential due to photorespiration, at global median conditions (Zhu et al., 2010). It has been hypothesised that expression of a functional CB, inside the chloroplast to encapsulate the Rubisco, especially in a C3 plant, would increase the efficiency of carboxylation, and generate more ATP from the solar energy captured during photosynthesis (Bugbee and Salisbury,

1988; Hanson et al., 2016; Jensen and Leister, 2014; Long et al., 2006; Slattery and Ort, 2015). If plants were able to produce more energy through higher photosynthetic efficiency their growth rate should increase (Long et al., 2006; Zhu et al., 2010). CBs are the most studied BMCs because they have the most direct applications for improving crop yields (Frank et al., 2013). Food security, agricultural efficiency and the environmental impact are of global importance. Therefore, finding novel synthetic biology methods for improving upon current crops is a highly active area of research. One of the long-term aims of this work, and the wider field, is to improve the yields of key C3 crop species by improving photosynthetic efficiency using recombinantly expressed CBs. This study hopes to improve our fundamental understanding of the mechanisms behind the CB's self-assembly in order to inform those attempting to express CBs recombinantly, including in the chloroplasts of C3 plants. This could potentially be applied to all crops, including C4 species, once the method is well established.

1.2.2 Various types of CB

There are many types of Rubisco, with a range of macromolecular structures, kinetics and localisations (Table 2-1) (Alfreider et al., 2012; Badger and Bek, 2008; Galmés et al., 2014; Tabita et al., 2008). It is important to define the form of Rubisco before undertaking any study such as this because they are so strikingly different, despite all catalysing the same reaction (Table 1-1). For example, the macromolecular organisation of the different Rubisco complex isoforms ranges from single subunit homo-dimers to multi-subunit octamers (Table 1-1) *(Badger and Bek, 2008)*.

Table 1-1 A summary of the forms of Rubisco found in nature

Form	Macromolecular organisation	Phylogeny
IA	L8S8	Cyanobacteria (Form IAc) & Proteobacteria (Forms IAc & IAq) - *Halothiobacillus neapolitanus* c2
IB	L8S8	Cyanobacteria (Form IBc), green algae, higher plants (Form IB) - *Synechococcus elongatus* PCC 7942
1C	L8S8	Proteobacteria
1D	L8S8	Non-green algae
II	L2	Proteobacteria, archaea and dinoflagellate algae
III	L10	Archaea
IV	L2	Bacteria, archaea both photosynthetic and non-photosynthetic

A summary of the forms of Rubisco found in nature, and the broad phylogeny of where they are found (Badger and Bek, 2008). In the second column (Macromolecular organisation) L refers to the large sub unit of Rubisco and S refers to the small subunit of Rubisco (Spreitzer, 2003).

CB's can encapsulate form IAc and form IBc Rubisco. These distinct forms are encapsulated by two distinct types of CB, the α-CB encapsulating form IA and the β-CB encapsulating form IB (Rae et al., 2013) (Figure 1-4). These types of CB seem to have evolved separately in independent populations, for example, α-CB in mostly marine cyanobacteria and β-CB in mostly freshwater cyanobacteria (Rae et al., 2013). These two types of CB are very different (Figure 1-4). Previous studies have indicated they have alternative genetic arrangements, differ in size, and have variations enzyme content and protein composition (Cai et al., 2015a; Cannon et al., 2001; Cannon and Shively, 1983; Faulkner et al., 2017; Menon et al., 2008). β-CB are also bigger and have paracrystalline luminal packing of Rubisco (Figure 1-4).

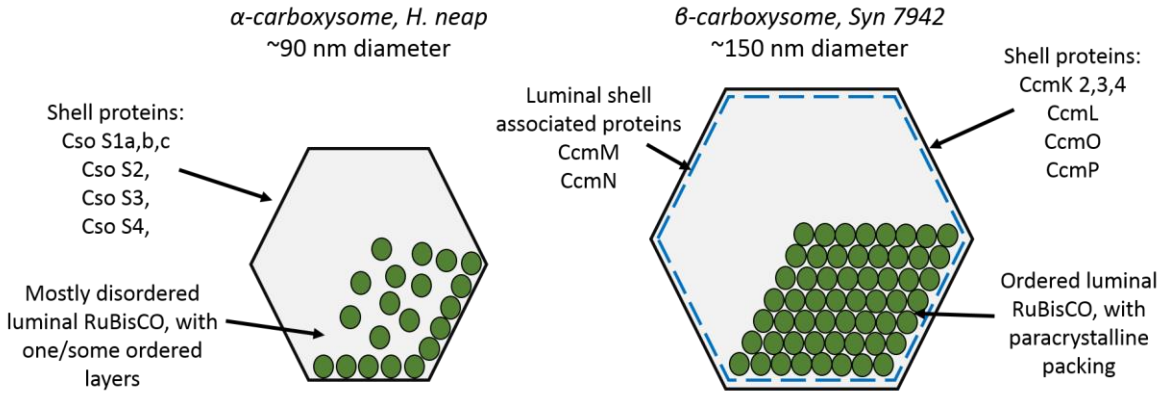


Figure 1-4 Schematics of the differing architecture of α andβ-CB. Left, a model representing the generally accepted structure and organisation of α-CB, such as in the model organism *Halothiobacillus neapolitanus C2* (*H. neap*). Right, a model representing the generally accepted structure and organisation of β-CB, such as in the model organism *Synechococcus elongatus PCC 7942* (Syn7942). Adapted from (Rae et al., 2013).

1.2.3 The model organisms used for the study of α and β CB

The model organism chosen for the study of α-CB was *Halothiobacillus neapolitanus* c2 (henceforth *H. neap*). *H. neap* is a species of chemolithoautotrophic Gammaproteobacterium. *H. neap* fix CO_2 via their using energy derived from the oxidation of reduced sulphur compounds (Kelly and Wood, 2000). *H. neap* has been widely studied for many years since the initial discovery and isolation of CB (Cannon and Shively, 1983; Shively et al., 1973a, 1973b). Therefore, there is an existing toolbox for *H. neap* CB isolation and some structural and functional study.

The model organism chosen for the study of β-CB was *Synechococcus elongatus* PCC 7942 (henceforth Syn7942), a single-celled, photosynthetic, freshwater, cyanobacterium. Syn7942 fix carbon via their CB using energy derived from captured photons of light (Frenkel et al., 1950). The fresh-water single-celled cyanobacterium Syn7942. This organism is naturally competent and genetically tractable, making the fluorescent labelling or knock-out of CB proteins relatively simple (Golden, 1988). Many studies have used this approach to study β-CB in vivo but to date, until this study, no successful isolation of native, intact and functional β-CB has been reported (Cameron et al., 2013; Chen et al., 2013; Golden, 1988; Kinney et al., 2012; Sun et al., 2016).

These two species are highly diverged and therefore there external and cytosolic environments are very different, likewise the environment their CB are adapted to (Figure 1-5).

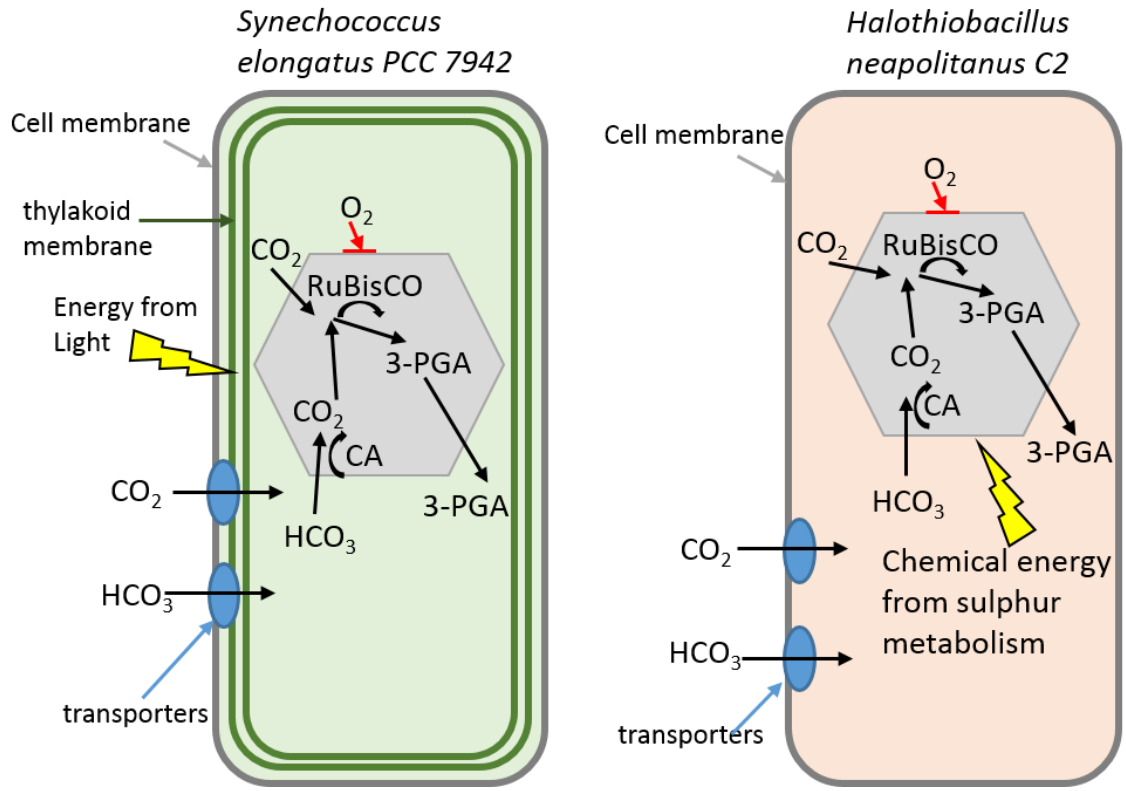


Figure 1-5 Cartoon representations of the Syn7942 and *H. neap* CCM. Left, a depiction of the *Synechococcus elongatus PCC 7942 (*Syn7942*)* bacterium including the relevant parts of the CCM and photosynthetic apparatus. Right a depiction of the *Halothiobacillus neapolitanus C2* (*H. neap*) bacterium including the relevant parts of its CCM.

1.3 Details of the individual α- and β- carboxysome protein components

1.3.1 The distinct genetic architecture of α- and β- CB

There are clear observed differences from previous studies of α- and β- CBs gene organisations. α-CBs are encoded by a single operon with only a single gene in a distant locus, or in an alternate direction, under the control of a second promoter (Abdul-Rahman1 et al., 2013; Baumgart et al., 2017). β-CBs have more complex arrangements usually involving more genes than α-CBs as well. β-CB usually have multiple distant loci encoding a few of the CB genes each. Interestingly the genes in these loci are

often grouped by similar function. In the case of Syn7942, there is a large central locus encoding the genes that have been defined as having structural roles within the CB (Long et al., 2005; Rae et al., 2012). This central locus is flanked by loci containing the minor proteins that have been defined as being non-structural. Many of these ascribed functions are predicted or require additional study to determine (Cai et al., 2012; Rae et al., 2013; Sommer et al., 2017).

The shell of the β-CB from Syn7942 is composed of the structural BMC-H proteins CcmK2, CcmK3 and CcmK4 (Figure 1-6), which appear as hexamers and form the shell facets (Kerfeld et al., 2005). CcmL, the BMC-P protein, forms pentamers that sit at the vertices between the shell facets (Tanaka et al., 2008). The final shell protein is CcmO, which is deduced to interface the edges of the shell facets along the icosahedral 2-fold axis of symmetry (Rae et al., 2012). The shell of the α-CB from *H. neap* is formed of the CsoS1ABCD BMC-H facets and CsoS4AB BMC-P vertices proteins (Figure 1-6) (Sutter et al., 2015; Tsai et al., 2007). The core of β-CBs is formed by Form 1B Rubisco L8S8 complexes formed by RbcL & RbcS, the β-CA called CcaA, linker proteins CcmM and CcmN (Gonzalez-Esquer et al., 2015). In α-CBs CbbL and CbbS proteins form the 1A L8S8 Rubisco complex and there are no homologues to CcmM or CmmN. CcmM has two active isoforms, CcmM58 and CcmM35, with distinct functions (Long et al., 2007). CcmM58 provides the interactions between the outer shell, β-CA and Rubisco molecules adjacent to the shell. Whereas, the 35 kDa truncated version CcmM35 is likely located in the CB's lumen, and crosslinks Rubisco enzymes (Long et al., 2011, 2010). CcmN acts as a bridge between CcmM and the shell by its two functional domains. The N-terminal domain of CcmN interacts with CcmM58 and the C-terminal peptide is capable of binding the major shell protein CcmK2 (Kinney et al., 2012). In addition, the β-CB's shell also contains the minor BMC-T protein CcmP that forms a dimer of trimers and likely modulates the shell permeability (Cai et al., 2013). RbcX is recognized as a chaperonin-like protein for Rubisco assembly, but its precise function in Syn7942 is still unclear (Figure 1-6) (Emlyn-Jones et al., 2006; Occhialini et al., 2016).

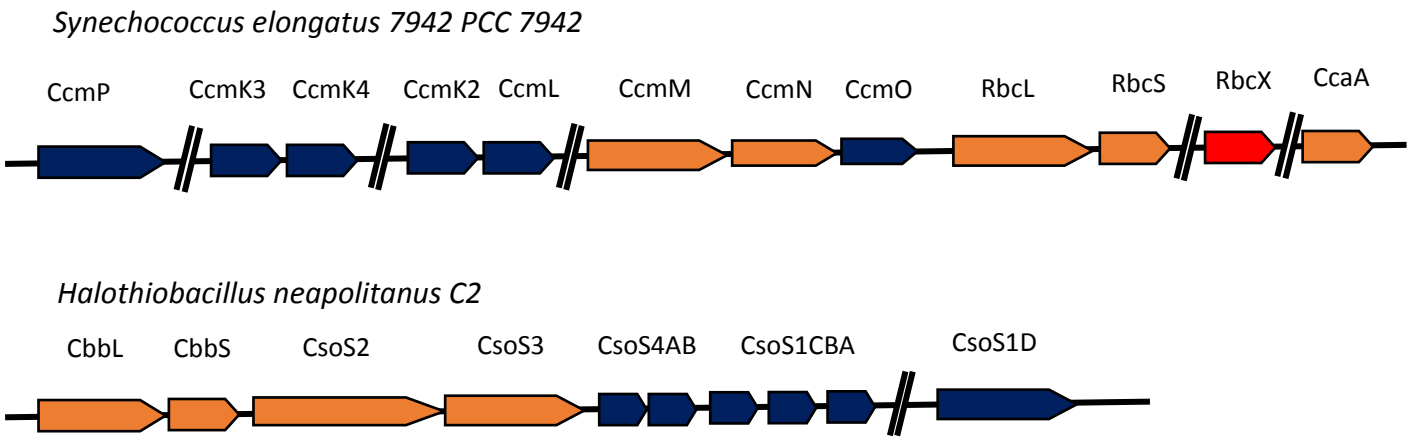


Figure 1-6 Gene architecture of the CB of the two model organisms described in this thesis. Genes encoding shell proteins are denoted in blue whereas genes encoding luminal proteins are encoded in orange. RbcX is a gene encoding a protein with unknown localisation and function within the Syn7942 CB so is denoted in red, RbcX functions as a Rubisco chaperone in other organisms (Saschenbrecker et al., 2007). (Figure Adapted from (Rae et al., 2013).

1.3.2 Techniques to study BMC protein structure

The structures of many different BMC shell proteins have been solved (Table 1-2). The structures of many luminal components have also been solved, namely the L8S8 Rubisco complex (Newman et al., 1993; Saschenbrecker et al., 2007). From these available structures, it is clear that all BMC-H hexamers fold in a highly conserved way despite primary sequence differences. In various studies presented in table 1-2, RMSD fits of aligned Cα atoms excluding the highly flexible C-terminal loop show how little the fold of the isoforms or orthologues vary (Kerfeld et al., 2005; Tanaka et al., 2008; Tsai et al., 2009). However, there are still many open questions about BMC proteins despite all the insight into their structure. For example, there are multiple packing models presented for the same structures (Cai et al., 2013; Kerfeld et al., 2005; Larsson et al., 2017; Tanaka et al., 2009). There is also no evidence to suggest how flexible or dynamic these packing arrangements may be under physiological conditions.

From these structures of WT BMC-H proteins and BMC-H proteins with point mutations the specific side chains involved in lateral hexamer-hexamer interactions have been discovered. Some studies have found the self-assembly properties of the BMC-H protein can be altered by point mutation, such as the propensity to form either sheets or nanotubes in PduA and PduA V51A (Pang et al., 2014). These are important observations and in part give rise to the idea of modular BMCs with engineered properties (Bonacci et al., 2012; Kerfeld and Erbilgin, 2015). Therefore, one of the aims of this thesis and the experiments described in chapter 3 is to further elucidate the factors and specific residues responsible for controlling BMC-H self-assembly (Sutter et al., 2016).

Table 1-2 A tabulated list of the relevant BMC shell protein structures currently available on PDB

Protein(s)	PDB code	Organism	BMC type	Reference
HOCH_5815, BMC-H	5DJB	*Haliangium ocraceum*	unknown	(Sutter et al., 2016)
CcmK1, BMC-H	3DN9	*Synechocystis sp. pcc 6803*	β-carboxysome	(Tanaka et al., 2009)
CcmK2, BMC-H	4OX7	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Cai et al., 2015b)
CsoS1, BMC-H	4OX8	*Prochlorococcus marinus MIT 9313*	α-carboxysome	(Cai et al., 2015b)
BMC-H, 3xBMC-T, BMC-P	5V74	*Haliangium ocraceum*	unknown	(Sutter et al., 2017)
HOCH_5812, BMC-T	5DIH	*Haliangium ocraceum*	unknown	(Aussignargues et al., 2016)
CcmK4, BMC-H	2A10	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Kerfeld et al., 2005)
CcmK4, BMC-H	2A18	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Kerfeld et al., 2005)
CcmK2, BMC-H	2A1B	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Kerfeld et al., 2005)
CcmK1, BMC-H	3BN4	*Synechocystis sp. pcc 6803*	β-carboxysome	(Tanaka et al., 2008)
CcmK1, BMC-H	4LIW	*Synechocystis sp. pcc 6803*	β-carboxysome	(Thompson and Yeates, 2014)
CcmK2,-c terminus BMC-H	3CIM	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Tanaka et al., 2009)
CcmK2, BMC-H	3DNC	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Tanaka et al., 2009)
Cso-CbbQ, BMC-T	5C3C	*Halothiobacillus neapolitanus*	α-carboxysome	(Sutter et al., 2015)
CsoS1A, BMC-H	2EWH	*Halothiobacillus neapolitanus*	α-carboxysome	(Tsai et al., 2007)
CsoS1C, BMC-H	3H8Y	*Halothiobacillus neapolitanus*	α-carboxysome	(Tsai et al., 2009)
OrfA	2RCF	*Halothiobacillus neapolitanus*	α-carboxysome	(Tanaka et al., 2008)
CsoS1A	2G13	*Halothiobacillus neapolitanus*	α-carboxysome	(Tsai et al., 2007)
CcmP, BMC-T	4H5T	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Cai et al., 2013)
CcmP, BMC-T	4HT7	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Cai et al., 2013)
CcmP, BMC-T	5LSR	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Larsson et al., 2017)
CcmP, BMC-T	5LT5	*Synechococcus elongatus PCC 7942*	β-carboxysome	(Larsson et al., 2017)
PduU, BMC-T	3CGI	*Salmonella enterica*	Pdu	(Crowley et al., 2008)
PduT, BMC-T	3N79	*Salmonella enterica*	Pdu	(Crowley et al., 2010)
PduA, BMC-H	3NGK	*Salmonella enterica*	Pdu	(Crowley et al., 2010)
PduA, K26A, BMC-H	4PPD	*Salmonella enterica*	Pdu	(Sinha et al., 2014)

A tabulated list of the relevant BMC shell protein structures currently available on PDB **(**[https //www.rcsb.org/pdb](https://www.rcsb.org/pdb)**).** The conserved BMC domain, organism and BMC type are listed for those with such information available in their relevant primary citations.

1.3.3 Structure and organisation of BMCs as seen by EM

BMCs, specifically CBs, were first discovered by EM during the study of blue-green algae (cyanobacteria) (Drews and Niklowitz, 1956). Subsequently, a plethora of EM studies had observed similar polyhedral bodies within bacterial cells (Gantt and Conti, 1969; Shively et al., 1973a, 1973b). True structural details regarding the organisation of the shell and luminal proteins did not come until EM imaging was applied to purified BMC, (Cannon and Shively, 1983; Holthuijzen et al., 1986), including many more, up until the present day. Following this, the roles of individual CB proteins were suggested based on EM studies of knock out mutants, made possible by the discovery of the CB genes and operon (Cannon et al., 2001; Orus et al., 1995; Price and Badger, 1989).

From the comparison of EM images of different BMCs from a variety of bacteria, the heterogeneity in BMC organisation and size was discovered. This is how the original models for α- and β- CB self-assembly and organisation were developed (Figure 1-2, reviewed in detail in Rae et al., 2013). α- and β- CBs are classified based on the form of Rubisco they encapsulate, but there is a growing body of evidence that aside from this their luminal Rubisco is organised in very different ways (Rae et al., 2013), based largely on the available EM data. Since the expected CB structure under EM imaging has been widely reported, EM imaging is used in this work to verify the purity and integrity of CBs in samples from my newly developed purification protocols. There is a large body of TEM data for *H. neap* α-CB but there is no TEM data on purified Syn7942 β-CB. A single study utilising the Percoll method for CB enrichment has previously studied these CB in vitro, but they were not kept in their native state during the Percoll gradient and ~70% of the shell protein content was lost relative to the whole cell starting material (Long et al., 2005).

Thin section EM imaging has been used to study BMC biogenesis in vivo (Cameron et al., 2013), the roles of individual BMC proteins via BMC protein knock out mutants (Cai et al., 2009; Kinney et al., 2012), and the synthesis of heterologously expressed BMC proteins (Parsons et al., 2010). Observations of the elongated BMCs formed when the BMC-P, pentameric, vertices proteins are knocked out are clear evidence of the role of BMC-P proteins in creating curvature at the vertices (Cai et al., 2009). Mixtures of the Pdu shell proteins were observed forming nanotubes, lattices and empty shells, when expressed heterologously in *E. coli.* These observations give insights into how the different types of BMC proteins interact with each other, and the roles these interactions play in BMC formation (Parsons et al., 2010). Thin section EM cannot track dynamic processes, however, and can only observe snapshots of what is occurring inside the cell.

A high resolution structure generated by cryogenic electron microscopy (cryoEM) is yet to be reported for an entire native BMC complex. A structure solved by cryoEM has recently been reported for a synthetic partial BMC structure (Sutter et al., 2017). This full 3D structure provided the first clear experimental evidence, albeit in a non-native heterologously expressed system, of details such as the orientation of shell proteins in the BMC shell and packing angles between BMC shell proteins, in a full 3D structure. Many high resolution cryoEM structures of viruses have been reported (Adrian et al., 1984; Agirrezabala et al., 2015; Jiang and Tang, 2017; Lee and Gui, 2016; Sirohi et al., 2016), and these are thought to resemble the icosahedral BMC structure.

1.3.4 Atomic Force Microscopy and its application to BMC study

Atomic force microscopy (AFM) is an imaging technique where a microscopic tip typically only a few nanometers in size physically traces over the surface of the sample. This tip at the end of a cantilever

known as a probe, A laser-detector array tracks the vertical deflection of the probe resulting from the sample and this deflection is used to estimate the samples height. These height data are used to generate a 3D topology of the sample's surface. AFM is becoming a widely used technique for the study of bio-macromolecules due to the innate advantage of being able to image under buffered, controlled, near physiological conditions (Allison et al., 2010; Chang et al., 2012; Vahabi et al., 2013). More traditional bio-imaging techniques are not able to use controlled conditions in this way during imaging. AFM also has the capability to measure the mechanical and physical properties of the sample, due to direct contact of the AFM probe and the sample. Nanoindentation (NI) experiments where a known force is applied to the sample can be used to determine properties such as the Young's Modulus, spring constant and breaking force. This methodology has been widely reviewed (Marchetti et al., 2016; Roos, 2011; Schillers et al., 2017; Vinckier and Semenza, 1998).

High-speed AFM (HS-AFM) is becoming increasingly widespread, particularly in the study of biological samples (Ando et al., 2013; Casuso et al., 2011; Eghiaian et al., 2014; Imamura et al., 2015; Uchihashi et al., 2016). HS-AFM can capture high resolution images showing molecular level details of individual BMC proteins at speeds of up to ~1 second per image.

1.4 Engineering of bacterial microcompartments

1.4.1 Synthetic BMCs

The operon encoding *H. neap* α-CB has been cloned into *E. coli* previously (Bonacci et al., 2012). Heterologously expressed CBs and other BMCs are of interest because it remains to be seen if they self-assembly correctly, and are therefore comparable to the native compartment from the original species. There is no study to date to our knowledge that directly compares the structural and functional characteristics of the same CB, generated from the same operon expressed in two different

species. This comparison is essential in answering whether CBs assemble correctly in a different cytosolic environment, missing possible chaperones, under a different ion concentration and pH and other varied conditions. If there are measurable differences between CBs produced in different environments, it may mean they are not suited to some potential applications utilising heterologous expression; or that some species may have CBs more resilient to environmental change, and therefore should be used preferentially by synthetic biologists.

α-CBs from *H. neap* should be used for this purpose as no successful heterologous expression of fully formed, functional β-CBs has been reported to date. Due to the simpler genetic organisation of the *H. neap* α-CB in a single operon (Figure 1-6), the cloning of this CB is much simpler and more likely to succeed, than the cloning of β-CBs requiring complex assembly of multiple gene fragments. It is necessary however to compare the functionality of these two CB's directly in order to determine which has the highest specific Rubisco activity, and therefore the greatest potential activity when heterologously expressed.

1.4.2 BMC shell formation, BMC Shell structure and generation of empty BMC shells

The focus of some recent studies has been to generate synthetic scaffolds or empty BMC using luminal knock out mutants or heterologous expression of just shell proteins with no luminal contents (Fan and Bobik, 2011; Pang et al., 2014; Parsons et al., 2010; Sutter et al., 2017). There are many open questions in this area regarding shell formation.

It has previously been proposed that BMC formation occurs via several mechanisms, the 'inside-out' mechanism whereby the luminal contents form a pre-BMC assembly to which the shell binds, the 'simultaneous' mechanism whereby the luminal and shell proteins co-assemble in time and space and

the 'shell first' mechanism where at least a portion of the shell assembles first (Aussignargues et al., 2015; Cai et al., 2015a; Cameron et al., 2013; Chen et al., 2013; Kinney et al., 2012; Perlmutter et al., 2016; Yeates et al., 2007). Despite these studies, key questions such as the orientation of shell proteins within the BMC, the mechanism of shell protein-shell protein, shell protein-luminal protein and luminal protein-luminal protein interactions are still unclear.

The relative contributions of the shell protein and the luminal protein to the overall structure are not yet known. For example, the differences between empty and cargo loaded virus capsids have been previously reported (Carrasco et al., 2006; Marchetti et al., 2016; Mateu, 2012); the loaded and empty states of the nano-compartment encapsulin have also been studied (Snijder et al., 2016). A detailed study like this has not yet been done with BMC. There is no study to date that directly compares two different BMC of any kind. These questions are of paramount importance to unlocking the potential of BMC to synthetic biology. For example, if we treat BMC as modular, as the current theory states (Bonacci et al., 2012), and exchange one protein building block for another, how do the structural properties respond?

Questions such as these are more difficult to answer due to the added levels of complexity when studying the whole BMC. The structural roles of each protein are incredibly convoluted in the full BMC complex, and it is not possible to distinguish the different BMC-H proteins by AFM. Under native conditions without the addition of a tag the BMC-H proteins of any given BMC will likely all appear as identical hexamers in the topograph; for example, there are only minor differences in primary sequence and between X-ray crystal structures of Syn7942 BMC-H proteins CcmK2, CcmK3 and CcmK4 (Kerfeld et al., 2005; Rae et al., 2012; Samborska and Kimber, 2012). In a simplified system containing a single purified shell protein, some of these questions can be addressed more easily without the

convoluted environment of indistinguishable different BMC-H proteins; systems such as this have been established in previous studies (Lassila et al., 2014; Pang et al., 2014). Studying the major shell protein in isolation in vitro allows the investigation of molecular level detail of a single BMC-H protein.

1.4.3 How BMC respond to environmental changes

The bacteria which contain BMCs respond rapidly and effectively to environmental changes such as pH, temperature, and osmolality. This is an essential response for the organism to survive. Yet little is known about the way in which BMCs respond to changes in their environment. The cytosol of a bacterium is not a completely stable environment, so it stands to reason that BMCs must also adapt to environmental changes as bacteria themselves do. For many of the proposed future applications of BMCs, such as molecule delivery, this is the central obstacle that remains to be overcome before they can be developed further. This is discussed in many reviews and research articles describing the use of BMC as drug delivery vehicles, for example (Chessher et al., 2015).

The concept of pH dependent assembly and disassembly is widely discussed but has yet to be studied in detail in CB. It has been studied in Pdu BMC (Kim et al., 2014). It remains to be seen if other BMC, including CB, have similar responses to environmental factors such as pH.

1.4.4 Engineering the function of BMC

Given their ability to encapsulate a multi-step, multi-enzyme, complex chemical process BMCs are highly active areas of research. BMCs are of interest because this enzymatic encapsulation offers a whole host of possibilities if it can be manipulated by synthetic biology (Frank et al., 2013). If we could heterologously express functional BMC in the bacteria used in the biotechnology industry and

encapsulate the enzymes catalysing rate limiting steps, there could be significant gains in productivity (Frank et al., 2013; Kerfeld, 2016; Quin et al., 2016). Approaches like this could also allow the production of high value molecules currently not able to be produced biologically.

Before we can begin to manipulate these structures to form designed nanoreactors we first need to understand them at the molecular level, i.e. what role does each BMC protein play in self-assembly, substrate influx, or structural integrity? In order to modulate the specific enzyme activity or molecular flux of a BMC in a controlled and predictable manner the individual factors that affect these properties need to be understood (Chessher et al., 2015; Frank et al., 2013; Kerfeld and Erbilgin, 2015). There are very few examples of successful repurposing (changing the function from catalysing one reaction to another by encapsulating a non-native enzyme) of functional BMCs in literature and those that there are only encompass a small proportion of the possibilities, yet it is often discussed and widely reviewed (Frank et al., 2013; Plegaria and Kerfeld, 2018). From an extensive search of the literature available only five examples of such studies were found where the new enzymatic activity achieved surpassed the activity of the un-encapsulated heterologously expressed enzyme (Huber et al., 2017; Lawrence et al., 2014; Lee et al., 2016; Liang et al., 2017; Yung et al., 2017). In these studies, the specific activity the synthetic encapsulated enzyme was greater than the WT enzyme or the recombinantly expressed but not encapsulated enzyme. In other similar studies this comparison was not made, or the WT enzyme had greater specific activity than the encapsulated synthetic enzyme.

There are more studies where proof of principle has been demonstrated. In these studies a synthetic compartment, or compartment heterologously expressed from a cloned operon, or the successful encapsulation of a new protein inside an empty shell was reported (Aussignargues et al., 2015; Cai et al., 2016; Fan et al., 2012; Fan and Bobik, 2011; Quin et al., 2016; Sargent et al., 2013). The vast

majority of these used Eut, Pdu, or partial/chimeric CB, not full CB, with the notable exception of the heterologous expression of α-CBs in *E. coli* (Bonacci et al., 2012). It is possible that there are so few examples of improved enzymatic activity because it could not be achieved in these other cases. This suggests that the widespread use synthetic BMCs is limited by a lack of fundamental understanding in how the structure of the BMC relates to its function.

There are very few examples in the current published literature of CB being expressed in a chloroplast. One notable example is the study by Lin et al, in which they express β-CB shells. These β-CB shells assemble in the Nicotiana benthamiana chloroplast (Lin et al., 2014a). In a similar study cyanobacterial Rubisco was expressed in the tobacco chloroplast where the endogenous Rubisco had been knocked out, but the specific activity of the cyanobacterial Rubisco was lower than the original (Lin et al., 2014b). The resulting plant had a slower growth rate due to the Rubisco replacement. Despite this success in expressing the major shell components and forming a BMC like structure and a functional Rubisco replacement, functional fully formed CB including the cargo enzymes have not been expressed in plants to date. I believe this is also due to a lack of fundamental understanding of the CB structure and self-assembly mechanics (Sections **1.1.1-1.1.3**).

This page is intentionally black and marks the end of chapter one.

*"**Research is what I'm doing when I don't know what I'm doing**." **Wernher von Braun**, 1957.*

Chapter 2, Materials and methods

2.1 Bacterial culturing

2.1.1 *E. coli* culture conditions

All *E. coli* strains used in this work were pre-existing in this lab, all were cultured in liquid using premade LB broth (miller's) powder at 25 g/l (ThermoFisher Scientific 12795027 Miller's LB broth base) or on LB agar plates made using premixed powder 37 g/l ThermoFisher Scientific 22700041 Lennox LB agar powder). Three *E. coli* strains were used; TOP10 (TOP10 Chemically Competent *E. coli* ThermoFisher Scientific C404010) for plasmid retention and plasmid production purposes, BL21 DE3 (BL21 DE3 Chemically Competent *E. coli* ThermoFisher Scientific C6000) for heterologous protein expression purposes and BW251113 (Datsenko and Wanner, 2000) for use of the λ red recombination system. All TOP10 and BL21 DE3 cultures were grown from a starting OD 600 nm ~0.05 (Jenway 6300 Spectrophotomer, Jenway, UK) at 37 °C. All BW251113 cultures were grown from a starting OD 600 nm ~ 0.1 at 30 °C as the plasmid for the λ red recombinase is lost at 37 °C. All cultures were grown for between 12 – 16 hours until OD 600 nm > 1.2 then sub cultured or transferred to 4 °C for short term storage. Long term storage of strains was achieved using 20 % glycerol as a cryoprotectant and snap freezing with liquid nitrogen before storage at -80 °C. Growth curves can be found in **Appendix B**.

2.1.2 Syn7942 culture conditions

Syn7942 was pre-existing in this lab. Syn7942 cultures were grown in BG 11 liquid medium (Rippka et al., 1979) or on BG 11 plates with 1.5% bacteriological agar (Agar bacteriological LP0011B ThermoFisher Scientific). Cultures were grown at a constant 30 °C with constant white illumination at 60 µE m^{-2} s^{-1}. Syn7942 were sub-cultured every 5-6 days when they began to enter stationary phase and cells began to elongate at OD 750 nm 1.5-2. Cultures were inoculated to an OD 750 nm of 0.15.

Cultures were grown in either 8-10 ml in small filter capped culture flasks (Nunc™ Cell Clulture Treated EasYFlasks™ 156367 ThermoFisher Scientific) or 30-50 ml in large filter capped culture flasks (Nunc™ Cell Clulture Treated EasYFlasks™ 156499 ThermoFisher Scientific). Cultures were regularly screened for contamination using fluorescence microscopy to compare bright field and chlorophyll auto-fluorescence images (Zeiss axioplan 2), any cultures found to contain >~1% non-chlorophyll auto fluorescent cells were discarded. Short-term storage was at 23 °C under 20 µE m^{-2} s^{-1} constant white light illumination. Long-term storage of strains was accomplished using 5 % DMSO as a cryoprotectant and snap freezing with liquid nitrogen before storage at -80 °C (Faulkner et al., 2017; Sun et al., 2016). Growth curves are available in **Appendix B**.

2.1.3 *H. neap* culture conditions

H. neap used in this work was acquired from ATCC (The American Type Culture Collection, https //www.lgcstandards-atcc.org/en.aspx) as a freeze dried powder (*Halothiobacillus neapolitanus* Parker, Kelly and Wood ATCC 23641 C2) (Cannon et al., 2001; Hutchinson et al., 1965). Cells were recovered and propagated as per the instructions from ATCC in liquid ATCC medium 290 (*ATCC Medium 290 S6 Medium for Thiobacilli*, Hutchinson et al., 1967) or on ATCC 290 2% agar plates. Subsequent experiments were started from a cryogenic stock made from the initial culture using the Protect microorganism Preservation system (Technical Service Consultant Ltd). This was also used for long term storage at -80 °C. Cultures up to 20 ml were grown in universal tubes (Sterilin™ Polypropylene Universal Containers, ThermoFisher, UK) containing >50 % headspace or in conical flasks for larger volumes up to 2 litres also containing >50 % headspace. Cultures were grown for 6-7 days to near stationary phase at a constant 30 °C as according to (English et al., 1995). New cultures were inoculated using a 1 3 ratio of near stationary phase culture to fresh ATCC 290. Cultures were regularly screened for contaminants using bright field microscopy (Zeiss axioplan 2) (See 2.9.2). The growth of these cultures was tracked using OD 600 nm. Growth curves are displayed in **Appendix B**.

2.2 Transgenic bacterial strain generation

2.2.1 *E. coli* strains for heterologous protein expression

The vectors used in this work were synthesised by GeneArt™ (ThermoFisher Scientific) or generated during this work by restriction cloning. Gene fragments, namely HOCH_5815 & Syn7942 CcmK4 enhanced green fluorescence protein (eGFP), were generated by HF-PCR using CloneAmp high fidelity DNA polymerase (Takara Bio) and primers designed during this work that were synthesised by Eurofins genomics. Restriction enzymes, namely Nco1 & Xho1 used were all produced by NEB (New England Biolabs INC, USA) and used as per the manufacturer's instructions. For a full list of vectors and primers used see **Appendix A**. All vectors used contained T7 promoter sequences and therefore IPTG was used to induce expression in all strains described here.

BL21 DE3 *E. coli* were used for all heterologous protein expression during this work. Chemically competent *E. coli* were transformed with heterologous expression vectors using heat shock as described in detail by Froger and Hall (Froger and Hall, 2007). Transformants were grown on selective LB agar plates and screened by colony PCR. Colony PCR was conducted with a few whole cells for template DNA using Biomix red premixed PCR reagent (Bioline) as per the manufacturer's instructions.

2.2.2 Syn7942 strains with carboxysomal fluorescence labels

Fluorescence labelling in this work was based on the original protocol for mutagenesis of Syn7942 (Golden, 1988), the lambda red recombination system (Mosberg et al., 2010), as described in more recent work in this laboratory (Sun et al., 2016). All of the eGFP labelled strains used in this work were generated using the method as described. All the CB protein labelled strains used in this work were already available in the lab as cryogenic stocks, or as finished constructs in *E. coli* requiring only

plasmid mini prep, transformation into Syn7942 and genotype confirmation by colony PCR. GFP labelling primers and vectors are described in **Appendix A**.

2.3 Large-scale bacterial culture

2.3.1 Five litre Syn7942 conical flasks

Large conical flasks were used to grow 5 litre batches of Syn7942. These flasks, already containing BG 11 medium and a magnetic stirrer bar, were sterilised by autoclave; and were sealed with a rubber bung perforated with two hollow glass tubes, one long and one short. The long glass tube reaching close to the bottom of flask was submerged under the medium, and the short was in the headspace above the medium. After autoclave and cooling to RT, these flasks were placed on magnetic stirring platforms and the long glass tube was connected to an air pump via a 0.22 µm filter, thus mixing and aerating the culture. These cultures were inoculated to OD 750nm ~0.15 using an entire 50 ml batch culture grown for ~5 days to OD 750 nm ~1.5. Initial growth was under continuous illumination with light intensity ~60 $\mu Em^{-2}s^{-1}$ and increased to ~100 $\mu Em^{-2}s^{-1}$ after 2 days and ~140 $\mu Em^{-2}s^{-1}$ after 4 days. Cultures were grown for ~7 days until OD 750 nm ~2. Growth curves are shown in **Appendix B**.

2.3.2 Five litre Syn7942 bioreactor cultures

Syn7942 cultures were maintained in constant illumination in BG-11 medium supplemented with 10 m TES and 3 g/L $Na_2S_2O_3$ and grown in a 5-liter fermenter (BioFlo 115, New Brunswick Scientific, USA) at 30 ℃ under 100 $\mu E\ m^{-2}s^{-1}$ initially, with constant agitation (200 RPM) and aeration (250 SLPM). Mixing and aeration were gradually increased during growth (250 then 300 RPM & 500 then 750 SLPM). The light intensity was also increased to 200 then 300 $\mu E\ m^{-2}s^{-1}$ at approximately the 2 day and 4 day time points respectively. The growth of cultures was tracked using OD 750 nm measurements, starting OD 750 nm ~0.15 and final OD 750 nm ~3-3.5 after 5 days of growth. Growth curves can be found in **Appendix B**.

2.3.3 Large volumes of *H. neap* grown by chemostat

H. neap cultures were grown in the same 5 liter BioFlo fermenter in ATCC 290 liquid medium, at 30 °C with 350 RPM agitation, and 500 SLPM aeration. A chemostatic growth condition was achieved using a constant dilution with fresh medium at a rate of 5 % per hour. The pumped out culture was collected twice per day and the cells harvested by two step centrifugation, 300g to remove Sulphur precipitates followed by 8000g (Sorvall SLA-3000 rotor) to pellet cells. The resulting pellets were snap frozen and stored at -80°C for isolation at a later date, if not used immediately. Growth curves are available in **Appendix B**. (Cannon and Shively, 1983; Dou et al., 2008; Menon et al., 2008)

2.4 Individual Protein and intact carboxysome purification

2.4.1 Purification of overexpressed BMC shell protein, HOCH_5815

Single BMC shell proteins express readily in *E. coli* Bl21 DE3. After the presence of overexpression vector was confirmed by PCR, expression conditions were screened using a range of temperature and IPTG concentrations; four temperatures 18, 25, 20 and 37 °C each with five IPTG concentrations a negative 0 IPTG control, 10, 100, 250, 500 µM, making a total of 20 cultures. Each culture was sampled and checked by whole cell lysate SDS-PAGE after 4 hours and overnight from induction. From this test 500 µM at 37 °C for 4 hours was determined as the best condition trialled. Further IPTG concentration optimisation found 450 µM resulted in more expression. The same held true for all shell proteins in the pET26b vector (see **Appendix A**) in this study. The growth of cultures was tracked by OD 600 nm. Expression was induced in cultures at OD 600 nm 0.6-0.8 (growth curves shown in **Appendix B**).

The shell proteins form large aggregations when overexpressed (Lassila et al., 2014; Sutter et al., 2016). 500 ml of cells were harvested by centrifugation at 6000 g (Sorvall SLA-3000 rotor), at 4 °C and all subsequent steps were on ice or at 4°C, lysed using a lysis buffer (50 mM Tris-HCl, pH 7.8, 100 mM NaCl, 10 mM $MgCl_2$, 20 µL DNase 25 mg/mL, 20 µL lysozyme (Lysozyme from chicken egg white, Sigma

Aldrich, UK) 50 mg/mL, 1% Protease inhibitor cocktail (PIC) (Thermofisher, UK) and 3 rounds of French press at 350 MPa (Stansted Pressure cell homogeniser, Stansted Fluid Power LTD, UK). The insoluble components were separated from the lysate by centrifugation for 20 minutes at 20,000 g, (Sorvall SS-34 rotor). The pellet was washed and re-centrifuged in the same manner three times with strong detergent buffer (50 mM Tris-HCl, pH 7.8, 100 mM NaCl, 10 mM $MgCl_2$, 3% (v/v) Triton X-100) and then washed three times with 0% Triton buffer (50 mM Tris-HCl, pH 7.8, 100 mM NaCl, 10 mM $MgCl_2$) in the same fashion to remove the detergent.

2.4.2 Isolation of CBs from *H. neap*

H. neap is a model organism for CB study and as such there are many published protocols for CB isolation. The protocol used in this study was based on (Menon et al., 2008 & Dou et al., 2008). 5 litres of culture was grown as described in (Section **2.3.3**). 20 L of cells were pooled and harvested by centrifugation (**2.2.3**). All subsequent steps were carried out at 4 °C or on ice and the resulting samples were stored at 4 °C. The cell pellet was re-suspended in 20 ml of TEMB (Tris, EDTA & Magnesium buffer) (10 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 20 mM $MgCl_2$) and the presence of 2% cell lytic B, 1% PIC and 10 mg ml^{-1} lysozyme, on a rolling mixer for 1 hour prior to cell breakage by sonication, 20x 5 seconds bursts with an amplitude of 14.4 Å (Soniprep 150 plus, Wolflabs, UK). The cell lysate was then treated with 2% Nonidet P40 (Thermofischer, UK) for 1 hour. Cell debris was removed by centrifugation 10,000 g for 10 minutes, followed by a centrifugation at 50 000 *g* for 20 minutes to enrich α-CB in the pellet and discard some cellular components in the supernatant. The generated pellet was re-suspended in 2 ml TEMB and was incubated in the presence of 1% *n*-doceyl β-maltoside (Sigma) overnight while mixing by gentle inversion on a rotary mixer. The resulting sample was loaded on top of a step sucrose gradient 10 – 50 % in TEMB, in 10 % increments each step consisting of 2 ml giving a final volume of 12 ml. Followed by a 250,000 g ultra-centrifugation for 20 minutes using an RPS40T rotor. Fractions were characterized by TEM imaging, SDS-PAGE, and Rubisco assay.

2.4.3 Isolation of β-CBs from Syn7942

5 litres of culture was grown as described in (**2.3.2**). Cells were harvested at OD = ~3.5 before reaching stationary phase after ~5 days. All subsequent steps were carried out at 4 °C or on ice, and the resulting samples were stored at 4 °C. The cell pellet was re-suspended in 20 ml of TE buffer (10 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0) and the presence of 10% cell lytic B, 1% PIC and 10 mg ml^{-1} lysozyme, on a rolling mixer for 1 hour prior to cell breakage by sonication, 20x 5 seconds bursts 14.4 Å peak to peak (Soniprep 150 plus, Wolflabs, UK). The cell lysate was then hand homogenised and treated with 3% Triton X-100 (Sigma) for 1 hour. Cell debris was removed by centrifugation 10,000 g for 10 minutes, followed by centrifugation at 50 000 *g* for 20 minutes to enrich β-CB in the pellet and discard some cellular components in the supernatant. The generated pellet was re-suspended in 2 ml TE buffer and was incubated in the presence of 1% *n*-doceyl β-maltoside overnight while mixing by gentle inversion on a rotary mixer. Followed by 180,000 g ultra-centrifugation using an RPS40T rotor for 30 minutes, through a step sucrose gradient 10, 20, 30, 35, 40, 45 & 50 % in TE buffer. Each of the increments was 1.5 ml in volume, giving a final volume of 12.5 ml including the sample. Each sucrose fraction was characterized by fluorescence microscopy, TEM imaging, SDS-PAGE, and Rubisco assay to determine the presence and activities of β-CB components.

2.5 Assay for Rubisco specific activity

2.5.1 Assay on live bacterial culture

Cells were harvested by centrifugation from cultures during the exponential growth phase. Cell pellets from 3 ml of culture were re-suspended in 1 ml of Rubisco assay buffer (100 mM EPPS pH 8.0, 20 mM $MgCl_2$), Cell density was then checked by OD 750 nm for Syn7942 or OD 600 nm for *E. coli* or *H. neap*. OD was adjusted to 1 by dilution with Rubisco assay buffer and re-measured to ensure dilutions were made accurately. These prepared cells at equal densities were mixed with assay buffer containing 25 mM $NaH^{14}CO_3$ (Perkin Emler, US) in scintillation vials (6.5 ml snaptwist scintillation vial PP, 16x57mm,

VWR) for 2 minutes in a 30 °C water bath. All subsequent assay steps were carried out at 30 °C. To permeabilize the cells prior to the reaction cells were incubated for an additional 1 minute with 0.03% MTA (mixed alkytrimethylammonium bromide, Sigma, US). To initialise the reaction 2 mM D-ribulose 1,5-bisphosphate (RuBP, Sodium salt, Sigma, US) was added. After 5 minutes of reaction time 2:1 volume of 10 % acetic acid was added to terminate the reaction. The samples were dried for ~45 minutes at 95 °C in heat blocks to remove all buffers and drive off any unfixed $NaH^{14}CO_3$. Dried sample precipitates were re-suspended in 200 µl of MilliQ water and mixed with 2 ml of Ultima Gold XR scintillation cocktail (Perkin Elmer, US). Radioactivity measurements were then made using a scintillation counter (Tri-Carb, perkin elmer, US).

2.5.2 Assay on purified CBs

The general procedure was the same as described in (**2.5.1**). The protein concentration of purified CB samples was determined by Bradford assay. An aliquot of the sample was then diluted to 1 mg/ml in chilled assay buffer and kept on ice. 5 µl of the sample was then incubated with 25 mM $NaH^{14}CO_3$ in a final volume of 250 µl for 2 minutes and the reaction initialised with RuBP and proceeded for 5 minutes before termination and then continued in the same manner as in (**2.5.1**), but without the permeabilization step.

2.5.3 Determining Michalis-Menton enzyme Kinetics

A range of substrate concentration, RuBP from 0 to 5 mM, was used for Michaelis-Menten determination. The assay procedure was the same as described in (**2.5.1** and **2.5.2**) except the reaction initialisation step was carried out with varied concentrations of RuBP for the same sample. The resulting fixation rates were plotted against substrate concentration to create a Michaelis-Menten curve and a Lineweaver-Burk plot (Lineweaver and Burk, 1934), using Sigmaplot V14.0 (Systat

Software, Inc., San Jose California USA, www.systatsoftware.com) from which the V_{max} and K_m of Rubisco were estimated.

2.5.4 Assay data analysis

A standard scintillation count curve from a range of known concentrations was generated from simply drying and counting $NaH^{14}CO_3$ in MilliQ water. The same standard curve was used to calculate the amount of ^{14}C fixed from the scintillation count data in every experiment with the same batch of $NaH^{14}CO_3$. This value was calibrated for background by blank samples without providing RuBP and water only samples, then divided by 5 to calculate the rate of ^{14}C fixed per minute. These rates were normalised against protein concentration or OD to calculate the rate per mg/ml of CB or per OD of cells. The significance of these data was assessed by two tailed t-test.

2.6 Atomic Force Microscopy

2.6.1 Substrate preparation

Three kinds of AFM substrates (the surface on which AFM samples are imaged) were used in this study, mica (0.25 mm mica sheets, Agar Scientific, UK), HOPG (HOPG-ZYB, Windsor Scientific, UK) and glass slides (Corning microscope slides, Sigma Aldrich, UK). Mica disks were cut to a standard 8 mm diameter using a custom hole punch jig from a large mica sheet, HOPG was purchased as a 10 x 10 x 20 mm block. Both substrates were affixed to magnetic disks (SD-101, 12mm diameter, Bruker, UK) using a transparent two component resin adhesive (Araldite® Crystal, Screwfix, UK). Glass slides were cleaned prior to use as described in (Roos, 2011), and Mica disks were also affixed to the centre of glass slides using a transparent two component resin adhesive. Glass slides with or without mica disks affixed were used to enable simultaneous optical microscopy and AFM using the BioAFM system (JPK Nanowizard 3 AFM mounted on a Zeiss 880 LSM). All other substrates were also imaged using the

Multimode system (Bruker MultiMode8 AFM). Magnetic disks were used for all other AFM experiments as they provide a more stable support which induces less drift. Mica and HOPG were both cleaved using adhesive tape (Scotch® Magic™ tape) immediately before each use to ensure a clean surface.

2.6.2 Sample preparation

Purified shell protein samples were stored at ~80 mg/ml at 4 °C short term. Purified CB samples were stored at ~4 mg/ml 4 °C short term and -80 °C long term. Stock samples were taken from storage conditions and diluted to 2 mg/ml or 0.5 mg/ml respectively for adsorption to AFM substrates. Adsorption buffer (10mM Tris-HCl pH 7.5, 150 mM KCl, 25 mM $MgCl_2$) at room temperature and ambient pressure was used for adsorption of all samples described in this thesis. After absorption samples were washed three times in Imaging buffer (10mM Tris-HCl pH 7.5, 150 mM KCl) and subsequently imaged in the same.

2.6.3 Probe/cantilever stiffness calibration

Probes were calibrated using a contact based method in imaging buffer on a flat area of clean substrate with no sample prior to each AFM experiment. An average curve from three independent curves was used to estimate the deflection sensitivity. The probe was then withdrawn a large distance (>100 nm) away from the substrate before thermal tuning. The Lorentzian model was used to fit the data and estimate the spring constant of the probe. This same general approach was taken using either the point and shoot mode on our Multimode 8 or the force spectroscopy capture setting on our JPK Nanowizard 3 or Ultra S.

2.6.4 High-speed imaging

HS-AFM imaging at a scan rate of 30-60 Hz was used to study BMC shell protein assembly dynamics. The JPK Nanowizard Ultra S and Bruker dimension fast scan systems were used with USC-F0.3-K0.3 (Nanoworld, Windsor scientific UK) probes. Minimal forces (50-100 pN) were applied to avoid sample damage and artefacts induced by scanning. Patches of interest were found in overview scans (~3 x 3 µm) at relatively slow speeds (3-5 Hz). Once located the scan size was incrementally decreased and speed incrementally increased while optimising the other scan parameters to maintain high quality and high resolution images framing the patch of interest.

2.6.5 Combined AFM and confocal microscopy

AFM and confocal imaging were carried out simultaneously on the same sample. The AFM images were captured using the JPK Nanowizard Ultra S and the same method described in (Sections **2.6.1-4**). The confocal images were captured using the Zeiss 880 LSM and aligned with the AFM using a direct optical overlay, using the standard procedure detailed in the JPK user manual. CB labelled with GFP (Section **2.2.2**) were imaged using standard GFP excitation (458 nm) and emission (500-530 nm) settings (Lee et al., 2016) with a 63x objective with a pre-existing method from our research group as described in (Sun et al., 2016). The resulting images were processed using a combination of Fiji (Cardona et al., 2012) and the JPK data processing software (JPK, Germany). The confocal pixel size was calculated using the Zeiss black confocal software and the process described in the Zeiss 880 user manual.

2.6.6 CB/virus imaging and NI

CBs were imaged using either peak force tapping mode on the Multimode 8 or quantitative imaging (QI) mode on the Nanowizard Ultra S. DPN10 probes (Bruker) probes were used and calibrated before imaging (Section **2.6.3**). Minimal forces (50-100 pN) were applied to avoid sample damage and artefacts induced by scanning. Very high gains were used during scanning to ensure the probe tracks the sudden change in height correctly when moving from the substrate to CB, without this the probe may not track correctly and may apply very high forces to the CB or impact the side. Once found via overview images (~5 x 5 µm), CBs were 'framed' by zooming into a small scan (~400 x 400 nm) and imaged only once before NI. NI was performed using point and shoot on the multimode 8, and using contact mode force spectroscopy capture on the Nanowizard Ultra S. A single NI experiment consisted of 6 force curves, 1 captured on a clean area substrate near the CB to serve as a control curve and 5 centrally on top of the CB. The mean value of these 5 curves was taken to represent the CB.

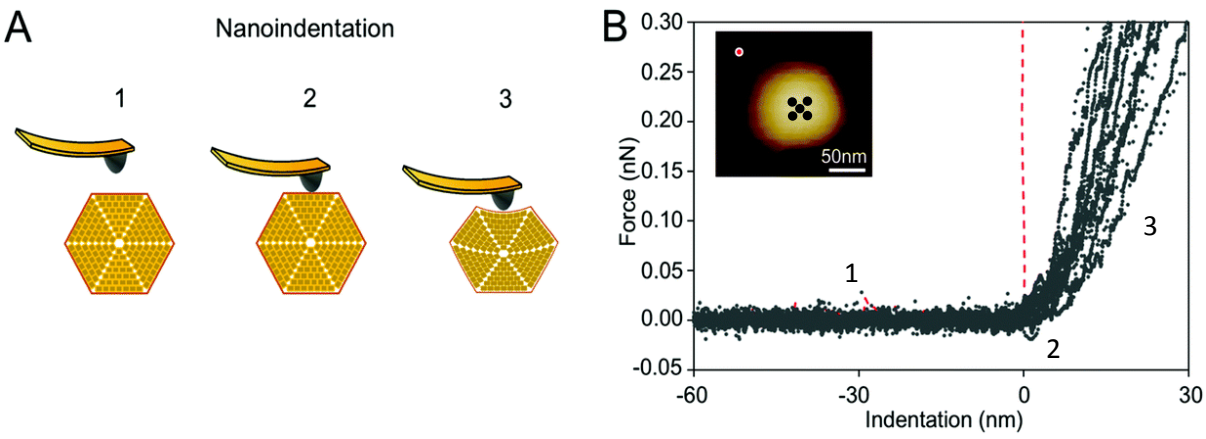


Figure 2-1 AFM-NI methodology A) Schematic representation of an AFM_NI measurement on top of a CB. 1 moving the probe down towards the CB, 2 contact with the surface of the CB, 3 indentation of the CB. B) Typical force curves generated during the process shown in B), the inset image indicates the locations at which these measurements were taken, CB measurements in black and AFM substrate control measurement shown in red.

2.6.7 AFM image processing

AFM images were flattened by plane fitting and setting the substrate regions of the image around the edges to minimum. Images with any discernible noise were median filtered and the LUT table was adjusted manually to increase contrast then applied to all future images in the same batch. This was done during or immediately after capture, to assess image quality, using the manufacturer's commercially available image analysis software purchased with our AFMs, Nanoscope analysis (Bruker, NanoscopeAnalysis version 1.8, USA) and JPK data processing (JPK, verison 6.058, Germany). Once in this form, subsequent processing was done on an image by image basis using Image SXM (https //www.liverpool.ac.uk/~sdb/ImageSXM/), Image J (Cardona et al., 2012), WSxM (Horcas et al., 2007) or Gwyddion (Nečas and Klapetek, 2011) depending on the desired output in each case e.g. 3D or Fourier transformed.

2.6.8 Calculating shell protein dynamics

Images processed as described in (**2.6.7**) were exported as batches of tif files corresponding to each patch. These were then batch processed in ImageSXM using a custom macro written expressly for this analysis. This macro automatically aligns the images frame by frame based on a custom designed auto-correlation function, applies Z thresholds and coverts the images into a binary format. These aligned binary images are subtracted by the following image to generate a difference image. Particle analysis can quantify the regions of difference in the difference image in terms of the number of hexamers gained and lost.

2.6.9 Determining CB mechanical properties

The analysis method to determine mechanical properties of CB like particles was described in detail (Faulkner et al., 2017).

Assuming that β-CB could have a mechanical behaviour similar to those of viruses, we used three typical models to determine its mechanical properties. The first one is the linear model, widely used to study virus rigidity (Table 4-1), where the cantilever and the particle are considered as two springs in series. The spring constant of β-CB k_{CB} was calculated using:

$$k_{\mathrm{CB}} = \frac{k_{\mathrm{total}} \times k_{\mathrm{cantilever}}}{k_{\mathrm{cantilever}} - k_{\mathrm{total}}} \quad (1)$$

Where $k_{cantilever}$ was the pre-calibrated spring constant of the cantilever and k_{total} was the slope measured in the range of interest of the force–distance curve recorded on top of the β-CB.

Stiffness is a property of the object and depends not only on the material it is comprised of, but also on its dimensions and geometry. Young's modulus provides a measurement of the intrinsic elasticity of the material. In the case of viruses, it is very common to use the thin-shell theory to estimate E_S. Young's modulus can be estimated using the following equation:

$$k_{\mathrm{CB}} = \alpha \frac{E_{\mathrm{S}} h^2}{R} \quad (2)$$

where α is the geometry-dependent proportionality factor (here we consider α = 1), which was reported to be a reasonable value for various virus capsids (Table 4-1) k_{CB} is the spring constant of β-CB, estimated using the lineal model (Equation 1), h and R are the shell thickness and the particle's radius measured by TEM (4.5 nm and 75 nm, respectively)

The third model used to estimate the mechanical properties of β-CB is the Hertzian model (Johnson and Johnson, 1987). This model is implemented in the commercial software of the Bruker and the JPK systems. If the sample is softer than the tip, Young's modulus E_H can be obtained using:

$$F = \frac{4}{3}\frac{E_{\mathrm{H}}}{1-\nu_{\mathrm{CB}}{}^{2}} R_{\mathrm{tip}}^{1/2} d^{3/2} \quad (3)$$

where F is the measured force, R_{tip} is the tip radius (for DNP cantilevers, R = 20 nm), and ν_{CB} is the Poisson coefficient of β-CB (here we consider ν_{CB} = 0.5, for soft biological samples) and d is the indentation depth, determined from the displacement z_p of the piezo-scanner, the initial contact distance z_0, and the deflection given by a hard wall $F/k_{cantilever}$:

$$d = z_{\mathrm{p}} - z_0 - \frac{F}{k_{\mathrm{cantilever}}} \quad (4)$$

Simulations of force-indentation curves were carried out using the Force Distance Curves tool in the Virtual Environment for Dynamic AFM (VEDA) software (Kiracofe et al., 2014), assuming a Hertz contact regime. The cantilever was assumed to have a tip radius R = 20 nm, spring constant k = 100 pN nm^{-1}, Young's modulus E_t = 130 GPa and Poisson coefficient μ_s = 0.3. For CB, a Poisson coefficient μ_s = 0.5 was used and Young's modulus was in the range from 0.5 to 500 MPa.

2.7 Electron microscopy

2.7.1 Sample preparation

The structures of isolated β-CBs were characterized using negative staining TEM as described previously (Arteni et al., 2008; Liu et al., 2008). Samples were adsorbed to the grid (Carbon-coated formvar-copper

grids, TAAB Laboratory & Microscopy, UK) for 5 minutes at ambient temperature and pressure before washing with ultrapure water and staining with 3% uranyl acetate.

2.7.2 Image capture and analysis

Images were recorded using a FEI Tecnai G2 Spirit BioTWIN transmission electron microscope. Initial image processing was conducted using Tecnai analysis (FEI). Further image analysis was carried out using ImageJ software (NIH Image). CB diameter results are presented as mean ± SD (Figure 4-6).

2.8 Carboxysome proteomic analysis

2.8.1 Sample preparation for proteomic analysis

The CB sample isolated from a 40%(w/v) sucrose gradient was initially in TE buffer (100mM TRIS-HCl pH 8, 10mM EDTA pH8, and 40 % sucrose) and was washed with PBS.

30µL of sample (~50µg) was made up to 80 µL with 25mM NH_4HCO_3 in a 0.5mL low-bind tube. 5 µL of 1 % (w/v) Rapigest in 25mM ambic was added and the sample incubated at 80 °C for 10 min. The sample was reduced by the addition of 5µL of 60mM DTT in 25 mM NH_4HCO_3 and the samples incubated at 60 °C for 10 min. Alkylation was carried out by the addition 5 µL of 180 mM iodoacetamide in 25 mM NH_4HCO_3 and the sample incubated at RT for 30 min in the dark. 5 µL (1 µg) of trypsin (200 ng/µL in 50 mM acetic acid) was added and the sample incubated in a heat block at 37 °C overnight. The following morning digests were terminated by the addition of 1 µL of TFA (trifluoroacetic acid) and incubated at 37 °C for 45 min, before centrifugation at 17,200 x g for 30 min and transfer of the clarified digest to fresh 0.5 mL low-bind tubes. Sample digest was mixed with an

equal volume of 50 fmol/µL of a yeast enolase standard digest in total recovery autosampler vials immediately prior to LC-MS analysis.

2.8.2 TOP3 Mass spectroscopy

Data-dependent LC-MS MS analyses were conducted on a QExactive quadrupole-Orbitrap mass spectrometer coupled to a Dionex Ultimate 3000 RSLC nano-liquid chromatograph (Hemel Hempstead, UK) in collaboration with the centre for proteome research. Sample digest (2µL) was loaded onto a trapping column (Acclaim PepMap 100 C18, 75 µm x 2 cm, 3 µm packing material, 100 Å) using a loading buffer of 0.1 % (v/v) TFA, 2 % (v/v) acetonitrile in water for 7 min at a flow rate of 9 µL min^{-1}. The trapping column was then set in-line with an analytical column (EASY-Spray PepMap RSLC C18, 75 µm x 50 cm, 2 µm packing material, 100 Å) and the peptides eluted using a linear gradient of 96.2 % A (0.1 % [v/v] formic acid) 3.8 % B (0.1 % [v/v[formic acid in water acetonitrile [80 20] [v/v]) to 50 % A 50 % B over 30 min at a flow rate of 300 nL min^{-1}, followed by washing at 1% A 99 % B for 5 min and re-equilibration of the column to starting conditions. The column was maintained at 40^{o}C, and the effluent introduced directly into the integrated nano-electrospray ionisation source operating in positive ion mode. The mass spectrometer was operated in DDA mode with survey scans between *m*/*z* 300-2000 acquired at a mass resolution of 70,000 (FWHM) at *m*/*z* 200. The maximum injection time was 250 ms, and the automatic gain control was set to $1e^6$. The 10 most intense precursor ions with charges states of between 2+ and 5+ were selected for MS/MS with an isolation window of 2 *m*/*z* units. The maximum injection time was 100 ms, and the automatic gain control was set to $1e^5$. Fragmentation of the peptides was by higher-energy collisional dissociation using a normalised collision energy of 30 %. Dynamic exclusion of *m*/*z* values to prevent repeated fragmentation of the same peptide was used with an exclusion time of 20 seconds. This was based on the method described in (Silva et al., 2006).

2.8.3 TOP3 Mass spectroscopy data analysis

The raw data file was imported into Progenesis QI for Proteomics (version 3 Nonlinear Dynamics, Newcastle upon Tyne UK; a Waters Company). Peak picking parameters were applied with sensitivity set to maximum and features with charges of 2^+ to 7^+ were retained. A Mascot Generic File, created by Progenesis, was searched against the *Synechococcus elagonatus* database from Uniprot (2,657 sequences) with the sequence of yeast enolase (UniProt P00914) added. Trypsin was specified as the protease with one missed cleavage allowed and with fixed carbamidomethyl modification for cysteine and variable oxidation modification for methionine. A precursor mass tolerance of 10 ppm and a fragment ion mass tolerance of 0.01 Da were applied. The results were then filtered to obtain a peptide false discovery rate of 1 %. Semi-quantification was carried out using a Hi-N (number of peptides quantified per protein) of 3 using the yeast enolase 50fmol/ µL as the reference protein (Silva et al., 2006).

2.9 Fluorescence imaging and confocal microscopy

2.9.1 GFP Fluorescence imaging using blue light illumination

For highly concentrated samples of CB with GFP tags, GFP emission was visible by eye, and was recorded with a standard digital camera, using broad range blue – UV illumination from a transilluminator (Dark Reader®, DR196 Transilluminator, Clare Chemical research, USA) and the accompanying filter.

2.9.2 Bright field and Fluorescence microscopy

Bright field and/or Fluorescence microscopy was conducted on a Zeiss Axioplan 2 (Carl Zeiss Axioplan universal microscope, Zeiss, UK); White illumination for bright field or broad range UV light excitation from a UV lamp. This excitation resulted in emission from both GFP and Syn7942 chlorophyll, which could be viewed independently using built-in red (588 nm) and green (509 nm) filter cubes in the microscope.

This page is intentionally blank and marks the end of chapter two.

*"**I don't believe it**. Prove it to me and I still won't believe it." **Douglas Adams**, (Life, the Universe, and Everything), 1982.*

Chapter 3, Visualisation of bacterial microcompartment facet protein assembly using high speed atomic force microscopy

3.1 Introduction

3.1.1 The previous studies of BMC shell proteins in vitro and scope of this study

The facets of the shell are predominantly composed of a homohexameric (BMC-H) shell protein (Kerfeld et al., 2005; Tanaka et al., 2010; Yeates et al., 2010), suggested to be either a single or a double layer (Kerfeld et al., 2005; Tanaka et al., 2009; Yeates et al., 2007). Empty shells, nanotubes and various other architectures can be visualized by transmission electron microscopy (TEM) after heterologous expression of shell proteins in *E. coli* (Parsons et al., 2010). Other studies of related BMC-H proteins such as PduA have begun to investigate molecular scaffolding approaches (Young et al., 2017). Whilst successful in creating scaffolds, these studies have thus far been confined to a very particular set of well controlled conditions within *E. coli* (Lee et al., 2017; Pang et al., 2014). Exactly how shell proteins self-assemble into higher order structures is an open question. One aim of this study is to find some of the missing details of the BMC self-assembly process by visualising the self-assembly of BMC-H using HS-AFM. Here I describe a combination of X-ray crystallography and HS-AFM to characterize a molecular sheet of HOCH_5815 and visualize the dynamics of BMC-H hexamer shell facet self-assembly for the first time.

3.1.2 Homology between BMC-H shell proteins

Aside from the afore-mentioned conserved BMC-H domains (**1.1.2**) in BMC major shell proteins, there is a wide range of highly homologous proteins reported (Cai et al., 2012; Sommer et al., 2017; Sutter et al., 2015). The BMC-H type major shell proteins, all form highly similar hexamer quaternary structures (Tanaka et al., 2009 & Table 1-2). BMC-H proteins contain a widely conserved lateral

interaction mediating (D/N-x-x-x-K RPH) motif, meaning the mechanisms of hexamer-hexamer interaction are likely similar (Kerfeld and Erbilgin, 2015; Klein et al., 2009). Therefore, for the work described later in this thesis regarding HOCH_5815 BMC-H proteins I work under the assumption my findings can be generalised amongst other BMC-H proteins. Despite this there are clear primary sequence differences in key regions of the hexamer, including the hexamer-hexamer interaction region (Cai et al., 2015b). Some of these differences are highlighted in the multiple alignment shown in Figure 3-1. For example, there are basic residues missing in CcmK2 such as L11 and Q53 which are conserved basic residues in both PduA and HOCH_5815 and thus create additional (D/N-x-x-x-K RPH) motifs.

Figure 3-1 A multiple alignment of three BMC major shell protein sequences, CcmK2 from Syn7942, PduA from Salmonella enterica and HOCH_5815. This alignment was generated using publically available gene sequences from the KEGG database (PduA t0836, CcmK2 Synpcc7942_1421) and the online Clustal Omega tool (https //www.ebi.ac.uk/Tools/msa/clustalo/).

3.1.3 The model organism *Haliangium ochraceum*

Haliangium ocraceum (henceforth referred to as *HOCH*) is a halophilic chemoautotrophic Myxobacterium. *HOCH* was used in this study as its BMC-H protein (HOCH_5815) has a high yield and purity when overexpressed and purified from *E. coli* (Lassila et al., 2014). Due to the similarity of BMC-H proteins, and the high de similarity with Syn7942 Ccmk2 (Figure 3-1), HOCH_5815 was deemed a suitable BMC-H protein to base wider conclusions upon. The same study also showed interesting stable protein assemblies persist under buffered conditions in vitro. These assemblies looked particularly suitable for imaging with HS-AFM. The reason for this high stability remains unclear,

however it may be due to the unusual genetic arrangement of shell protein genes in the HOCH BMC (Axen et al., 2014). These differences may result in longer periods of persistence in isolation or a change in stoichiometry relative to other orthologous BMC gene organisations.

3.2 Results and discussion

3.2.1 The cloning of HOCH_5815 into *E. coli*

The gene sequence of HOCH_5815 from start codon to stop codon (http://www.genome.jp/dbget-bin/www_bget?hoh:Hoch_5815) was synthesised by Geneart (ThermoFischer, UK) and was heterologously expressed in BL21* DE3 *E. coli* cells. Chemically competent *E. coli* were transformed by heat shock with pET11b_HOCH_5815, made using restriction cloning into the multiple cloning site of an empty pET-11b vector (Adgene, UK). Restriction digest was done using Xho1 and EcoR1 independently, empty pET-11b was transformed by heat shock into E. coli and purified using miniprep. See the materials and methods chapter (**2.2.1**) for more details and **Appendix A** for sequences.

The presence of expression vectors was confirmed by colony PCR using HOCH_5815 internal primers (**Appendix A**), after which these vectors were isolated from *E. coli* by miniprep and sequenced ('lightrun sequencing', GATC, UK) with primers in the T7 promoter and terminator to confirm the correct HOCH_5815 gene sequence was present.

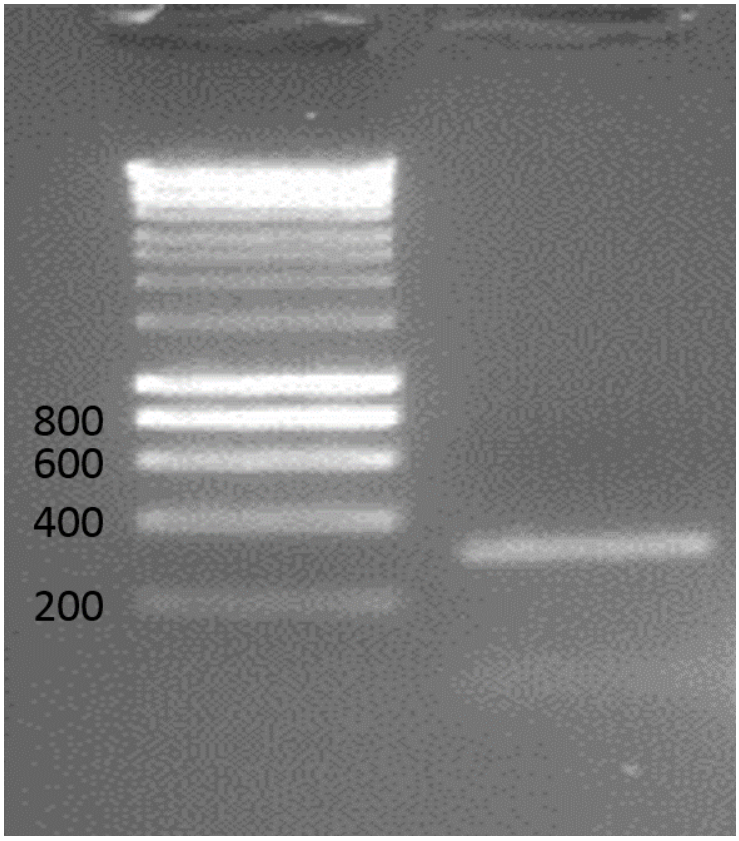


Figure 3-2 Agarose gel electrophoresis of HOCH_5815 screening PCR showing Hyper ladder 1 (bioline) in the left lane and PCR product band representing ~300 bp, the expected size for the HOCH_5815 gene fragment. The PCR product was amplified from 10 ng of a plasmid mini prep of BL21 DE3 transformed with HOCH_5815 Petm11 using primers directly after the ATG and before the terminator, inside the gene sequence. Thus the presence of this PCR band confirms the presence of the HOCH_5815 gene in the BL21 DE3 strain.

3.2.2 Overexpression and purification of HOCH_5815 protein

The expression conditions were trialled at various temperatures (18, 25, 30 & 37 °C) and IPTG concentrations (10, 100, 250, 500, 1000 μM), to confirm overexpression happened as expected, and to find optimal conditions (Figure 3-3, C). Protein levels were determined by SDS-PAGE using whole *E. coli* cells, compared to an IPTG free negative control. Following which expression was carried out at

37 °C with 450 µM IPTG as this is where to strongest band corresponding to the expected size of HOCH_5815 was observed. Overexpressed protein was then purified from the insoluble fraction of cell lysate using a series of washes with a strong detergent buffer, and a second series of washes to remove the detergent. The resulting sample was analysed by SDS-PAGE and the presence of a band of the expected size for HOCH_5815 was observed, with only minor contaminants visible (Figure 3-3).

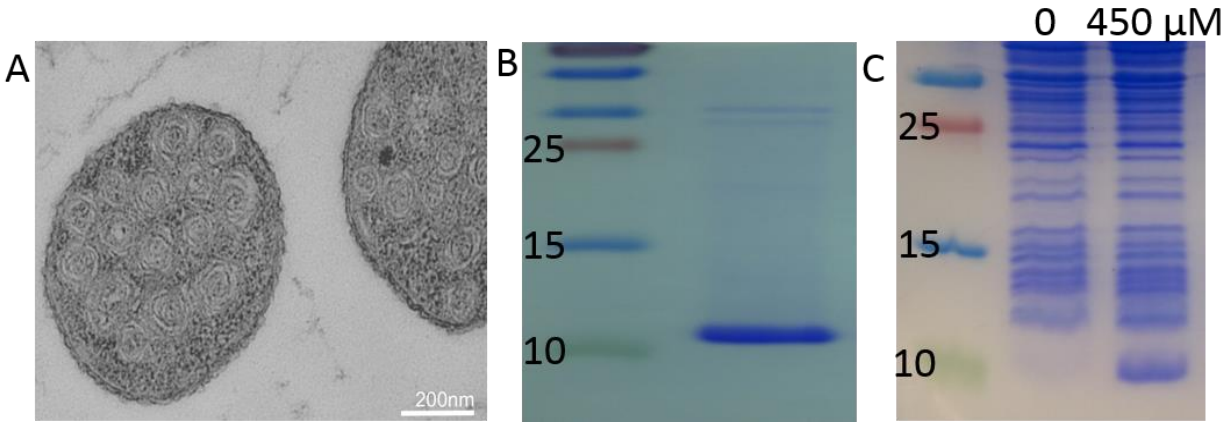


Figure 3-3 HOCH_5815 overexpression in *E. coli.* *(*A) Thin section electron micrograph of *E. coli* overexpressing the HOCH_5815 BMC-H proteins resulting in "swiss rolls" formed by sheets of hexamers. This thin section electron micrograph was gathered in collaboration with the Kerfeld lab, Berkeley USA. (B) SDS-PAGE showing two lanes, one containing 'PAGE ruler^{+}'$^{©}$ ladder (Thermofischer, UK) and the other containing purified HOCH_5815 protein with a single major band corresponding closely to the expected size of 10.9 kDa. The first three ladder bands from bottom to top represent, in green 10 kDa, in blue 15 kDa, in red 25 kDa. C) SDS-PAGE showing three lanes, The left contains ladder, the centre and right lanes contain whole *E. coli* cells induced with 0 and 450 µM IPTG respectively.

3.2.3 The structure of HOCH_5815

To characterize the building blocks of these higher order assemblies the HOCH_5815 protein was crystallised and its structure determined at a resolution of 1.8 Å by molecular replacement. These X-ray crystallography data were gathered in collaboration with the Kerfeld lab (PDB 5DJB). Eight monomers (four partial hexamers) form the asymmetric unit of the *P3* space group. The crystallographic 3-fold axis generates cyclic hexamers (Figure 3-4). Each monomer consists of the α/β fold characteristic of BMC-H monomers (Pfam00936 domain, Figure 3-4), as described for the previously reported BMC-H structures (Table 1-2). The closest structural homologue is CcmK2 (PDB 3DNC) a C-terminal deletion mutant of the major CB BMC-H shell protein with a root-mean-square deviation (RMSD) of 0.7 Å over 84 (out of 90) aligned Cα atoms, excluding the flexible C-terminal loop. The lateral edges of the HOCH_5815 BMC-H hexamer are 37 Å long and it has an overall thickness of approximately 36 Å including the C-terminal extensions (Figure 3-4.B). The hexamers have a distinct sidedness, a concave and a convex face (Figure 3-4.B), and pack into uniformly oriented layers in the crystal (Figure 3-4.C). The lateral interface between two hexamers buries 307 $Å^2$ per hexamer. A total of 12 non-bonded contacts are made among 8 residues per interface. Residues K28 and R78 of two separate monomers on adjoining hexamers make relatively large contributions to the lateral interface. The concave–concave interface primarily consists of interactions made by the N- and C- termini of the proteins. Convex–convex interfaces are mediated by hydrogen bonds and two salt bridges between charged residues (R66 and R62 to E65) on the surface of the protein.

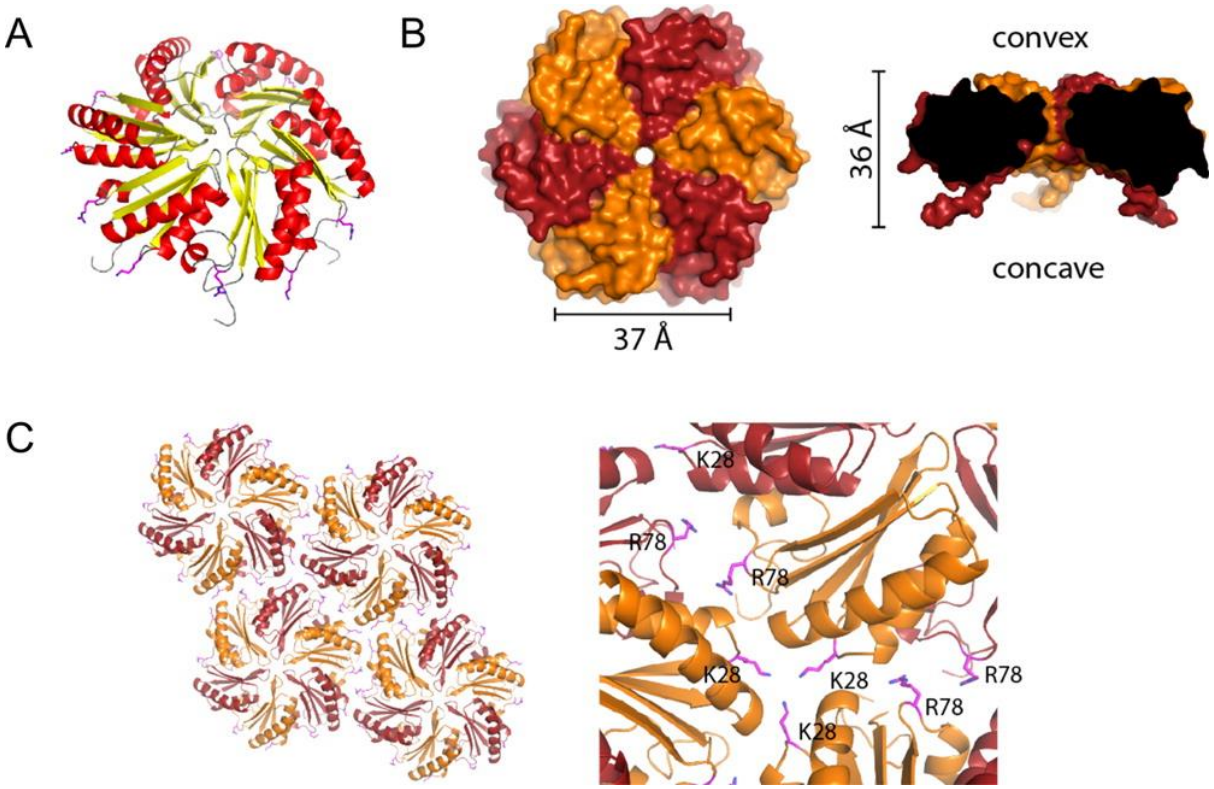


Figure 3-4 Crystal structure of the HOCH_5815 BMC-H protein. (A) HOCH_5815 adopts the characteristic α/β BMC-H fold (strands in yellow, helices in red) (Kerfeld et al., 2005). The edge residues K28 and R78 are represented by sticks. (B) A top down surface representation and a side-on centre sliced view with edge and thickness dimensions. The hexamer has a pronounced sidedness with distinct convex (top) and concave (bottom) surface. Protomer chains are alternatingly coloured dark red and orange. (C) The hexamers pack into uniformly oriented layers in the crystal. The zoomed-in view on the right shows laterally interacting residues K28 and R78 at the hexamer-hexamer interfaces (shown as magenta sticks). This X-ray crystallography data was generated in collaboration with the Kerfeld Lab.

This structure was integral to understanding the hexamer-hexamer interactions of HOCH_5815, as the interactions between hexamers, such as the salt bridges involving side chains of K28 and R78 were observed. The structure also served as a good model for what the expected dimensions and topology

should be under AFM imaging. Any differences observed during AFM imaging might suggest that the AFM buffered conditions were effecting the HOCH_5815 structure. The dimensions and dual sided nature of HOCH_5815 hexamers were observed and these observations gave rise to the initial ideas for the AFM experiments. The concave and convex sides have different net and localised areas of hydrophobicity and charge, implying that the interaction between each of these faces and the mica substrate is different and their topology under AFM imaging should be easily distinguishable.

3.2.4 HOCH_5815 organisation observed by AFM

Atomic force microscopy (AFM) in solution was used to characterise sheet formation by the HOCH_5815 BMC-H protein, which allows the observation of protein organisation and assembly under buffered near physiological conditions (under ambient temperature and pressure). The proteins were observed to form two-dimensional patches with varying sizes on the surface of AFM substrate (Figure 3-5.A). Cross-sectional analysis illustrates that the HOCH_5815 BMC-H forms flat protein sheets of ~3.5 nm thick (Figure 3-5.A), consistent with the thickness of a single hexamer in the crystal structure. This indicates that the sheets are composed of a single layer of hexamers or tiled array, reminiscent of (Lassila et al., 2014). The double layer formed by interactions between the C-terminal tails along the concave face, described by (Samborska and Kimber., 2012), was not observed.

AFM images of HOCH_5815 sheets captured at 17 seconds per frame (s/frame) using higher-magnification and smaller scan-size, were able to resolve the regular angles and straight edges of the hexamers within the tiled arrays (Figure 3-5.B & Figure 3-8). We were able to resolve the molecular details of protein organisation within these sheets by minimising the AFM scanning force, applying less than ~100 pN. The two distinct surface morphologies of the hexamer patches could be distinguished in AFM topographs based on the relative sizes of the central depression, measured as

the distance between the protruding regions of a single hexamer on each surface (Figure 3-5.C,D). These two morphologies corresponded to the concave and convex faces of the hexamers that were observed in the crystal structure (Figure 3-4.B). The AFM results show that BMC-H patches are composed of a monolayer of uniformly oriented HOCH_5815 proteins (Figure 3-5, C), which were also observed in the crystal packing. Patches were not observed with a mixture of orientations.

The hexamers have distinct electrostatic properties on their concave and convex faces, evident from the crystal structure. These oritentation specific electrostatic properties likely account for the different rates of shell protein attachment to the mica surface (Figure 3-6). When the concave face (which is relatively nonpolar) of the hexamer is exposed to the AFM probe and the relatively polar convex surface is attached to the negatively charged mica substrate, the sheets are more stable than sheets in the opposite orientation (i.e., convex up, concave on mica). Hence in this work, the focus remained on the concave face up oriented patches.

Higher resolution images using a small scan area (100x100 nm), within a large patch (>1 μm^2) were captured. These images allowed us to investigate the organisation and hexamer packing within the HOCH_5815 patches, representing the facet of a BMC, (Figure 3-5, D). In these images the edges of each individual hexamer are clearly defined and it is possible to overlay these hexamers with an accurate schematic representation of the hexamer packing (Figure 3-7, A). From this schematic representation more detailed measurements of the packing were possible such as the periodicity (~6.5 nm) and angle between adjacent hexamers (~60°) (Figure 3-7.B & C). This information sheds new light on how the major shell BMC-H proteins are organised within a BMC.

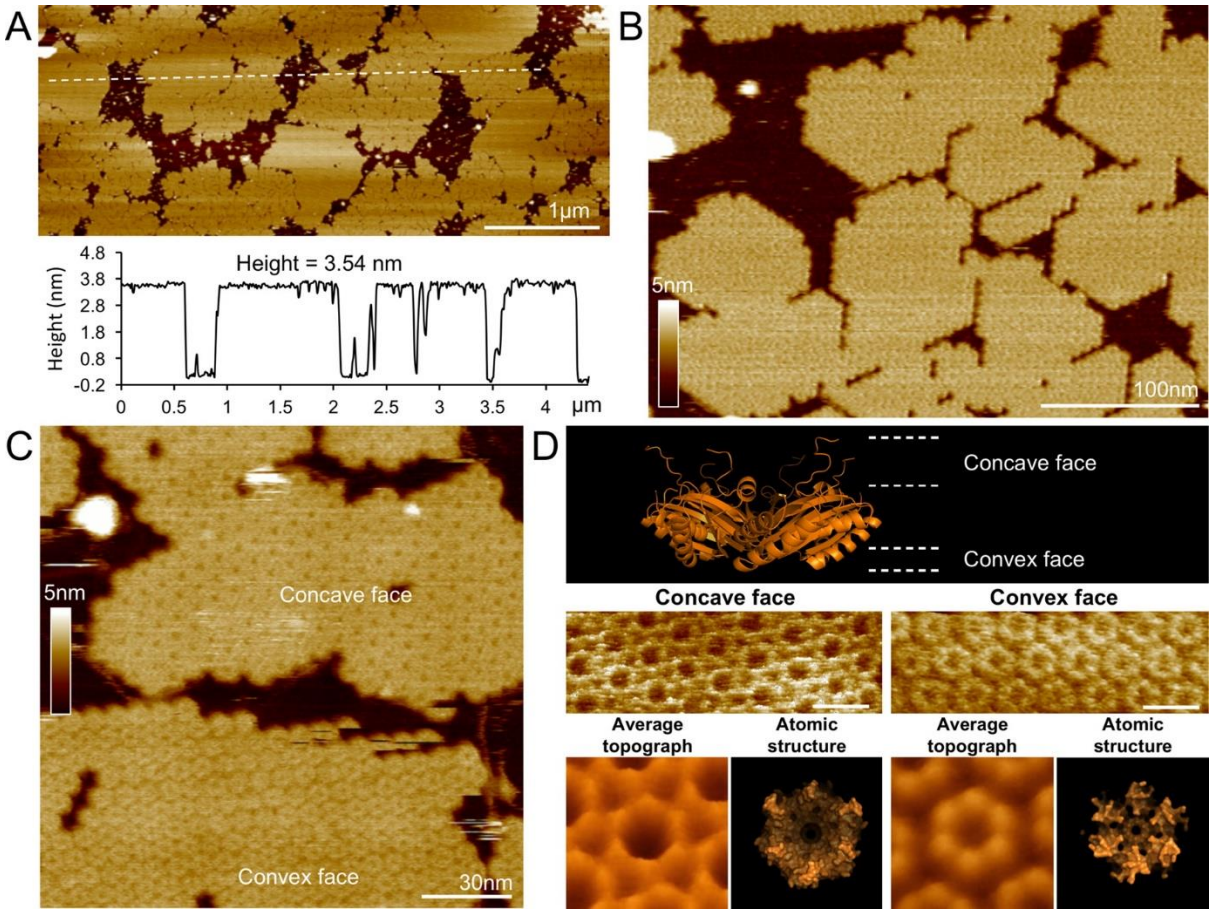


Figure 3-5 AFM analysis of the HOCH_5815 sheet. (A) Cross-section analysis using topographs captured of hexamers adsorbed to the mica surface indicates a monolayer 3.54 nm in thickness. (B) Higher magnification reveals greater detail about the organisation of hexamers as seen by straight edges and regular 60° angles (see Figure 3-7). (C) Between patches of hexamers two distinct surfaces morphologies can be observed, and only a single orientation was observed in any one given patch. (D) The relative sidedness of the sheet (convex versus concave face) accounts for the differing surface topographies and can be distinguished during AFM scanning by the diameter of the central depression that surrounds the pore. The concave face has a depression diameter of 52.8 Å whereas the convex face has a diameter of 47.1 Å measured by AFM cross-section analysis, compared with 51.6 and 45.7 Å, respectively, based on the crystal structure. The atomic structures were generated from the X-ray crystallography data gathered in collaboration with the Kerfeld Lab.

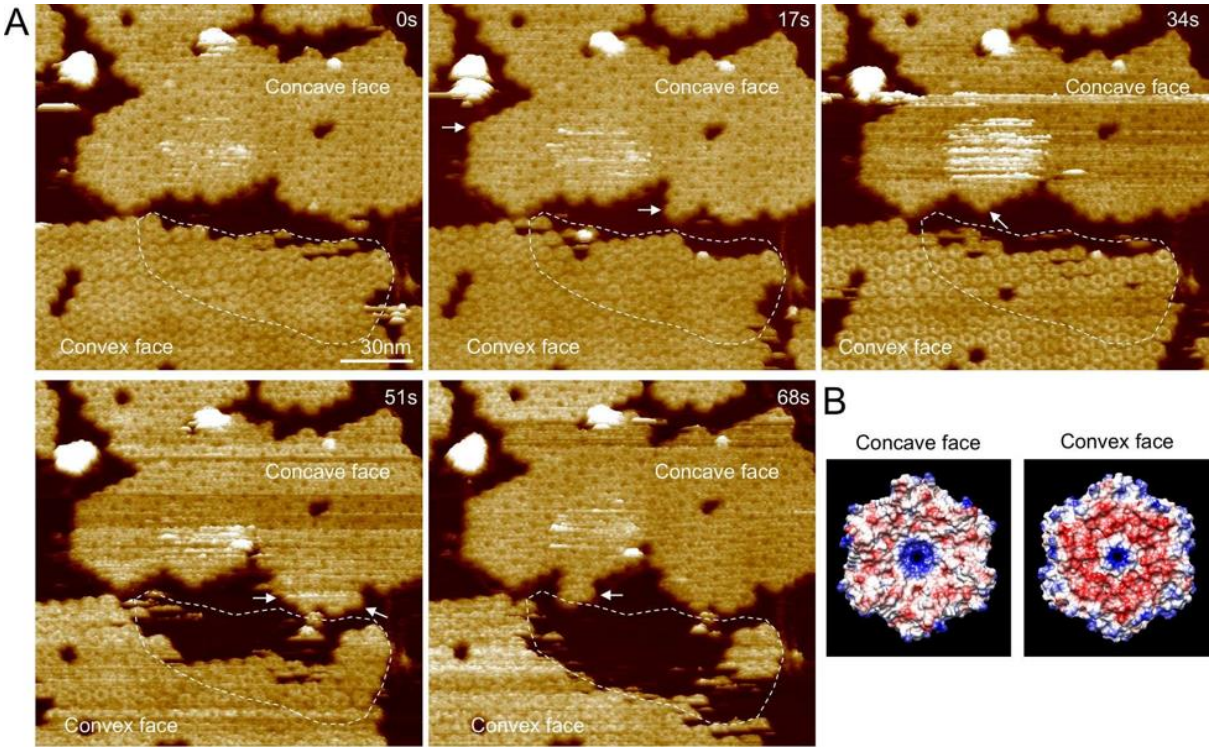


Figure 3-6 Different dynamic features of hexamer self-assemblies when concave and convex faces are attached to the mica substrate. (A) Five aligned AFM frames captured at 17 s/frame, showing that the two patches with distinct orientations present different dynamic features. The protein patch with convex face up displays greater protein dynamics (dashed area), whereas the other sheet with concave face up is relatively stable (arrows indicate a few single-protein motion events). (B) The electrostatic properties of the hexamer, showing that the concave face is relatively less uniformly polar than the convex face. The atomic structures were generated from the X-ray crystallography data gathered in collaboration with the Kerfeld Lab.

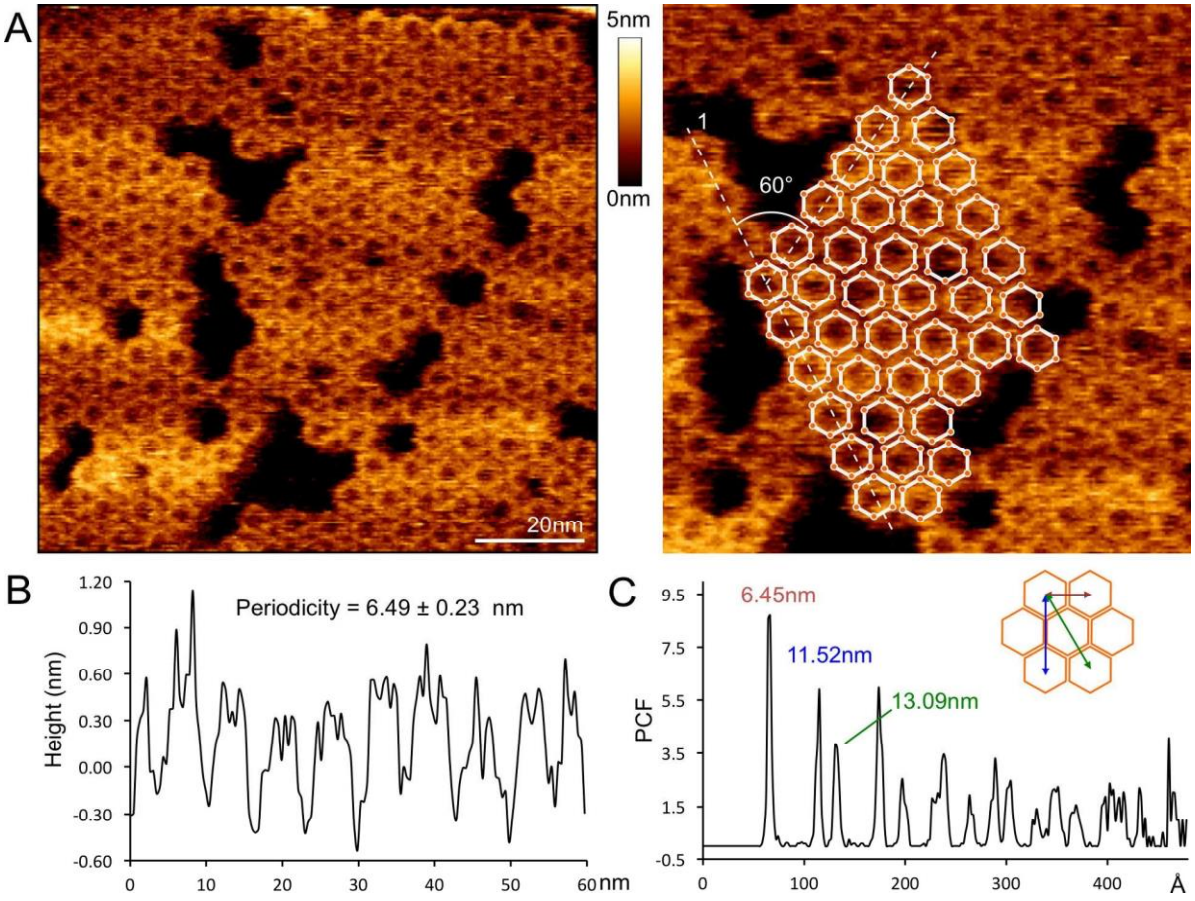


Figure 3-7 Organisation of HO hexamers in the self-assembled patches. (A) High-resolution AFM image showing the structure of individual hexamers and the "honeycomb" packing pattern of hexagons (60° angle) in the self-assembled layers. (B) Height profile along Line 1 in panel (A) indicates the periodicity of hexamer packing (6.49 ± 0.23 nm, n = 10) in the patch. (C) Pair correlation function analysis of the hexamer organisation in protein sheets. The closest interacting distances are 6.45 nm, 11.52 nm and 13.09 nm respectively, representing the typical assembly of hexamers as illustrated in the structural model (insert, red, blue and green arrows).

3.2.5 Tracking HOCH_5815 dynamics by HS-AFM

HS-AFM has evolved into a powerful tool for exploring the structure and dynamics of biomolecular systems (Ando, 2014; Ando et al., 2013; Casuso et al., 2011; Eghiaian et al., 2014). To investigate how HOCH_5815 BMC-H proteins self-assemble into BMC facets, I therefore observed the dynamics of HOCH_5815 hexamer sheets using HS-AFM (17 s/frame, 30 Hz scan rate). The lattice-like tiled-arrays of BMC-H hexamers were clearly identified, due to the ability of HS-AFM to resolve individual hexamers (Figure 3-8). Two hexamers (depicted by white arrows at 0 s) were observed dissociating from the sheet at 17 s. Another hexamer subsequently dissociated (white arrow shown at 17 s) and one new hexamer assembled into the protein sheet at 34 s (yellow arrow). Such translational motions of hexamers were continuously observed during this series of images and in many similar imaging series (Figure 3-8 & Figures 3-9 to 3-15). Similar observations were made during higher scanning speed (2 s/frame, 60 Hz can rate, Figure 3-10), revealing the dynamic nature of self-assembly of the shell proteins and flexible interactions between proteins in the shell. These results demonstrate that sheets are formed by the incorporation of preassembled hexamers.

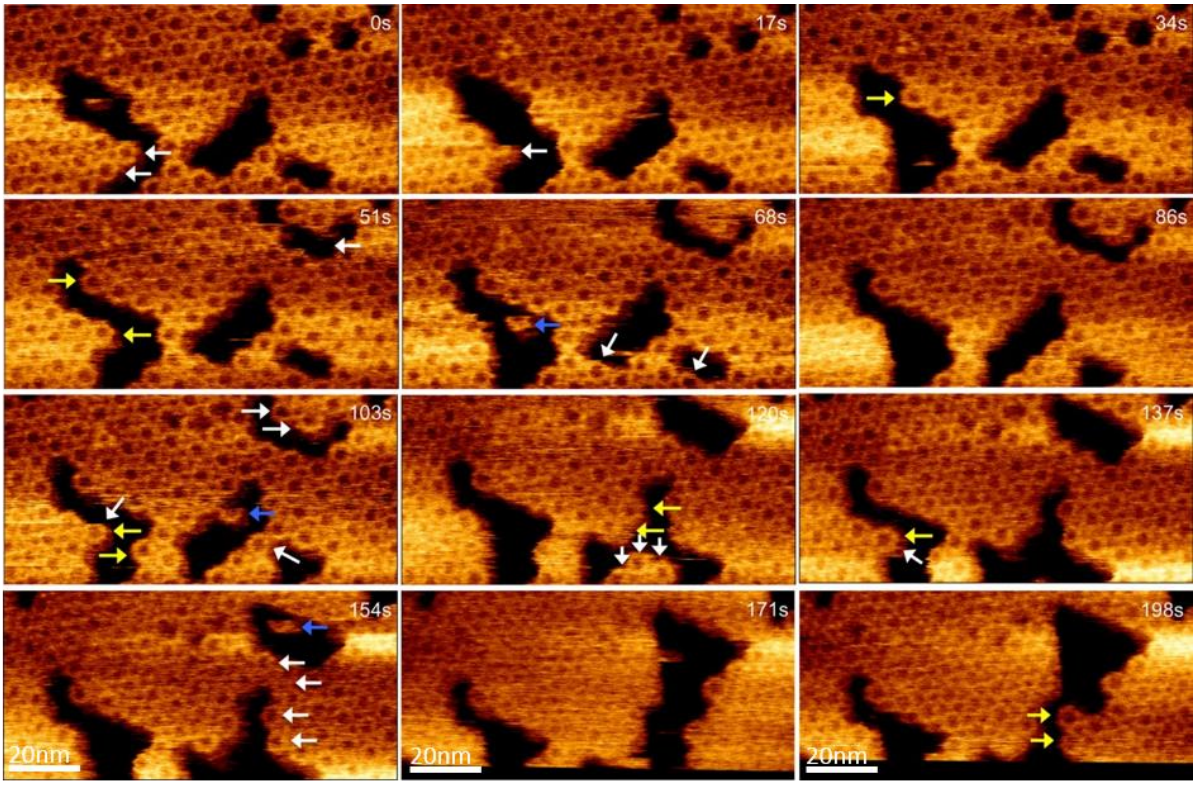


Figure 3-8 HS-AFM imaging reveals the dynamics of HOCH_5815, BMC-H, sheet formation. Hexamers are both removed from (white arrows) and incorporated into the sheet (yellow arrows) during the course of scanning. Blue arrows depict hexamers not associated with the sheet that are translocating across the mica surface. The Twelve aligned HS-AFM images (100 x 47.4 nm) shown were captured at 17 s per frame as part of a 20 minute HS-AFM imaging series.

3.2.6 Observation of alternative types of protein dynamics

Other than individual hexamers moving freely from one location to another within the buffer in a relatively open area of mica inside a patch (Figure 3-8), a range of different dynamic events was also observed. The different kind of events observed appear to depend on the local organisation of HOCH_5815 patches. Fewer dynamic events were observed in more densely packed regions containing more or larger patches. As indicated in the series of images summarised in (Figure 3-13),

the rate of dynamic events is faster within the patch 'inner' and slower along the edge of the patch 'edge'. Where patches have a particularly long straight edge individual hexamers were seen 'rolling' along the edge apparently looking for a site where more than a single contact could be made (Figure 3-9). In patches where this straight edge ends in a pocket where a hexamer can make contact with two others individual hexamers 'roll' until they reach the 'pocket' and then remain there for many subsequent frames (Figure 3-11). Through a combination of these observed behaviours, patches were also notably seen growing rapidly on one edge with several of these 'pockets' compared to a much straighter edge with fewer 'pockets' (Figure 3-12). HOCH_5815 hexamers were also observed disassembling on mass from areas of the patch where they could only make two contacts and reassembling in an area with more pockets where they could make three contacts or more (Figure 3-14). All of these observations suggest that a single face in contact with another hexamer is not a stable interaction, at least two faces are required to be in contact with two other hexamers. BMC_H proteins tend to arrange in patterns where the highest number of contacts can be made, even at the cost of large scale reorganisations. Therefore, the tiled arrays of protein we see all have a similar macro organisation and the likelihood is so will the facets of a BMC as they assemble from BMC-H proteins free in solution.

In subsequent AFM images it was further observed that the inclusion of additional hexamers into existing sheets occurs only when the docking hexamer is in the same orientation as the sheet (Figure 3-15). Hexamers with concave face up could assemble together but could not associate with the hexamers with convex face down. This is most likely ascribed to the particular contacts at the protein interfaces and, moreover, that the same surface orientation of shell proteins is prerequisite for the generation of BMC shells. (Figure 3-15.D,E).

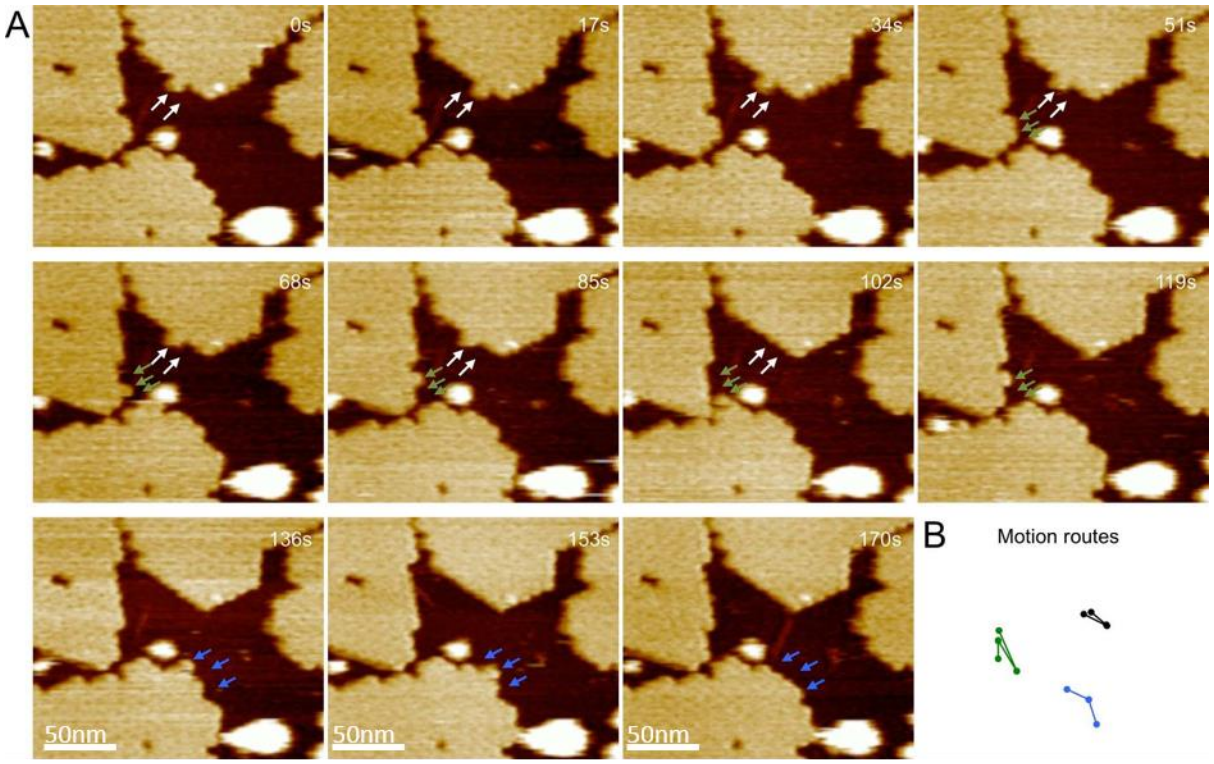


Figure 3-9 Dynamic associations among hexamers at the non-crystalline patch edges. (A) Eleven aligned HS-AFM frames of 250 x 250 nm captured at 17 s/frame. Three individual dynamics events of proteins were visualized (white, green and blue arrows). (B) The motion routes of the three dynamics events in (A) were shown in black, green and blue respectively, by tracking the hexamers between frames.

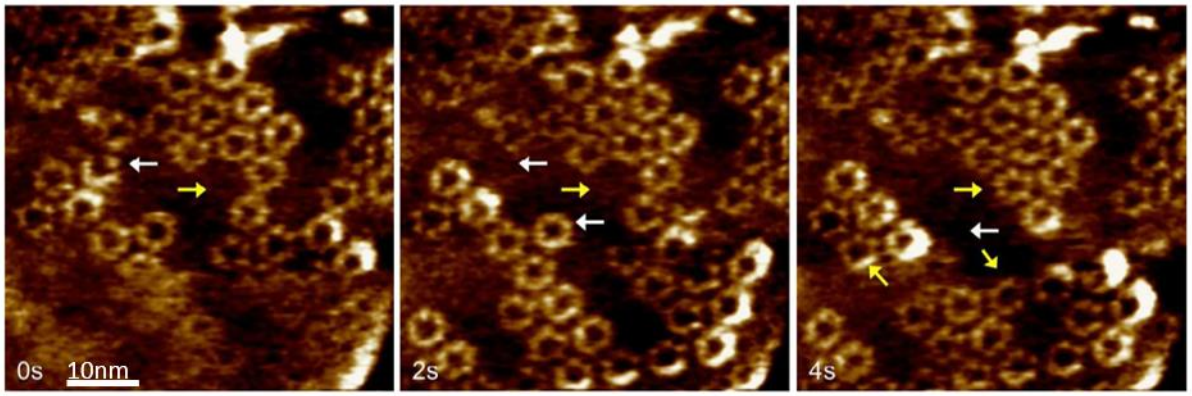


Figure 3-10 HS-AFM images of the self-assembled hexamer sheets. Three aligned HS-AFM images (50 × 50 nm) were captured at higher speed (2 s/frame). The assembly and disassembly of HOCH_5815 hexamers in the patches are shown in yellow and white arrows, respectively.

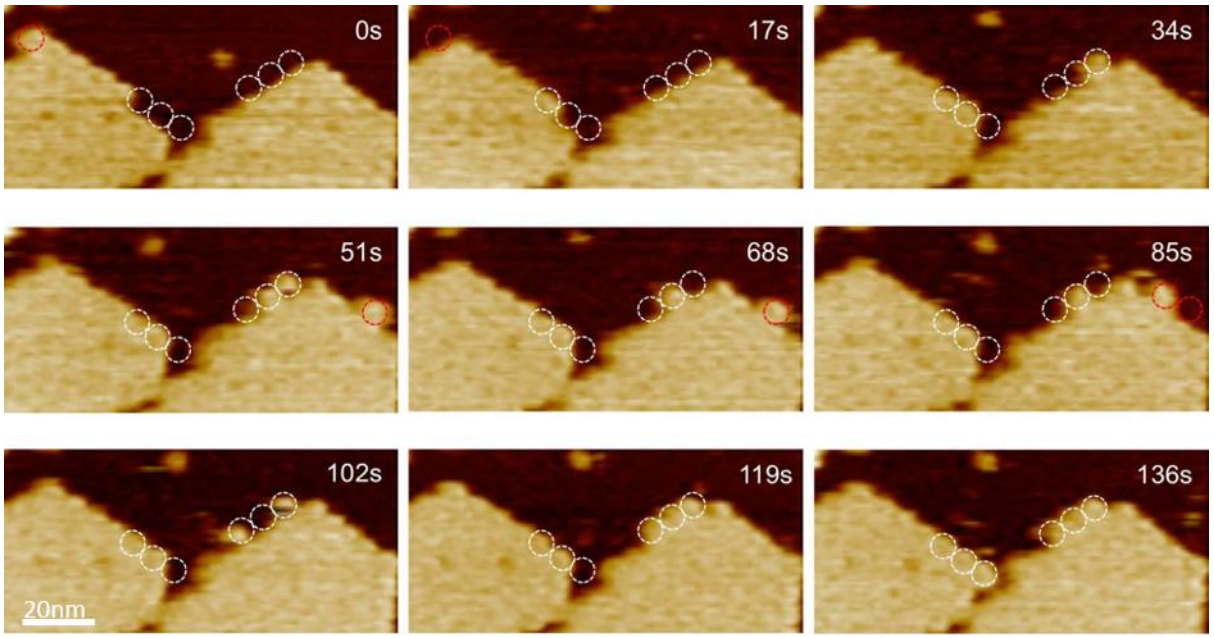


Figure 3-11 Formation of new edges of hexamer sheets by the assembly of individual hexamers. Nine aligned HS-AFM frames were captured at 17 s/frame, 100 x 100 nm. The association of hexamers to the sheet edges and their disassociation from the edges are shown in white and red circles, respectively.

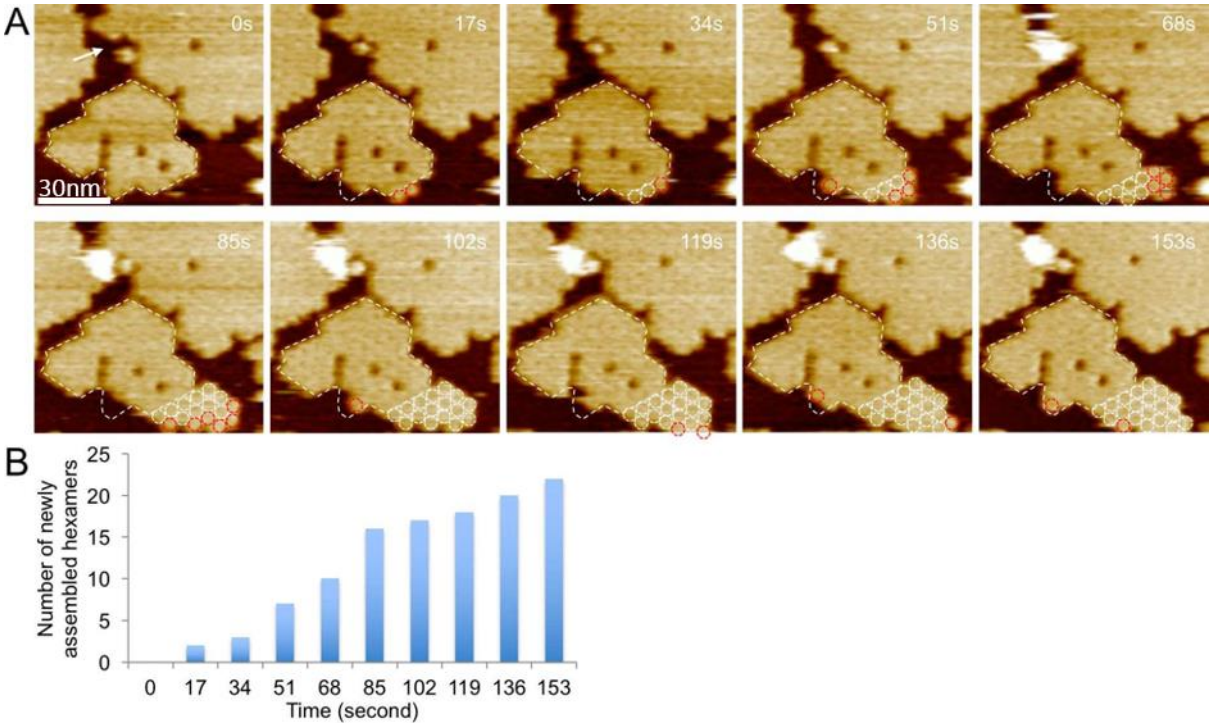


Figure 3-12 Progression of shell patch formation by the assembly of individual hexamers. (A) Ten aligned HS-AFM frames captured at 17 s/frame, 250 x 250 nm. The dashed line indicates the original protein patch at the beginning of HS-AFM imaging. Red circles in the frames indicate the newly assembled proteins, compared to previous frame, whereas white circles show the settled proteins during the assembly process. (B) The number of self-assembled hexamers observed in the patch per HS-AFM frame as a function of time.

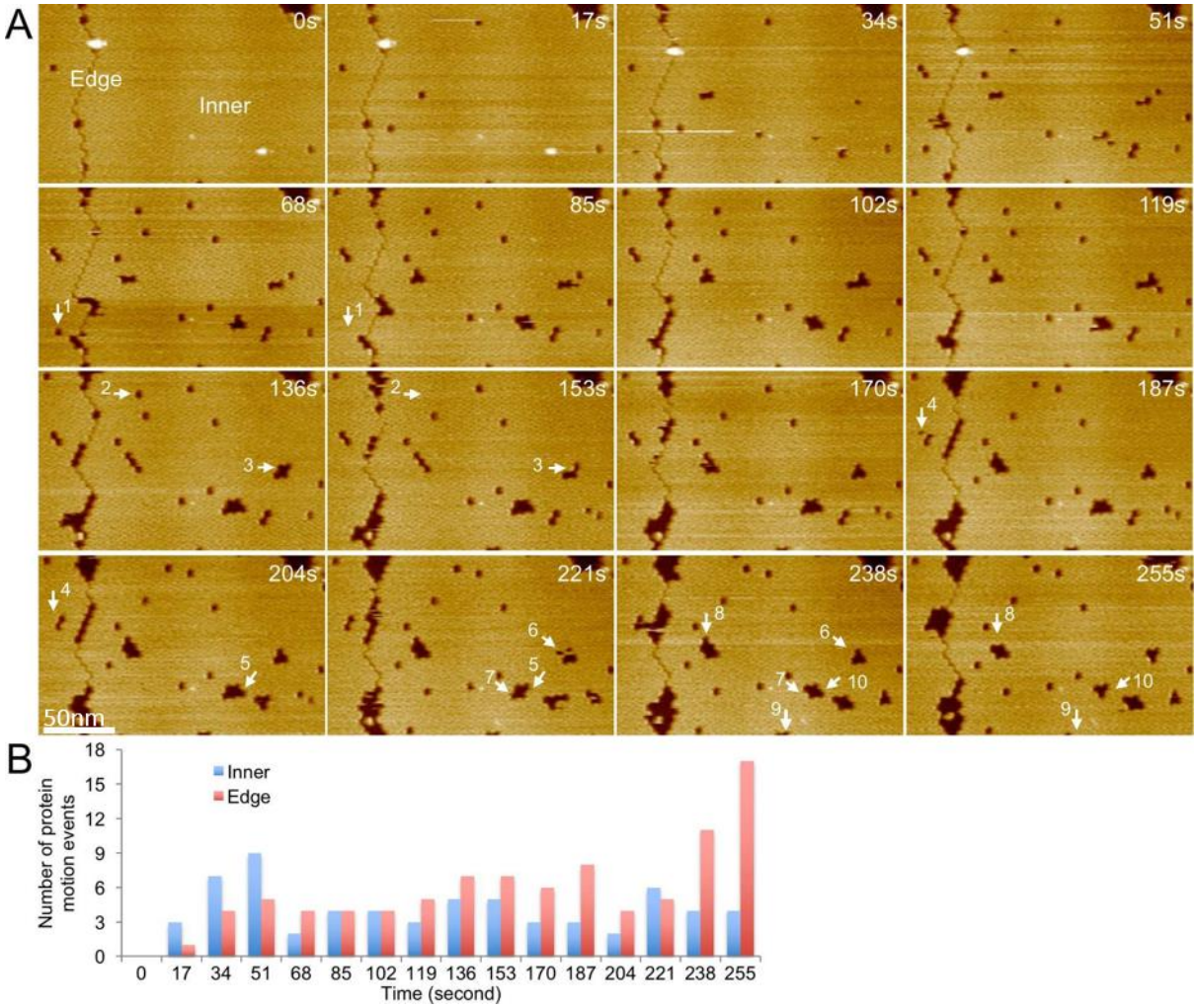


Figure 3-13 The motions of HO hexamers within the protein patch. (A) Sixteen aligned HS-AFM frames captured at 17 s/frame, 300 x 300 nm, display both protein assembly and disassembly events. Arrows represent the assembly events in which diffusing hexamers were trapped by other hexamers within the protein sheet. (B) The number of protein motion events that occurred at both the patch interior and edge per HS-AFM frame as a function of time.

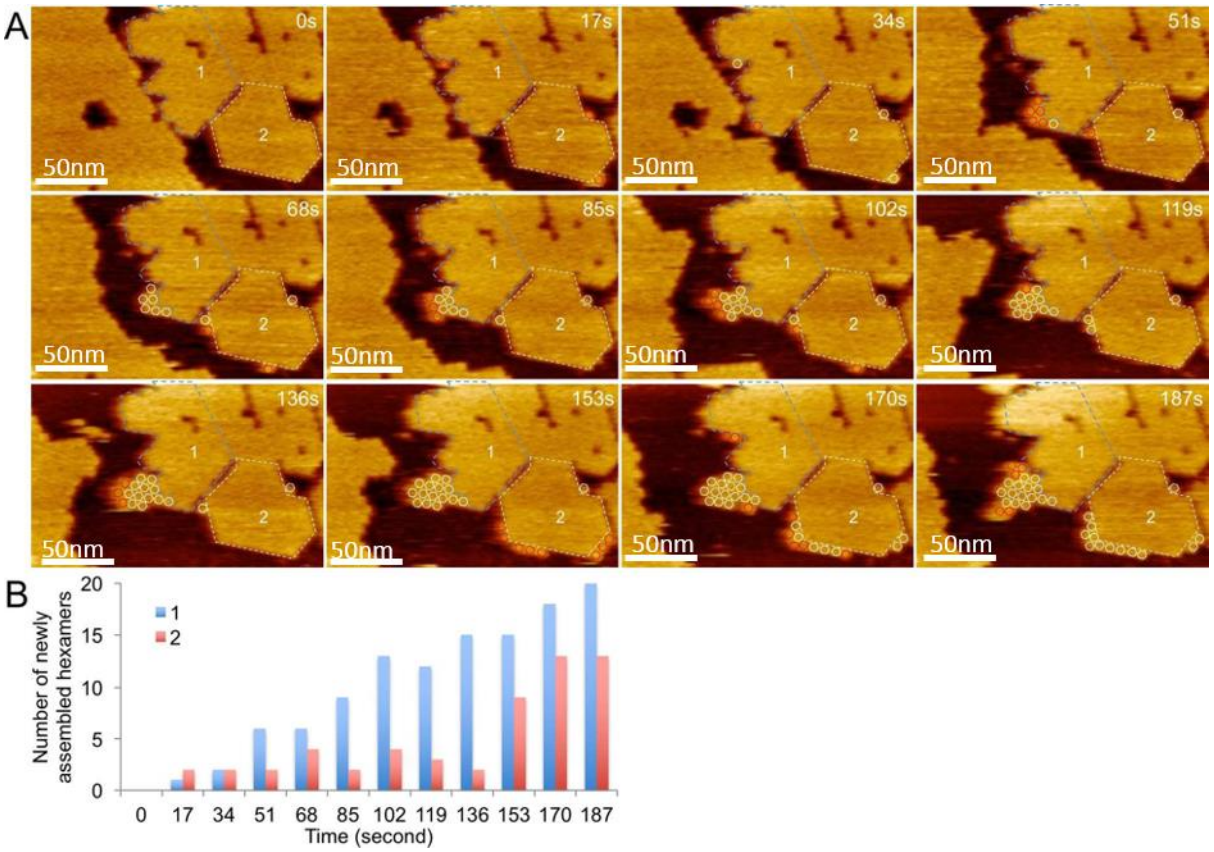


Figure 3-14 Assembly and disassembly dynamics of hexamer sheets. (A) Twelve aligned HS-AFM frames 350 x 350 nm captured at 17 s/frame. Dashed lines 1 and 2 indicate two protein patches that are growing, whereas the large patch on the left side in the frames shows disassociation events. Red circles in the frames indicate the newly assembled proteins, compared to previous frame, whereas white circles show the stationary proteins during the assembly. (B) The numbers of newly assembled hexamers in patch 1 and 2 per HS-AFM frame as a function of time.

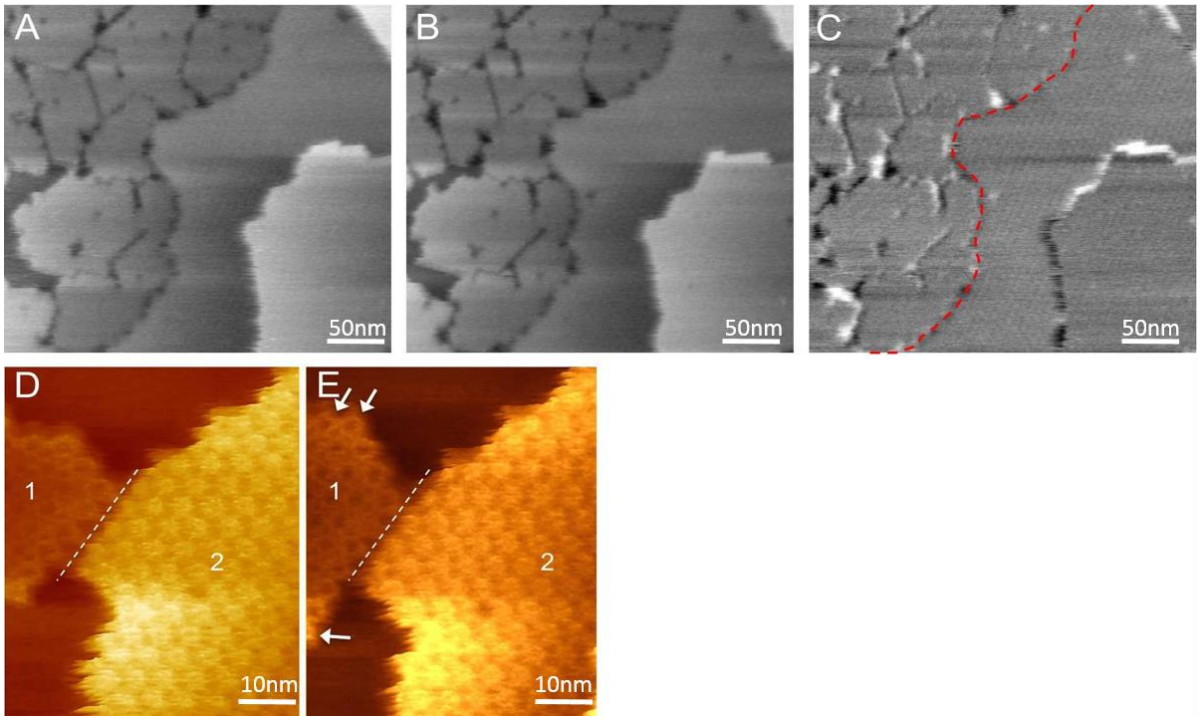


Figure 3-15 The assembly of protein patches depends on the protein orientation. (A) An AFM topograph 500 x 500 nm of protein patches in the same region at 0 min. (B) An AFM topograph 500 x 500 nm of protein patches in the same region at 3 min. (C) The difference AFM image by comparing the two AFM topographs 500 x 500 nm (A) and (B). Black and white areas in the difference AFM image indicate the regions where protein dynamics were detected. The red line represents the interface between sheets with different orientations. The results revealed that the dynamic events were only observed at interfaces between sheets with the same orientations, rather than at the interface between sheets with different orientations. It suggested that the same orientation is essential to ensure the protein sheet development. (D, E) 100x100 nm topographs showing the interface of protein patches (dashed lines) with distinct orientations did not merge during three minutes of AFM imaging. Areas 1 and 2 represent the protein sheets with concave and convex faces exposed to the AFM probe, respectively. Arrows indicate the protein dynamics events captured.

3.2.7 Reorganisation also involves groups of preassembled HOCH_5815 hexamers

The dynamic events previously described in (**3.2.6)** all involve a single hexamer moving by breaking and remaking interactions. Groups of bound hexamers were observed moving together, without losing their interactions with each other. Small 'oddly' shaped patches near a larger more typically organised patch often merged, providing that the two patches were in the same orientation (Figure 3-15). It appears there may be a minimum number of hexamers required for a patch to be stable, under these AFM imaging conditions (Figure 3-16). This merging behaviour may be one of the mechanisms by which the minimum size of a CB facet is maintained.

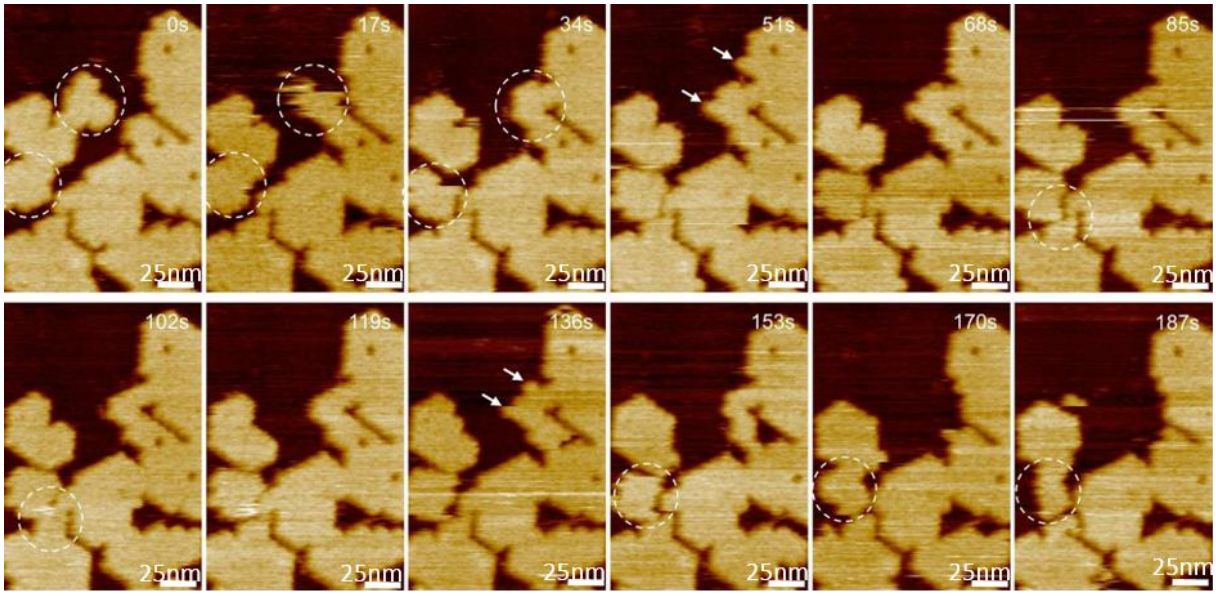


Figure 3-16 Development of shell patch by the merging with other patches. Twelve aligned HS-AFM frames 300 x 300 nm were captured at 17 s/frame. Circles indicate the merging and dynamic large protein patches during HS-AFM imaging. Arrows show the dynamics individual proteins during HS-AFM imaging.

3.2.8 Characterisation of HOCH_5815 mutants

Interestingly, CcmK2 of Syn7942 had been previously expressed in *E. coli*, purified and the structure solved by X-ray crystallography. However, despite this, no similar self-assembled structures were reported in *E. coli* or in vitro, supporting the different assemblies hypothesis (Cai et al., 2015b), although a double layered sheet structure has been reported previously (Samborska and Kimber, 2012). Two residues tha differ between HOCH_5815 and CcmK2 of Syn7942 in the hexamer-hexamer interaction region, K28 and R78, whilst most of the other amino acid residues whose side chains mediate lateral interactions are highly conserved (Kerfeld et al., 2005; Tanaka et al., 2009; Tsai et al., 2009, 2007) (Figure 3-1). To test this hypothesis and the effect of these two specific differences, site directed mutagenesis was used, and is subsequently described in this work.

To begin to dissect the structural determinants governing the self-assembly of shell proteins, single-point mutations of residues at the lateral interface of adjoining hexamers were made, using the PCR based In-Fusion (Clonetech) technique (Allemandou et al., 2003). Residues K28 and R78 are strongly conserved residues in BMC-H proteins and they are located at the junctions between adjacent hexamers in the layers observed in the crystal packing (Figure 3-3). AFM images (Figure 3-16) show that K28A mutant typically forms larger patches (∼1200 nm in diameter), whereas the patches of R78A mutant are much smaller (85 nm in diameter) than those formed by the wild-type (WT) protein (705 nm in diameter). Unexpectedly, the K28A sheets tend to appear as stacked layers, whereas the patches formed by WT or R78A proteins are only single layers (Figure 3-17, C). Molecular-resolution AFM topographs indicate that the double K28A layer is formed by two protein sheets with convex surfaces making contact with each other (Figure 3-18) In addition, the protein dynamics are modified by the K28A and R78A mutations (Figure 3-17, E & Figure 3-19). Calculating the number of hexamer movements in each frame indicates that for many images in a series of images of the K28A mutant there were no dynamic events. There were no frames in which no dynamic events were observed for

the WT or R78A HOCH_5815 (Figure 3-19). The results indicate that K28 and R78 residues of the HOCH_5815 BMC-H protein play key roles in governing the assembly of HOCH_5815 shell hexamers and that the strength of lateral hexamer interactions within the sheets can be modulated without complete disruption. In addition, we found that subtle changes (e.g., a single-point mutation) can promote stacking of the shell protein hexamers.

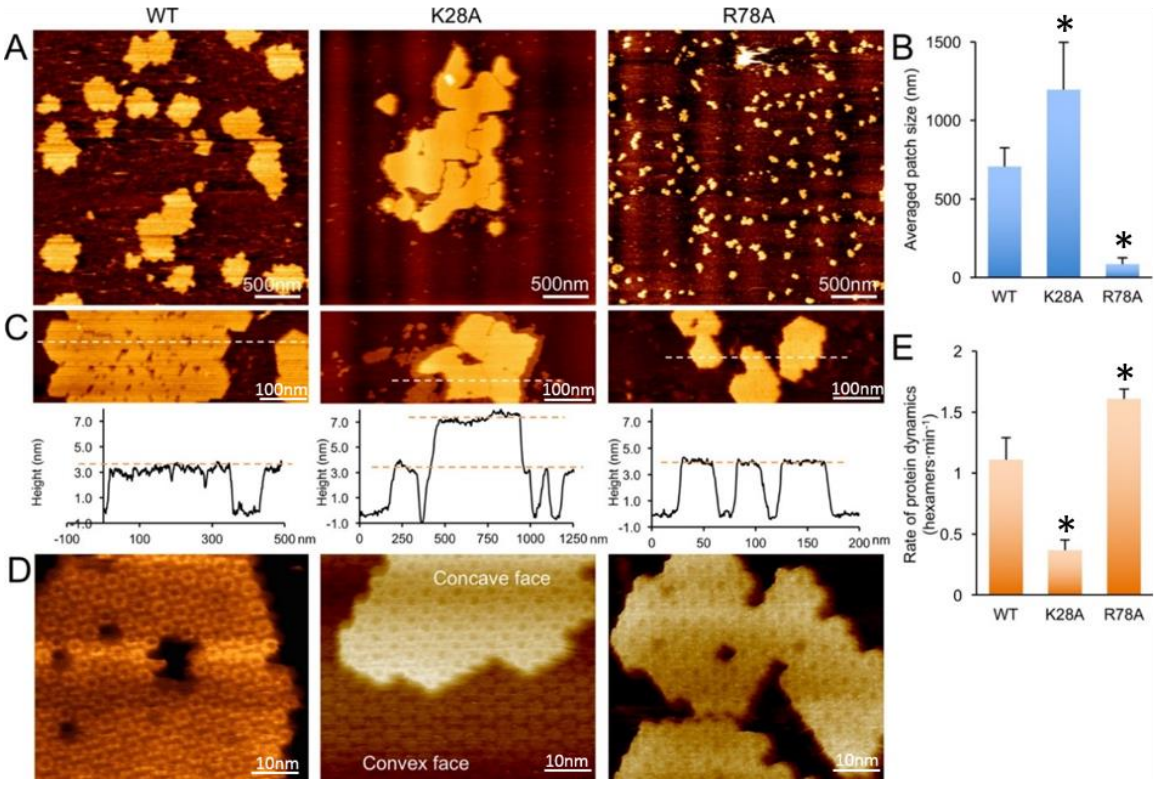


Figure 3-17 Characterization of the impact on assembly of point mutations of the HOCH_5815 BMC-H protein. (A) AFM images of patches of WT, K28A, and R78A BMC-H protein. (B) Quantitative particle analysis shows the average patch size of the R78A mutant (85 nm) is smaller than the WT (705 nm) while that of the K28A mutant (1196 nm) is larger than the WT. (C) Cross-section analysis shows the thickness of sheets formed by the WT and R78A mutant to correspond to that of a single protein layer (3.5 nm). The K28A mutant sheet has thickness consistent with a double layer (7.0 nm). (D) The double layer of the K28A mutant is formed by convex–convex contacts. (E) Normalized rates of protein dynamics in the WT, K28A, and R78A hexamer sheets. The dynamics features of these assemblies are variable R78A proteins in the self-assembled patches present higher translational dynamics than WT proteins, whereas the K28A sheets appear relatively stable during AFM imaging. Measurements were made based on a series of AFM images taken at multiple distinct regions (n = 12). The data were normalized to correct for differences in frame capture time and the scan area which has diverse ratios of protein to mica and the scan size, relative to the WT dynamic events. Error bars represent standard deviation, asterisks indicate statistically significant difference from WT by T-test, $p<0.005$ with 95% confidence.

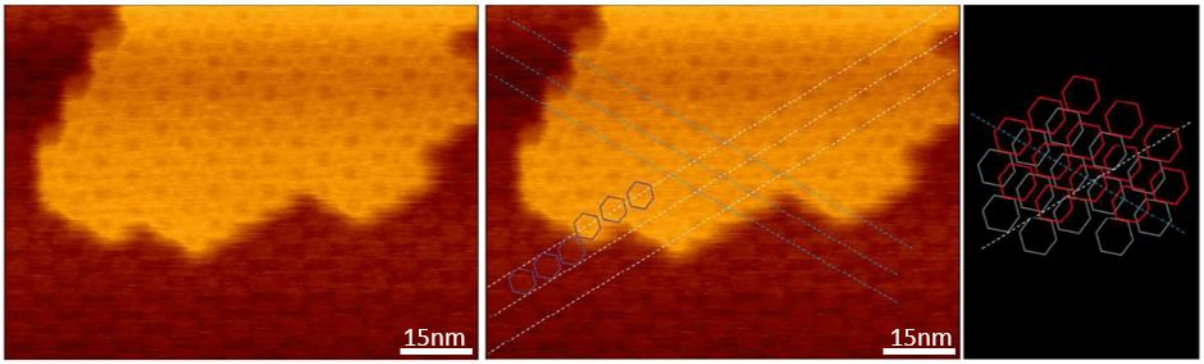


Figure 3-18 Stacking pattern of K28A hexamer double sheets characterized by AFM imaging. White lines show the directions where the protein centres in the upper layer are facing the interfaces of hexamers in the bottom layer, whereas blue lines indicating the centres of proteins in both layers are in identical line. Hexamers in the upper layer are shown in red, those in the bottom layer are shown in grey. The data indicates that there is no direct pore superpositioning between the double shell sheets. The AFM images are 100 x 100 nm.

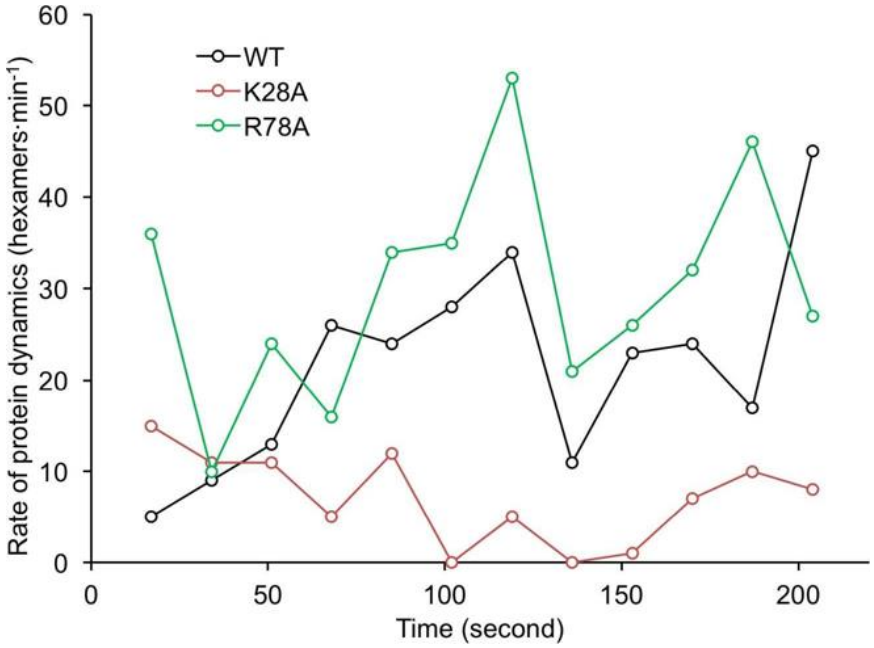


Figure 3-19 The rates of protein dynamics in WT, K28A and R78A HOCH_5815 per HS-AFM frame, as a function of time. (See Figure 3-17) where the data were normalized to correct for differences in frame capture time and the scan area which has diverse ratios of protein to mica and the scan size, relative to WT dynamic events.

3.2.9 Observing the assembly of HOCH_5815 sheets by adding more free protein

The dynamics events described here could be due to AFM tip effects on the HOCH_5815. To test if continuous HS-AFM scanning prevents or hinders patch new protein assembly by sweeping away protein, additional HOCH_5815 was added to the buffer by fluid perfusion during scanning (Figure 3-20). The same patch was then observed for the following 6 minutes. Hexamers were observed assembling around the entire perimeter of the existing patch. The orientation of the patch relative to the motion of the AFM tip had no apparent effect (Figure 3-20).

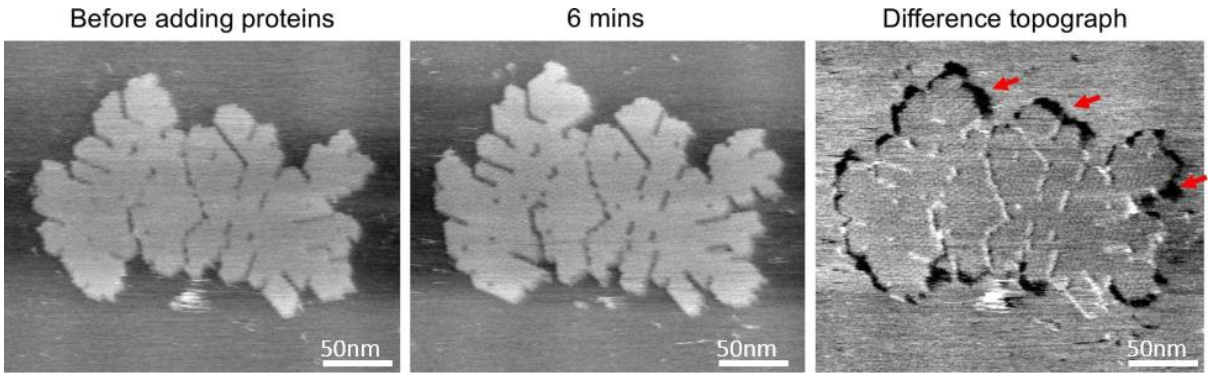


Figure 3-20 The growth of protein patches triggered by the addition of extra hexamers into the AFM imaging buffer. Left, AFM topograph of a large protein patch before adding extra hexamers, via perfusion into the imaging buffer. Middle, AFM topograph of the same protein patch captured at 6 min after adding extra hexamers. After the addition of proteins, a series of AFM images were captured at several locations for 6 minutes, all showing a similar growth in patch size (n = 9). Right, difference AFM image by comparing panel left and middle. Black areas (shown in arrows) indicate the newly developed patch areas.

3.2.10 Ordered orientations of BMC shell superstructures

The clear morphological differences between the concave and convex faces of HOCH_5815 observed by AFM have a definite structural role. The observations that HOCH_5815 hexamers in a sheet are always in the same orientation (Figure 3-4 & 3-5), and that two sheets comprised of hexamers with opposing orientations cannot merge (Figure 3-6 & Figure 3-15), show that the macro-molecular assemblies of HOCH_5815 and possibly all BMC-H proteins must have orientation exclusivity. This is intuitive and logical and has been previously suggested for entire BMC (Kerfeld and Melnicki, 2016; Sutter et al., 2017). The fact that the BMC-H proteins alone assemble into such ordered structures, in the absence of other BMC proteins, makes them uniquely well suited to molecular scaffolding approaches.

3.2.11 Modulated assemblies and dynamics of BMC shell superstructures

The dynamic aggregation of the trp-RNA binding attenuation protein (TRAP) cages has recently been explored using HS-AFM imaging. These cages are self-assembled by the TRAP ring-shaped proteins, reminiscent of the HOCH_5815 hexamers (Imamura et al., 2015). The disassembly of TRAP cages was also monitored In the presence of the disulphide bond reducing agent dithiothreitol. Interestingly TRAP cage formation could not occur on mica, unlike the dynamic assembly of BMC shell proteins on mica surface observed in this study, this difference is likely due to the different capacity for interactions between these proteins and the substrate. Nevertheless, both studies on TRAP cages (Imamura et al., 2015) and BMC-H hexamer sheets (in this work), demonstrated the highly dynamic nature of the building blocks of large protein complexes.

3.2.12 Preferential hexamer docking during sheet formation

The HS-AFM data in this study show the first clear evidence for preferential docking and local ordering of BMC-H hexamers during the self-assembly of a BMC shell facet. At least two hexamer-hexamer lateral interactions are required for an individual hexamer to be stable, within a facet-like sheet environment. This suggests that BMC-H proteins may only dock in the central region of the facet, most resembling the flat-sheets or tiled arrays seen here, during BMC assembly.

Similar HS-AFM experiments were conducted involving a mixture of different BMC-H proteins from β-CBs (Garcia-Alles et al., 2017), where the HOCH_5815 paralog hexamers alone formed similar tiled arrays, and the other BMC-H hexamers were found to introduce curvature. This would be interesting to study in the HOCH system in future work, to see if the same holds true or if this differs between the various types of BMC. In another study (Sutter et al., 2017), the same HOCH_5815 mixed with BMC-T and BMC-P proteins from HOCH formed a 6.5 MDa shell. The BMC-T and BMC-P proteins are responsible for introducing the distinct curvatures expected at the 2-fold and 5-fold axis of symmetry, in the icosahedral BMC structure, and that the BMC-H proteins predominantly form the flat regions of the facet (Sutter et al., 2017). It would also be interesting in future work to investigate these mixtures, and the stepwise self-assembly of the 6.5 MDa shells, using HS-AFM.

These same kinds of investigations of BMC-H self-assembly can be performed in silico, using coarse grain molecular dynamic simulations, another study observed many of the same self-assembly principles using this computational approach (Mahalik et al., 2016). It would be interesting to apply this approach to the mixtures of HOCH BMC proteins or the formation of the 6.5 MDa shell (Sutter et al., 2017).

3.3.3 Concluding remarks

In summary, these data provide insights into the organisation of BMC shell facets in terms of the formation, interactions, and dynamics at the single molecule level. This study shows that preassembled shell hexamers form single layer sheets of uniform orientation. In addition to this information regarding the arrangement of hexameric proteins, the HS-AFM data indicates an overall flexible intermolecular interaction as it shows that individual hexamers can dissociate from and incorporate into assembled sheets. These data also show that specific contacts at the interfaces of neighbouring proteins influence the dynamic features of shell proteins, and thereby the self-assembly of the shell facets. The design and construction of synthetic BMCs have attracted intense interest for the bioengineering of nanoreactors and molecular scaffolds. Understanding the details of self-assembly of BMC shell proteins is a prerequisite for control and engineering of BMC-based architectures with the aim of building designed nanoreactors and molecular scaffolds. This work, in particular the site directed mutagenesis has elucidated more details surround BMC-H self-assembly and has informed us of the role these specific conserved residues play in controlling it.

This page is intentionally blank and marks the end of chapter three.

*"**Biology is the most powerful technology ever created**. DNA is software, protein are hardware, cells are factories." **Arvind Gupta**.*

Chapter 4 Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes

4.1 Introduction

4.1.1 The current model of Syn7942 CB structure

The model cyanobacterium *Synechococcus elongatus* PCC7942 (Syn7942) contains β-CB (Section **1.1.2**, Figure 1-4). To date, models of the β-CB are based on crystal structures of individual β-CB proteins with the assumption of icosahedral symmetry (Table 1-2, Tanaka et al., 2008). The molecular details of the β-CBs' structure remain unclear. Therefore, in this work I intend to improve upon previous purification methods and obtain more meaningful TEM data from intact Syn7942 β-CB.

Three distinct assembly pathways of CB modules have been deduced. In Syn7942, *de novo* assembly of β-CBs exploits the "inside out" mode, Rubisco and CcmM first forming the core, followed by the encapsulation of this core by the shell proteins (Cameron et al., 2013; Chen et al., 2013). In contrast, the formation of empty α-CB shells in a Rubisco-knockout mutant of *H. neap* led to the implicit assumption that the shell forms first during α-CB biogenesis (Baker et al., 1998; Menon et al., 2008). In addition, partial α-CBs composed of the fractional shell and attached layers of Rubisco enzymes were imaged in *H. neap* and no Rubisco aggregations were observed (Iancu et al., 2010), suggesting a simultaneous assembly pathway for CB biogenesis.

Within the cytosol, which is a crowded and changing environment (Parry et al., 2014), it is important that CBs are sufficiently robust to ensure the proper protein assembly, encapsulation of Rubisco enzymes and functional architecture. On the other hand, they are also flexible and dynamic to allow metabolite passage, turnover of building modules and interactions with other cellular components. Indeed, protein modules in the BMC shell facet are highly dynamic (Sutter et al., 2016). Through specific interactions with the

cytoskeleton, β-CBs in Syn7942 are evenly positioned along the longitudinal axis of the cell, ensuring equal segregation of these essential organelles to daughter cells (Savage et al., 2010). The biosynthesis and spatial organisation of β-CBs in Syn7942 also have a close correlation with photosynthetic electron flow regulated by light (Sun et al., 2016). In such a dynamic context, the inherent physical properties of CBs are important for the structural and functional integrity and flexibility of these icosahedral organelles. Until now, the exact mechanical nature of CBs has not been characterized.

4.1.2 NI studies on viruses and encapsulin nanocompartments

BMC shells appear similar to a viral capsid, but are CB shells mechanically similar to viral capsids due to the shared icosahedral structure? NI has also been widely applied to viruses, encapsulin, and other particles similar to BMCs (Bolhassani et al., 2017; Boyd et al., 2015; Carrasco et al., 2006; Mateu, 2012; Snijder et al., 2013) (Table 4-1).

Table 4-1 A summary of the previously reported physical measurements made on viruses, using AFM NI.

K(particle) (N/m)	YM (MPa / GPa)	Failure (nN)	Height (nm)	type	Ref
0.25	0.05 GPa	0.6	24	encapsulin	(Snijder et al., 2016)
0.22	n/a	0.7	40	virus	(Hernando-Pérez et al., 2016)
0.56	0.3 MPa	n/a	80	virus	(Ortega-Esteban et al., 2015)
0.21	n/a	1.6	50	virus	(Boyd et al., 2015)
0.20	140 MPa	0.81	30	virus	(Michel et al., 2006)
0.10	0.37 GPa	0.98	25	virus	(Uetrecht et al., 2008)
0.43	n/a	2.01	90	virus	(Snijder et al., 2013)
0.04	n/a	n/a	100	virus	(Li et al., 2014)
0.3	1.2 – 1.6 GPa	0.6	50	virus	(Ivanovska et al., 2004)
0.23	n/a	1.6	40	virus	(Ivanovska et al., 2007)
0.12	0.9 GPa	0.9	70	virus	(Roos et al., 2012)
0.8	1.25 GPa	n/a	60	virus	(Carrasco et al., 2006)

A summary of the previously reported physical measurements made on viruses, using AFM NI. Where a value was not reported in a given study n/a has been indicated. The values given are approximate, rounded to 2 decimal places For K, E & failure and 1 nm for height, for the purpose of simplicity in comparison.

4.1.3 The scope of this study

In this work we purified functional β-CBs from Syn7942. This enabled the first detailed characterization of the three-dimensional structure, topography and intrinsic nanomechanics of native β-CBs, using transmission electron microscopy (TEM), atomic force microscopy (AFM) and proteomics. These results revealed three distinct structural features of intact β-CBs that were previously unobserved, the native protein organisation of the shell, the specific protein interactions in partial CBs and the internal organisation of the luminal proteins. Though structurally resembling viral capsids, β-CBs present significantly softer mechanics. This study provides novel insights into the inherent structural and physical elasticity of native β-CBs. It will empower our toolbox for the design and construction of functional metabolic machinery, this has applications in bioengineering and nanotechnology.

4.2 Results and discussion

4.2.1 Optimising growth conditions to obtain the maximum yield of CBs in the starting material

A CcmK4 eGFP Syn7942 strain was used to develop the procedure for β-CB purification. The eGFP tagging, with undetectable effects on the β-CBs' structure and physiology (Sun et al., 2016), enabled us to fluorescently screen the β-CBs' fractionation during the isolation and characterization processes. Syn7942 cells were grown under high light (~100 μE m^{-2} s^{-1}) to increase the CB abundance per cell, according to the findings of my co-authored previous study (Sun et al., 2016). A gradual increase in light intensity was found to increase the growth rate of CcmK4 eGFP Syn7942 (See Syn7942 photobioreactor growth curves in **Appendix B**) as opposed to the longer lag phase then very rapid growth in a sudden shift to high light reported previously (Sun et al., 2016). Crucially this still increased CB numbers by a similar amount per cell (Sun et al., 2016). By increasing mixing and aeration in a similar stepwise fashion the growth rate of CcmK4 eGFP Syn7942 was further improved (**Appendix B)**, similar to previous growth optimisation studies (Kuan

et al., 2015; McEwen et al., 2013; Sarnaik et al., 2017; Yang et al., 2015). By increasing the CB number cell, the growth rate, and the maximum attainable culture density it was possible to increase the amount of CB in the starting material for the purification.

4.2.2 Determining the optimal point in the growth cycle for CB isolation

Syn7942 cell morphology changes in response to environmental stimuli and quorum sensing, division stops and cells elongate (Hood et al., 2016). Once Syn7942 cultures approach stationary phase and begin to elongate, despite the increase in biomass, the yield of CBs from this purification procedure decreases. Isolation of CBs from 1 litre aliquots of photo-bioreactor Syn7942 culture, every 24 hours from 2-7 days post inoculation, determined that under these culture conditions day 5 gave the greatest yield (see growth curves in **Appendix B**). The change in Syn7942 cell morphology may result in increased CB-cytoskeleton interaction (Savage et al., 2010; Sun et al., 2016) and therefore less efficient separation of CB.

4.2.3 Determining isolated CB purity and function

Following Triton X-100 treatment, β-CBs were enriched in the pellet by two steps of centrifugation and many cellular components remained in the supernatant. After sucrose gradient centrifugation (Figure 4-3, A), most of the rest cellular components appeared in the top and pellet of sucrose gradient fractions by proteomics and TEM (data not shown). The majority of β-CBs were determined in the 20%, 30% and 40% fractions by fluorescence imaging (Figure 4-1, B). Most of the strong GFP spots appear in the 40% sucrose gradient fraction (Figure 4-3, B). SDS-PAGE illustrates the polypeptide patterns of β-CB in each fraction (Figure 4-1, C). Rubisco enzymes are the most abundant components in all fractions, in agreement with immunoblot data (Figure 4-2). Carbon fixation assays of each β-CB sucrose fraction reveal that the 40% fraction presents the highest Rubisco activity (Figure 4-1, D). Proteomic analysis (in collaboration with the Centre for Proteome Research) of the 40% fraction allows the identification of a total of seven β-CB components, including the shell proteins (CcmK2, CcmK4, CcmL), shell-associated proteins (CcmM, CcaA) and internal

proteins (RbcL, RbcS) (Table 4-2 & Table 4-3). These results verify the proper fractionation of functional β-CBs from Syn7942 (40%), whereas the 20 and 30% fractions may contain β-CB subcomplexes.

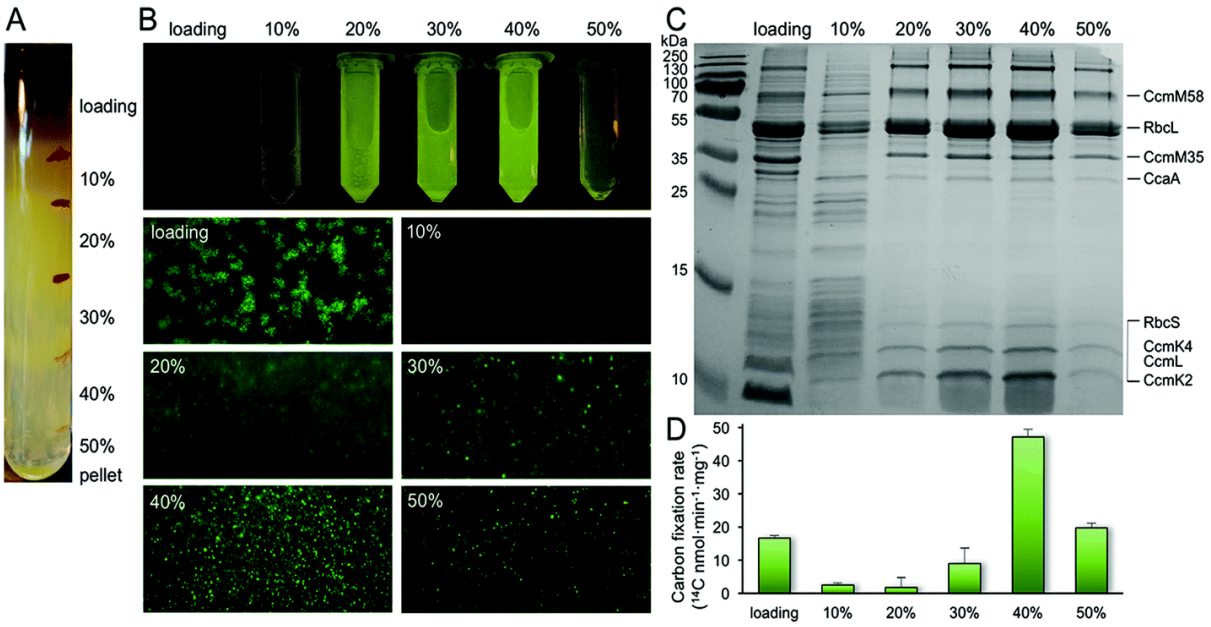


Figure 4-1 Isolation and characterization of CcmK4 eGFP β-CBs from Syn7942. (A) Step sucrose gradient separation of CcmK4 eGFP β-CBs. (B) Fluorescence detection of β-CBs fused with GFP in different sucrose fractions. (C) SDS-PAGE of individual fractions from the β-CB purification, showing the polypeptide composition of isolated β-CBs. The presence of Rubisco was verified by immunoblot analysis (Figure 4-4). Determination of β-CB proteins was confirmed by proteomic analysis (Tables 4-2 & 4-3), in collaboration with the Centre for Proteome Research. (D) Rubisco specific activity of each sucrose fraction as determined by ^{14}C radiometric assay.

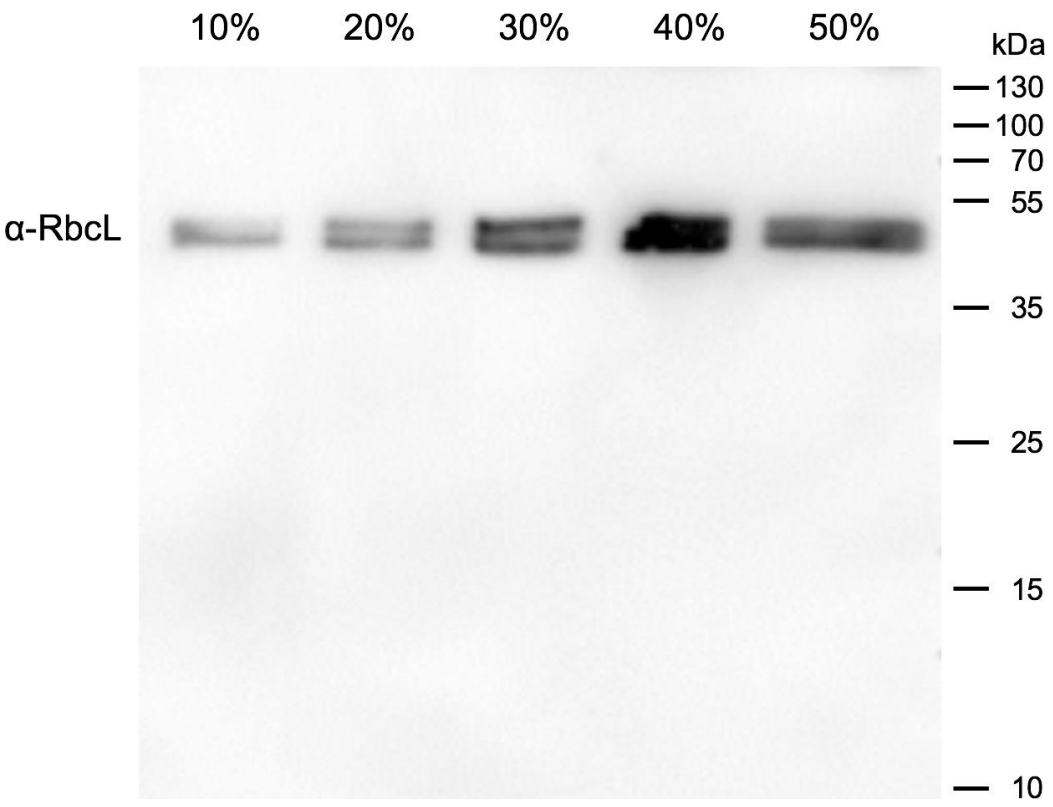


Figure 4-2 Immunoblotting analysis of different β-CB fractions using anti-RbcL antibody. Immunoblotting assays were carried out on the SDS-PAGE gel shown in Fig. 1C. RbcL (~50 kDa) was detected in all sucrose fractions and was most abundant in the 40 % sucrose fraction.

In addition to Rubisco, CcmM was also relatively abundant in the β-CB (Table 4-3). This is in agreement with their deduced roles, interacting with the shell and Rubisco and in interlinking Rubisco enzymes and thus forming the paracrystalline arrays (Rae et al., 2013). We could not differentiate the two CcmM isoforms, CcmM58 and CcmM35, in the isolated β-CBs, due to the absence of specific peptide sequences identifiable by mass spectroscopy in the N-terminus of CcmM58. The minor shell protein CcmL was identifiable in the isolated β-CBs (Table 4-3). According to the icosahedral shape there are 12 vertices, and therefore twelve CcmL pentamers, per CB. Although CcmO was deduced to occupy 10–30% of the shell s,(Rae et al., 2012) it was not detectable in the isolated β-CBs by mass spectroscopy. CcmN, CcmP and RbcX were likewise not detected, either in this work or in the previous study(Long et al., 2005). This is indicative of their low

abundance in the β-CB (compared to CcmL), the weak interactions with other CB proteins, or changeable CB composition in different conditions. Further exploration is needed to examine the accurate stoichiometry and function of these undetectable components in β-CBs.

Apart from the predominant β-CB components, four cytoskeletal proteins (ParA, MreB, FtsZ, Ftn2) were identified in relatively high abundances in the 40% fraction (Table 4-2), supporting the notion that there are inherent strong interactions between β-CBs and the cytoskeleton. These interactions are thought to be key to the spatial positioning of β-CBs in Syn7942 (Savage et al., 2010). It is feasible that the GFP tags of CcmK4 eliminate potential associations between β-CBs and other cellular structures somehow, *albeit* the underlying mechanism remains unclear.

Table 4-2 Protein categories identified by TOP3 MS analysis of the isolated CB sample from Syn7942.

Protein category	Number of identified proteins
Carboxysome proteins	8
Cytoskeletal proteins	4
Bicarbonate binding protein	1
Photosynthetic proteins	48
Transporters	9
Ribosomal proteins	52
DNA binding proteins	51
tRNA protein	17
Circadian proteins	3
Transferase	27
Kinases	17
Hypothetical or uncharacterised proteins	100
others	249
Total	496

A table showing the protein categories identified by TOP3 MS analysis of the isolated CB sample from Syn7942. The proteins described here were all detected by at least one unique peptide by mascot search and represent >1% instrument detection sensitivity. The quantitative data set for the CB proteins identified is available Table 4-3. The functions ascribed here are the function as listed on the uniprot database. 3 of the 5 most abundant proteins were CB proteins (RbcL, RbcS & CcmM) and all the CB proteins except CcmL and CcmK4 are in the top 150 in terms of abundance indicating that the most abundant proteinaceous species in the sample was the CB and the sample was reasonably pure for later experiments. These data were collected in collaboration with the Centre for Proteome research.

Table 4-3 Top 3 Proteomic results of isolated β-CBs from Syn7942.

Protein	Normalized amount (fmol)	Molar concentration (M)	Estimated number of functional units per β-carboxysome
RbcL	4530.8 ± 432.8	86.5 ± 37.7	67962.2 ± 4315.3
RbcS	1744.5 ± 765.1	131.2 ± 10.7	25386.6 ± 5193.1
CcmM	1567.1 ± 412.8	41.7 ± 24.3	84007.7 ± 11329.3
CcmK2	116.6 ± 22.0	10.7 ± 5.8	1117.7 ± 106.8
CcaA	81.8 ± 12.6	2.7 ± 0.4	2439.1 ± 299.4
CcmK4	18.7 ± 0.4	1.6 ± 0.7	206.2 ± 0.6
CcmL	1	0.091	12

A table showing the Top 3 Proteomic results of isolated β-CBs from Syn7942. The column "Normalized amount" displays the amount of each of the carboxysomal proteins detected in isolated β-CBs using TOP3-MS, converted to fmol using the intensity of the known standard peptide, normalized against the amount of the least abundant carboxysomal protein CcmL. "Molar concentration" was calculated considering the normalized amount and mass of each protein from the Top 3 semi quantitative abundance data. The "estimated number of each functional unit per β-CB" was calculated by dividing the concentration of monomers (molar concentration) by the aggregation states (i.e. 6 for hexamers, 5 for pentamers etc.) and then normalized based on the notion that there are 12 CcmL per β-CB.

4.2.4 Characterising CB sub-complexes and potential assembly intermediates

The structures of isolated β-CBs were examined using TEM and AFM. Electron micrographs of negatively stained specimens demonstrate that the 20 and 30% sucrose gradient fractions contain predominantly the β-CB substructures (Figure 4-3, A, and Figure 4-4). Shell facets with straight and regular edges as well as proteins attached to the shell were visualized. AFM imaging in solution was used to characterize the native topography of β-CB subcomplexes at near physiological conditions (Figure 4-3, B, D & E, and Figure 4-5, B, C & D). Cross-section analysis reveals that the thickness of these CB fragments is 18.03 ± 8.11 nm (*n* = 20), with a range from 12.1 to 25.3 nm (Figure 4-3, C). They are thicker than a single shell protein layer that is about 4.0 nm thick (Kerfeld et al., 2005; Sutter et al., 2016). Three-dimensional AFM image and cross-section analysis suggest that the β-CB fragment observed is composed of two shell facets with a joint edge that is raised from the AFM substrate surface (Figure 4-3, C & D). Individual shell hexamers and their spatial organisation in the shell facets could be seen (Figure 4-3, E, & Figure 4-5, B & D).

In addition to the relatively flat shell sheets, more curved shell fragments were also imaged in solution. One curved shell patch was observed, where the native large-scale organisation of shell hexamers can be viewed in particularly high resolution (Figure 4-5, A & B). This organisation is reminiscent of the organisation of shell hexamers in synthetic BMC-H shell self-assemblies observed using HS-AFM (Sutter et al., 2016). Cross-section analysis reveals the highly regular periodic arrangement of shell hexamers and their centre-to-centre distance is ∼9 nm (Figure 4-5, B & C). The protein structures and arrangement in the shell were better discerned in the 3D height AFM image (Figure 4-5, D), where the shell fragment approximately resembles a pyramidal shape.

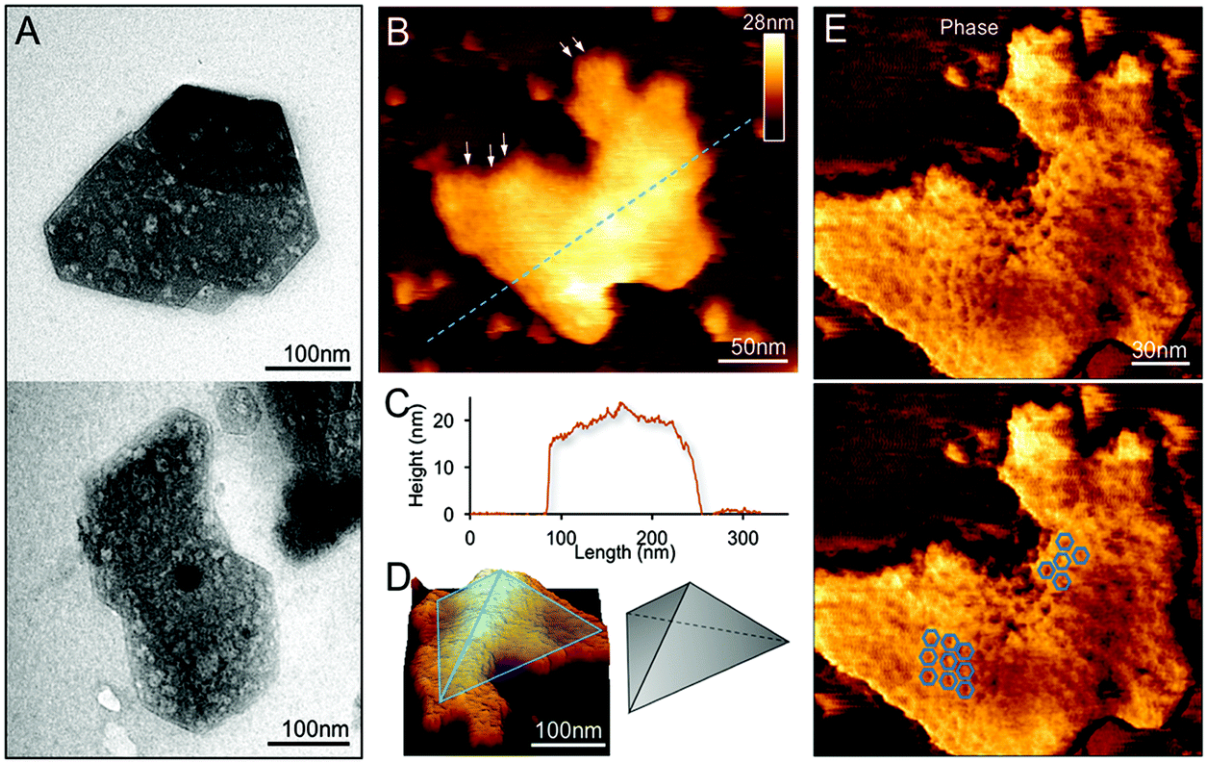


Figure 4-3 Characterization of the β-CB fragments in the 20% and 30% fractions. (A) TEM images of β-CB fragments captured in the 20 and 30% fractions. Regular and straight facet edges and proteins associated with the facets were observed, more TEM images are shown in (Figure 4-4). B) AFM topograph of a typical β-CB fragment illustrating the spatial organisation of individual shell proteins (indicated by arrows). An AFM topograph of a larger, curved β-CB fragment is shown in (Figure 4-6). (C) Cross-section analysis of the β-CB fragment along the dashed line indicated in (B). (D) 3D representation of the β-CB fragment, showing the possible shell substructure comprises two shell facets that have a joint facet edge. (E) AFM phase image recorded together with the height image (B), displaying the native protein organisation in the shell facets, with patterns of individual shell hexamers highlighted in blue hexagons.

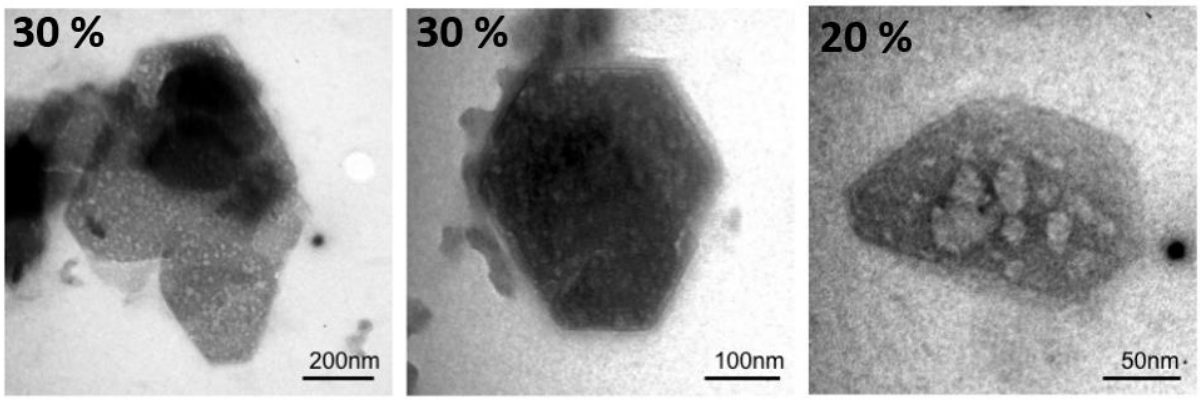


Figure 4-4 TEM images of partial β-CB fragments in the 20 and 30% sucrose fractions. The images left and middle were captured from the 30 % fraction, right was captured from the 20 %. There are clearly resolved luminal proteins still associated in the middle of these fragments, in particular in the right hand image. The larger fragments form the 30 % (left and middle) appear to have large aggregates of what may be luminal proteins associated (heavily stained dark clouding), not visible in the smaller fragment from the 20 % fraction (right).

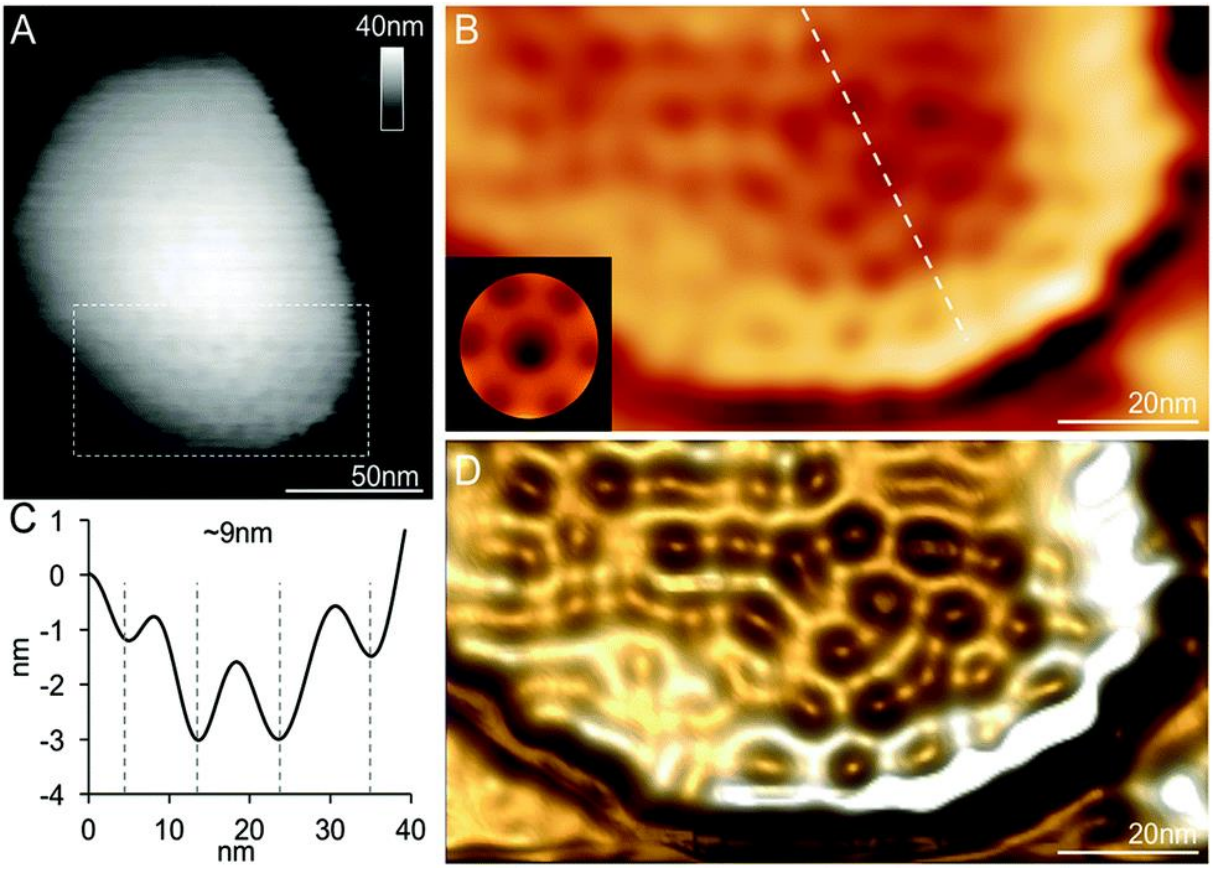


Figure 4-5 spatial organisation of proteins in a partial β-CB from the 30% fraction. (A) High-resolution AFM topograph of a partial β-CB fragment in imaging buffer. (B) High-pass-filtered AFM image showing the protein organisation in the shell fragment depicted in (A). The white line of the cross-section was used to calculate the pattern of hexamer organisation in (C). (C) The cross-section profile illustrates the periodic arrangement of hexamers and the centre-to-centre distance between neighbouring hexamers is ~9 nm. (D) Three-dimensional height image of the shell showing the shell protein structures and arrangement.

4.2.5 TEM and AFM imaging of purified CBs

These EM and AFM results, together with the SDS-PAGE (Figure 4-1, C), Rubisco assay (Figure 4-1, D) and immunoblot analysis (Figure 4-2), reveal that the observed specimens are partial β-CB modules (12.1–25.3 nm thick) that comprise shell facets, plus shell-associated proteins and 1–2 layers of Rubisco enzymes given that 1 L8S8 Rubisco octamer would add between 9 and 11 nm in height (PDB 1RBL and 3ZXW)(Taylor et al., 2001). Despite the possible artefacts in sample purification, these β-CB substructures resemble the partial α-CBs observed previously (Iancu et al., 2010) probably acting as intermediates generated in the β-CB biogenesis or degradation pathways.

In contrast, EM images of the 40% fraction show the regular and polyhedral shape of intact β-CBs from Syn7942 (Figure 4-6 and Figure 4-7). These organelles exhibit an average diameter of 149.90 nm (Figure 4-6, B), & Table 4-5), larger than the isolated α-CBs from *H. neap*, (Schmid et al., 2006) *Synechococcus* WH8102 (Iancu et al., 2010, 2007), and *Prochlorococcus marinus* MED4 (Iancu et al., 2010) Table 4-4). Interestingly, the size of isolated β-CBs is slightly smaller than that determined from previous thin-section TEM results (Rae et al., 2012; Whitehead et al., 2014). Nevertheless, unlike typical icosahedral viruses, β-CBs vary in size, ranging from 100 to 200 nm (Figure 4-6, B), consistent with the observations from *in vivo* confocal fluorescence microscopy and TEM results (Cameron et al., 2013; Long et al., 2011, 2010, 2005; Rae et al., 2012; Sun et al., 2016). The structural heterogeneity implicitly indicates the inherent dynamics of β-CB formation and biogenesis *in vivo*, which might be of physiological importance to the generation of new β-CBs from pro-CBs or pre-existing CBs and the degradation of mature β-CBs during cell growth and division (Cameron et al., 2013; Savage et al., 2010). Moreover, two closely associated β-CBs were occasionally seen (Figure 4-6, C & Figure 4-7, B). Despite the possibility of being artefacts in sample preparation, whether they are generated by potential interactions between neighbouring CBs or in the budding events (Cameron et al., 2013) remains unknown.

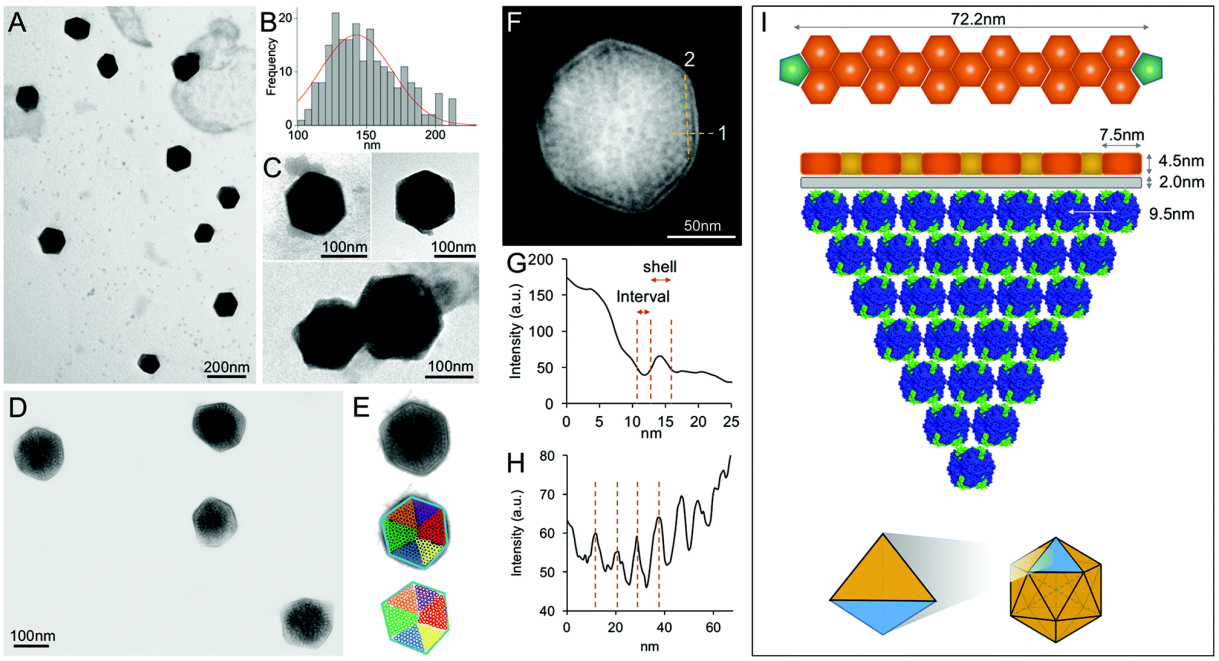


Figure 4-6 Characterization of the intact β-CBs in the 40% fractions. (A) An overview TEM image of the 40% sucrose fraction showing individual β-CBs with the polyhedral shape. (B) Histogram of the diameters of β-CBs measured from TEM images shows the size heterogeneity of β-CBs (*n* = 90). Each measurement is the mean of the three vertex-to-vertex measurements from a single CB as described in (Figure 4-7, A). (C) Typical TEM images of individual intact β-CBs (top) and β-CB aggregations (bottom). More TEM images are shown in (Figure 4-7). (D) High-resolution TEM imaging allows the direct visualization of both β-CB shell and internal structures. (E) Zoomed-in view of a single β-CB with the resolved structural features highlighted. The outer shell is highlighted by a blue hexagon, and the Rubisco molecules are highlighted by circles colour coded by individual interior Rubisco-organising pyramids. (F) Measurement of the shell thickness, inner gap (line 1) and Rubisco packing (line 2). (G) Profile analysis along the line 1 in (F), indicating the shell thickness of 4.5 nm and the inner gap (2.0 nm) between the shell and Rubisco-organising structure. (H) Profile analysis along the line 2 in (F), indicating the periodicity of Rubisco arrangement (~9.5 nm). (I) A structural model of the β-CB from Syn7942, based on the TEM observations. Top, the model of β-CB facet edge organisation (CcmK in orange, CcmL in green at the vertices), according to the shell hexamer length (7.5 nm) and the edge length (72.2 nm) measured from TEM images. Middle, the structural model of one β-CB module including one shell facet (orange), inner shell layer (grey) and a triangular pyramid Rubisco-organising core (blue). Bottom, twenty of such β-CB modules assemble to form the entire icosahedral β-CB in Syn7942.

Table 4-4 The diameter of isolated CB as determined by EM imaging

Type	Species	Diameter (nm)	Range (nm)	Reference
α-CB	*H. neap*	117.3 ± 6.9	97 – 132	(Holthuijzen et al., 1986)
α-CB	*H. neap*	100	88 – 108	(Schmid et al., 2006)
α-CB	*H. neap*	134 ± 8	116 – 169	(Iancu et al., 2010)
α-CB	*Synechococcus WH8102*	123 ± 5	114 – 137	(Iancu et al., 2007)
α-CB	*Prochlorococcus marinus MED4*	90	70 – 100	(Roberts et al., 2012)
β-CB	Syn7942	149.90 ± 13.78	100 – 200	This study

Besides the overall shape of β-CBs, EM images also provide detailed information about the shell architecture and internal organisation of the β-CB, which advances the model of β-CB structure and assembly (Rae et al., 2013) (Figure 4-6, D–F). It is evident that the intact β-CB comprises an outer shell that incorporates paracrystalline arrays of Rubisco enzymes (Figure 4-6, F). The average length of shell facet edges (vertex to vertex) is 72.16 ± 7.51 nm (n = 240). The thickness of the outer shell is 4.51 ± 0.22 nm (n = 60) (Figure 4-6, G), in agreement with the thickness of a single CcmK2 protein (Kerfeld et al., 2005). It demonstrates that the β-CB shell is constructed with a single layer of shell proteins. Intriguingly, we also observed a 2.0 nm low-density interval between the shell layer and Rubisco arrays (2.00 ± 0.24 nm, n = 60). This "gap" may accommodate a layer of loosely-packed proteins attached to the inner surface of the shell, *e.g.* CcaA, CcmM and CcmN, which play key roles in linking the shell and Rubisco-organising internal structure (Rae et al., 2013).

In contrast to the relatively disordered and less densely packed α-CB lumen (Holthuijzen et al., 1986; Iancu et al., 2010, 2007; Schmid et al., 2006), the β-CB internal structure is highly defined with paracrystalline arrays of Rubisco, in line with EM results of the ruptured Syn7942 cells (Kaneko et al., 2006). Individual Rubisco molecules inside the β-CB, notably those located in the outer layers of Rubisco arrays and adjacent to the shell, are clearly discriminated in the highly-ordered β-CB lumen. Approximately 9.5 nm Rubisco centre-to-centre distance was resolved (Figure 4-6, H). Given the 3.5 nm edge length of a CcmK2 hexamer (Kerfeld et al., 2005), the edge of a shell facet (72.16 ± 7.51 nm) is capable of accommodating 6 pairs of hexamers and 5 single hexamers between two CcmL pentamers at the vertexes (Figure 4-6, I, top). About 7 Rubisco proteins (~10 nm each) can be located along each facet edge under the outer shell (Figure 4-6, I, middle). Such protein organisation will result in a triangular pyramid β-CB substructure, which contains one shell facet with a single hexamer thick, a layer of shell-associated proteins and a Rubisco-organising triangular pyramid. Twenty of these β-CB modules eventually construct the entire icosahedral β-CB architecture (Figure 4-6, I, bottom). A Rubisco-organising pyramid under a triangular shell facet is estimated to contain 84 Rubisco proteins; a total of 1680 Rubisco enzymes may be encapsulated in one β-CB, roughly consistent with the previous estimation (Rae et al., 2013). Due to the paracrystalline packing, the Rubisco content of the β-CB is 7-fold higher than that of the α-CB (Iancu et al., 2010, 2007), though absolute quantification is required to explore the exact abundance of protein modules in CBs.

The different interior organisations of α- and β- CBs could result in the distinction in their hierarchical assembly processes. The biogenesis of α-CB was proposed to be initialized by the formation of outer shell (Baker et al., 1998; Menon et al., 2008) or follow a simultaneous assembly pathway (Iancu et al., 2010), whereas β-CBs seem to assemble from the inside out (Cameron et al., 2013; Chen et al., 2013). Characterization of partial β-CBs in this work suggests the strong protein–protein interactions within the "outer shell–inner layer–Rubisco" structures. The shell proteins, shell-associated proteins and Rubisco enzymes could potentially co-assemble to form large CB modules, which may serve as the assembly intermediates during β-CB assembly, biogenesis or degradation (Figures 4-3, 4-4 & 4-5). Concomitantly, our

EM results of intact β-CBs (Figures 4-6 & 4-7), showing more ordered Rubisco arrays at the outer surface of Rubisco arrays and less ordered Rubisco packing in the β-CB lumen, likely implies the potential "outer shell–inner layer–Rubisco" interactions (Figure 4-6, F).

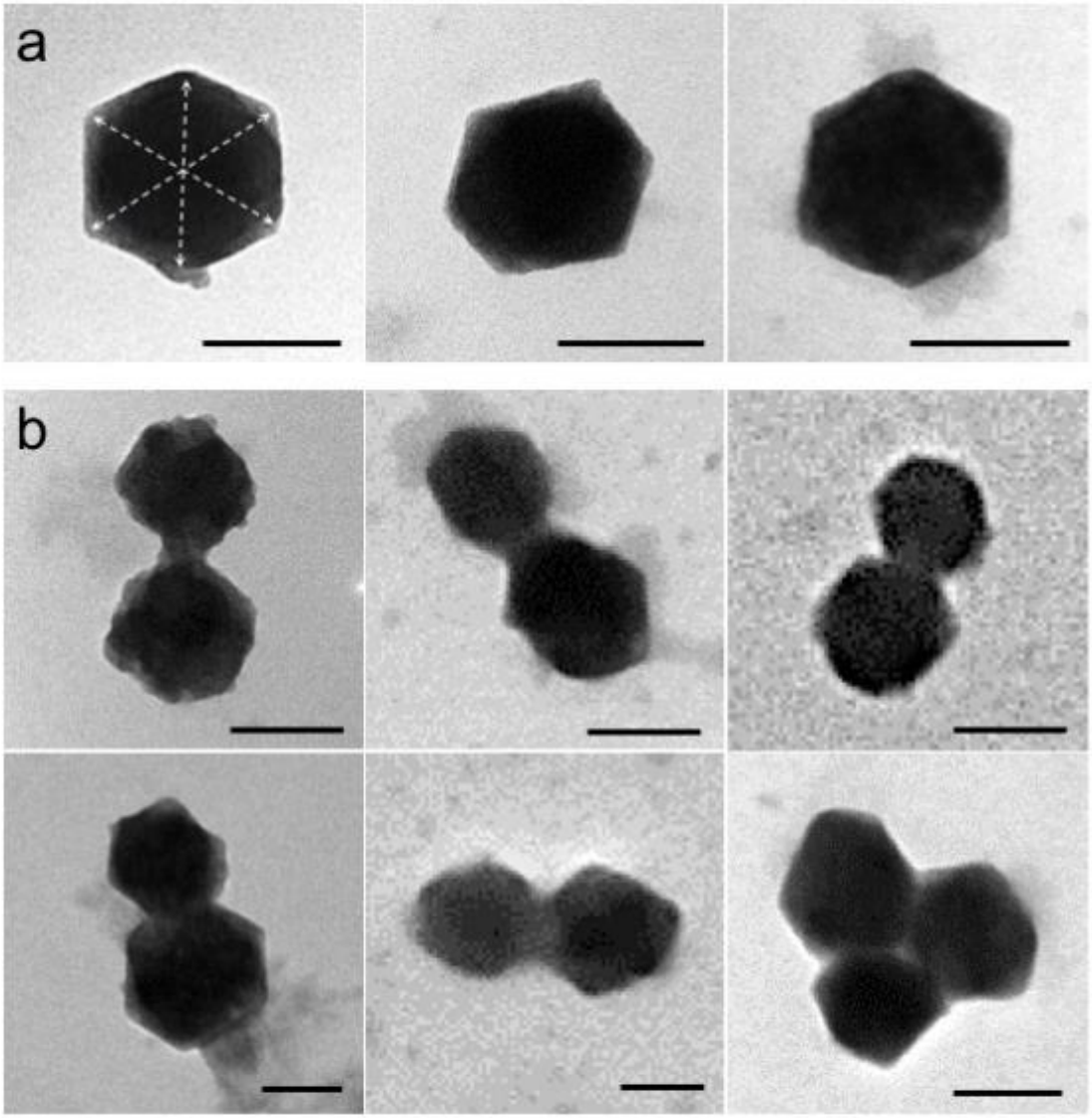


Figure 4-7 TEM images of intact β-CBs in the 40% sucrose fraction. (a) TEM images of individual intact β-CBs. The dashed arrows represent the three vertex-to-vertex measurements for determining the β-CB diameter, as described in (Figure 4-6, B). (b) TEM images of 2-3 aggregated β-CB (Figure 4-6, C). Scale bar 100 nm.

4.2.6 AFM imaging or purified CBs

High-resolution AFM imaging in solution has become a matured and powerful single-molecular tool in studying the structures of macromolecular complexes (Liu and Scheuring, 2013). By applying AFM imaging in solution, we characterized for the first time the topography and spatial protein organisation of intact β-CBs under near physiological conditions (Figure 4-8). The identification and structural integrity of β-CBs fused with eGFP were confirmed by simultaneous AFM-fluorescence imaging (Figure 4-10). AFM overview images illustrate the proper immobilization and distribution of individual β-CBs on the substrate surface (Figure 4-8). High-resolution AFM images enable the direct characterization of the topography and dimension of individual β-CBs (Figure 4-8, B & C, Table 4-4 & Table 4-5). The average height of β-CBs is 135.23 ± 23.02 nm (n = 50) (Figure 4-8, C), consistent with TEM results. Substructures in the β-CB surface were readily discerned at this resolution, which represents the molecular organisation of the β-CB shell (Figure 4-8, B-D). The facet boundaries could be occasionally observed in single CBs (Figure 4-8). Individual shell protein structures on intact β-CBs could not be distinctly discerned at this resolution, compared to partial β-CB structures that are better supported by the AFM substrate (Figure 4-5). These observations suggest the softness and flexible conformation of β-CBs. The aggregation of two β-CBs was also visualized in AFM (Figure 4-9. B), in line with our EM observation (Figure 4-6, B, and Figure 4-7, B).

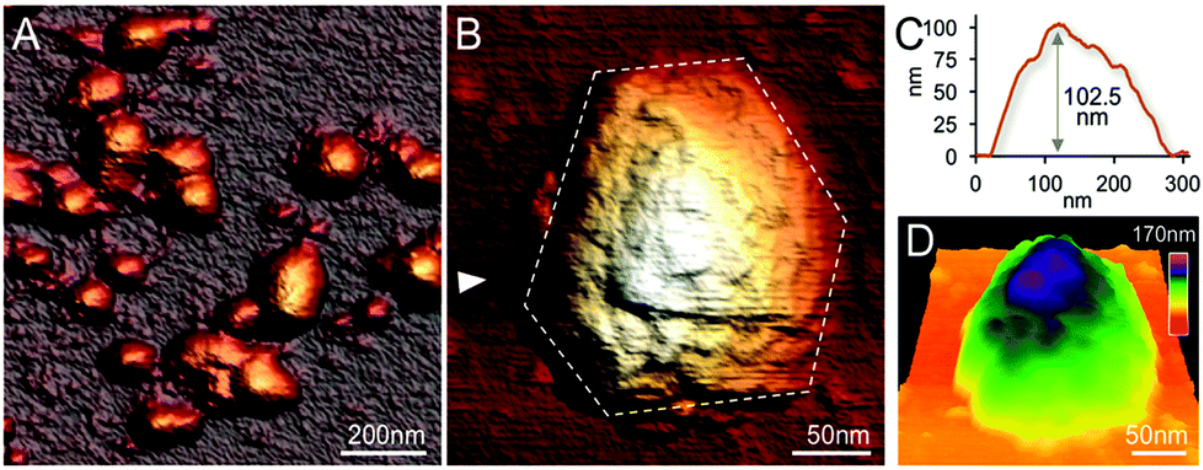


Figure 4-8 Native AFM topographs of intact β-CBs from Syn7942. (A) An overview topograph of isolated intact β-CBs in the 40% fraction captured by AFM in solution. (B) High-resolution AFM image of a single intact β-CB, showing the morphological features of β-CBs. Several surface protrusions can be distinguished. The polyhedral shape of the β-CB is outlined by white dashed lines. (C) Height profile of the β-CB, taken along the white arrow indicated in (B). (D) 3D representation of the native architecture of the same β-CB from Syn7942.

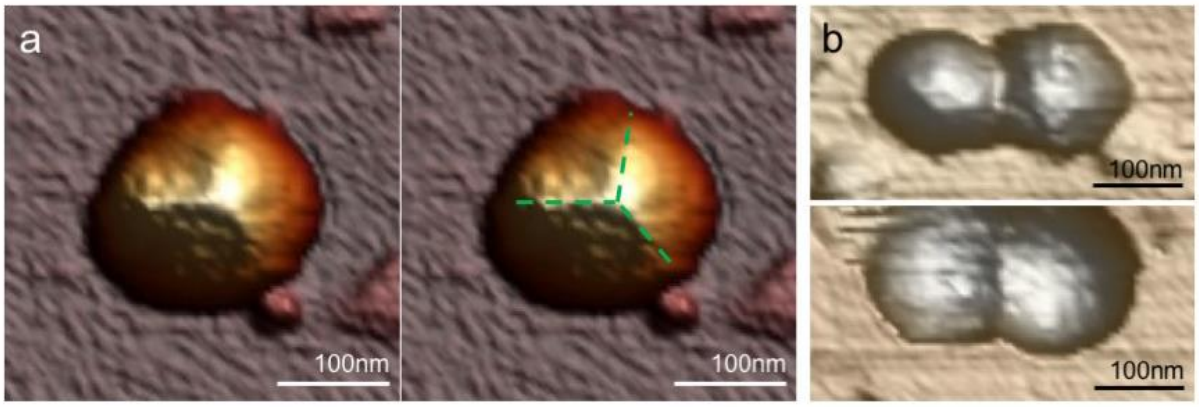


Figure 4-9 AFM images of intact β-CBs. (a) An AFM topograph of single β-CB with a vertex and three facet boundaries resolved, indicated by the green dashed lines. (b) AFM topographs of aggregated β-CBs, reminiscent of EM results (Figure 4-6, C and Figure 4-7).

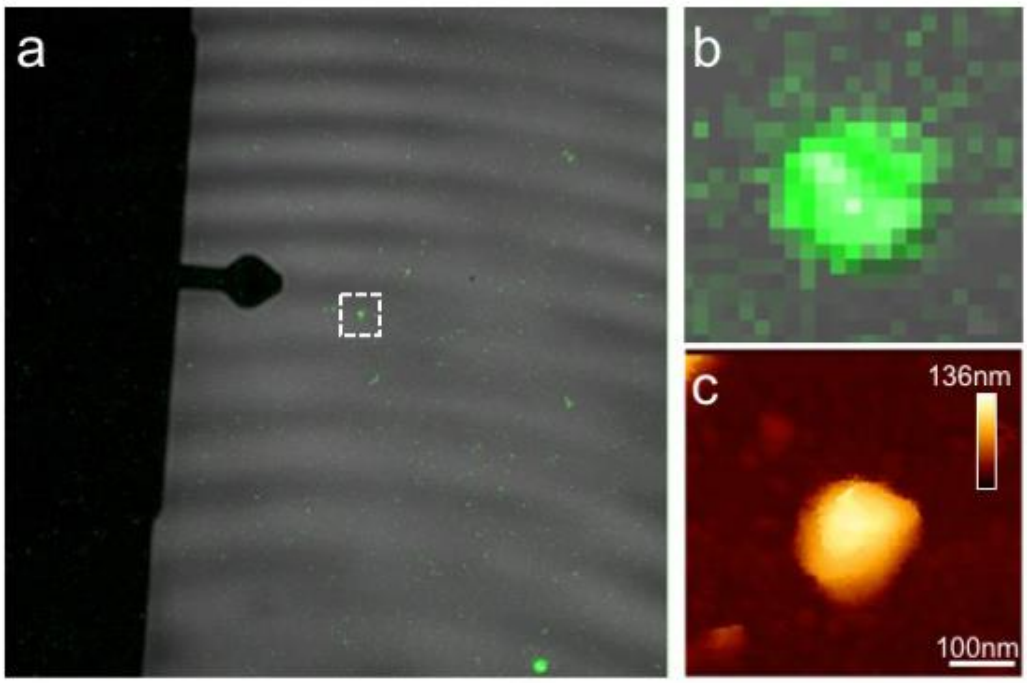


Figure 4-10 Combined confocal and AFM imaging of β-CBs fused with GFP. (a) A merged image of the transmitted and GFP channels captured using a hybrid JPK AFM-Zeiss 880 confocal microscope with a 63x objective. The white dashed square represents a 10 × 10 µm field of view of AFM after the engage. (b) Interpolated confocal image of a single β-CB in the centre of the view highlighted by the white square in panel A. (c) AFM topograph of the same β-CB captured simultaneously with the fluorescence image (b). The combination of AFM-confocal fluorescence imaging ensures the identification of β-CBs on the AFM substrate. The direct optical overlay calibration had a standard error of ±0.763 pixels, which equates to ±305.2 nm, allowing for drift in the AFM image and reengage error this overlay confidence distance was estimated at ±600 nm therefore only CB at least 700 nm from the next closest neighbour were chosen for imaging by this combined technique.

4.2.7 Mechanical properties of purified CBs

In the crowded and dynamic cellular environment, the physical properties of BMCs are essential for their stability, functionality and regulatory responses (Mitragotri and Lahann, 2009). Using AFM-based NI that has been exploited in studying viral capsid mechanics (Marchetti et al., 2016), the spring constant and Young's modulus of β-CBs was determined, to unveil the mechanical properties of CBs at near physiological conditions (Figure 4-11). A relatively low force (100 pN) was applied for AFM imaging and 1 nN force was applied for AFM NI on targeted β-CBs. (Figure 4-12, A) shows schematically a typical NI event performed on an intact β-CB. After locating the β-CB by AFM imaging, the AFM tip was positioned over the centre of the β-CB (Figure 11, B, inset) and pushed towards the CB (stage 1). There is zero force with *z*-displacement until the tip and CB contact (stage 2). As the tip pushes down, there is an increase in the force, resulting in the deformation of the β-CB structure (stage 3). (Figure 4-11, B) exhibits a collection of force-indentation curves of β-CBs. Within the range of 0–300 pN, the indentation on β-CB is up to 20 nm, which represents about 10% of the particle height, according to the previous study (Ortega-Esteban et al., 2015). No typical rupture/breaking events, as seen in viruses, were observed in β-CBs above 300 pN (Figure 5-5). A typical force-displacement curve, as depicted in (Figure 4-11, C) illustrates an initial nonlinear response, followed by a relatively linear deformation of the β-CB. The slope of the linear-like regime of the force-indentation curve is the spring constant *k* of β-CBs (∼20 pN nm^{-1}, *n* = 25, (Equation-1, Section **2.6.8**.) (Figure 4-11, C, and Figure 4-12), which represents the stiffness of β-CBs (Figure 4-12 C&D). The spring constant of β-CBs is lower than those typical for viruses (*k* 40–1250 pN nm^{-1}) (Table 4-1), revealing that the β-CB is softer than the viruses with protein-based shells.

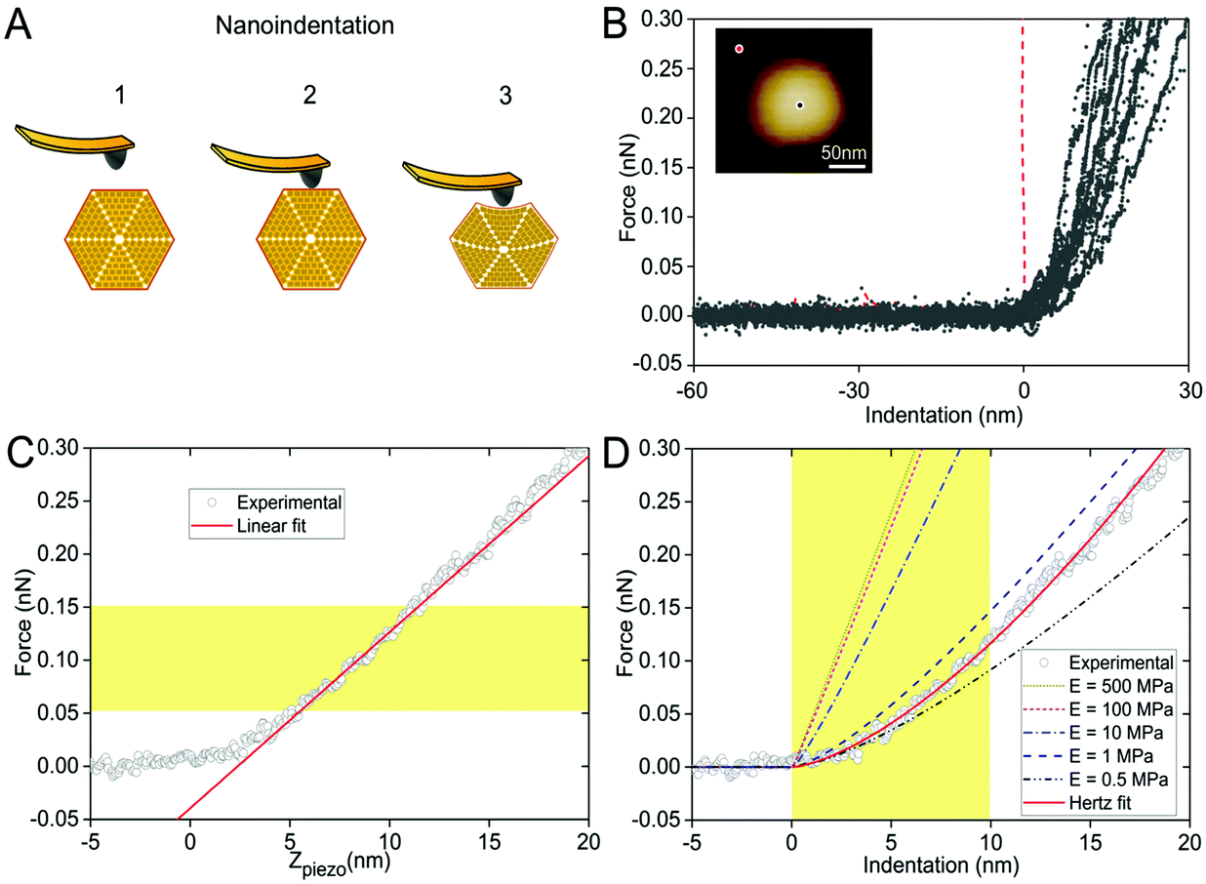


Figure 4-11 Mechanical characterization of intact β-CBs using AFM NI. (A) Schematic of an AFM NI experiment, including AFM tip engagement (1), tip-CB contact (2) and indentation (3) with increasing force. (B) Force-indentation curves of individual β-CBs. The red curve is the reference curve on the mica substrate. Inset, AFM image of a single β-CB during AFM NI. The black dot represents the indentation position on the CB, whereas the red dot represents the indentation position on the mica surface. (C) A typical force-displacement curve of a single β-CB. The red line is the fitting using the linear model based on the 0.05–0.15 nN region of the force curve. (D) A typical experimental (circle) force-indentation curve of a single β-CB and simulated force-indentation curves (colored dash lines) using a Hertz contact model in a sample with Young's modulus ranging from 0.5 to 500 MPa. The red curve is the fitting using the Hertzian model based on the 0–10 nm region of the force curve. Young's modulus of β-CBs (E_H) is 0.59 ± 0.34 MPa (n = 25), sitting between the predicted Young's moduli of 0.5 and 1 MPa. Additional details of the mechanical property calculations can be found in (Section **2.6.9**).

The thin-shell model has been widely used to determine Young's moduli of viruses, which have a linear elastic response to the indentation (Carrasco et al., 2006; Roos et al., 2010). In contrast, the force–distance curves of β-CBs present evidently the nonlinear nature (Figure 4-11, B-D) indicating the flexibility of the CBs' structures. Thus, the overall force-indentation curve is fitted to the Hertzian model (Johnson and Johnson, 1987) in the 0–10 nm region of the force curves (Fogure 4-11, D) and (Equation-3), to obtain Young's modulus of β-CBs (E_H = 0.59 ± 0.34 MPa, n = 25, Table 4-5 and Figure 4-13). (Figure 4-11, D) also shows the force-indentation curves obtained from experimental data and simulations with Young's modulus ranging from 0.5 to 500 MPa. The experimental curve (0.59 ± 0.34 MPa) sits exactly between the simulated curves with Young's moduli of 0.5 and 1 MPa. It exhibits notably lower Young's modulus than the bacterial nanocompartment encapsulin (E_H for encapsulin 30–60 MPa) (Snijder et al., 2016).

As β-CBs structurally resemble the virus capsids, we performed the comparison of the physical mechanics of CBs and viruses. We calculated Young's modulus E_S of β-CBs (77.90 ± 23.89 MPa, n = 25, Figure 4-13, and Table 4-5) using the thin-shell model (Equation-2) in the 50–150 pN region of the force curves, based on the spring constant k, shell thickness (h = ∼4.5 nm) and the size of β-CBs (R = ∼75 nm) which were determined from our AFM and TEM imaging (Table 4-5). The estimated E_S is much lower than those of viruses (140 MPa–1.8 GPa) and encapsulin (1.2–2.0 GPa) (Roos et al., 2010; Snijder et al., 2016). It is worthy to note that the exact thickness of the β-CB "shell" remains to be determined to obtain accurate Young's modulus of β-CBs, given the presence of the shell inner layer that is composed of shell-associated proteins observed in the TEM images (Figure 4-6, F). It may also be the case that the β-CB behaves as bulk material and does not have a "shell" as described by the thin-shell model, the mechanical properties of the β-CB could be similar throughout the entire complex.

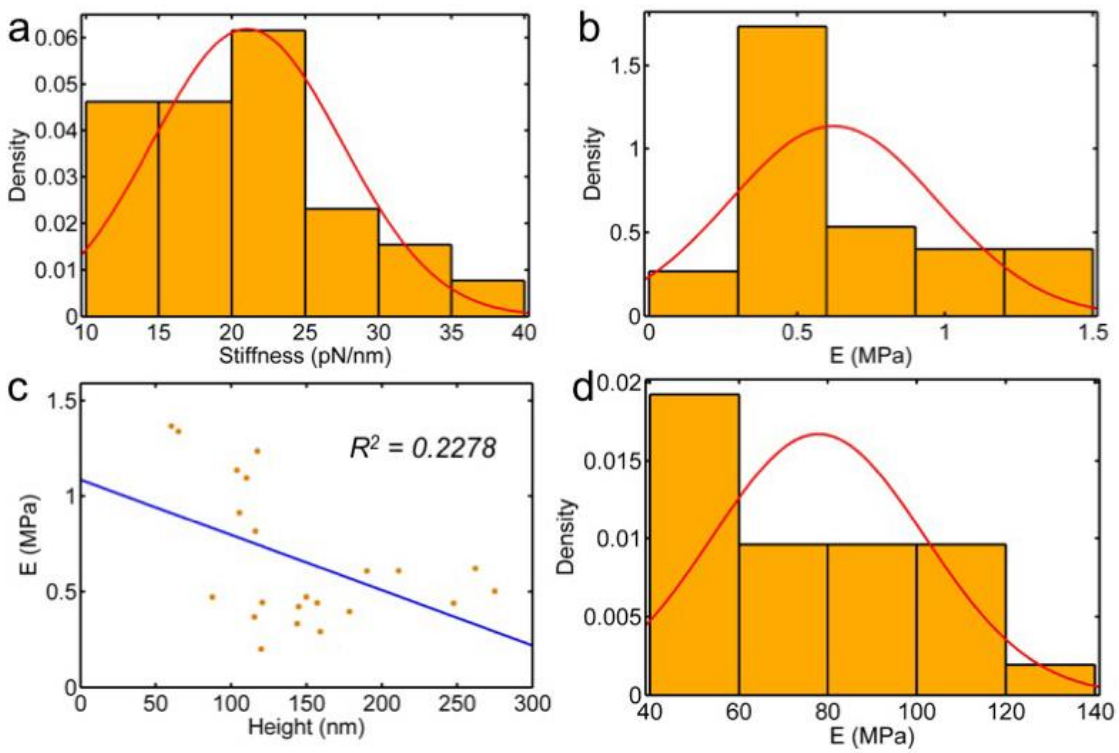


Figure 4-12 Statistical analysis of the mechanical properties of β-CBs. (a) Histogram of the β-CB stiffness (k_{CB}, n = 25). (b) Histogram of Young's moduli of β-CBs (E_H, n = 20) using the Hertzian model. (c) There is no correlation between Young's moduli and the CB diameter (y = -0.0039x + 1.1, R^2 = 0.2278). (d) Histogram of Young's moduli of β-CBs (E_S, n = 25) using the thin shell model.

Table 4-5 Physical properties of the CcmK4 eGFP β-CBs determined using EM, AFM and AFM NI in this study.

Parameter	Value	n	Detection method
Diameter	149.90 ±13.78 nm	90	EM
Facet length	72.16 ±7.51 nm	40	EM
Shell thickness	4.51 ±0.22 nm	60	EM
Rubisco packing periodicity	9.50 ±0.70 nm	30	EM
Shell-core interval width	2.00 ±0.24 nm	60	EM
Height	135.23 ±23.02 nm	50	AFM
Spring constant, k_{cb}	20 ±9 pN nm^{-1}	25	AFM NI
Young's Modulus, E_H	0.59 ±0.34 MPa	25	AFM NI Hertz model
Young's Modulus, E_S	77.90 ±23.89 MPa	25	AFM NI linear model

The physical properties of the CcmK4 eGFP β-CBs as determined using EM, AFM and AFM NI in this study. The "value" column illustrates the means, the standard deviation errors and the units of the physical properties. The "n" column indicates the number of individual β-CBs examined. The "Detection Method" column shows what techniques and methods were exploited to obtain the data. The structural dimensions of β-CBs measured by EM (Figure 4-6) these were used to build the model illustrated in (Figure 4-6, I) and measure the mechanical properties (Figure 4-11, Figure 4-12).

4.2.8 Comparison of β-CB with P22 VLP mechanics

The nanomechanical features of β-CBs and *Salmonella typhimurium* bacteriophage P22 were further compared (Figure 4-13). The physical mechanics of P22 has been characterized previously (Bhardwaj et al., 2014; Llauró et al., 2016a, 2016b). To confirm the reliability of our mechanical measurement, we applied the same procedure of AFM imaging and NI to P22 (Figure 4-13, and Figure 4-14). The height of P22 particles is 65.1 ± 5.9 nm (n = 20) (Figure 4-14) and the spring constant of P22 is 192.38 ± 63.77 pN nm^{-1} (n = 8), comparable to previously published results (Figure 4-13) (Llauró et al., 2016a, 2016b). Young's moduli of P22 fitted to the linear model and the Hertzian model are 101.04 ± 32.29 MPa and 11.06 ± 8.77 MPa, respectively (n = 8). They are both higher than those of β-CBs, suggesting that the β-CB exhibits softer mechanics than P22 (Figure 4-13).

CBs architecturally resemble icosahedral virus capsids. However, there is no evidence for sequence or structural similarity of CB shell proteins to known viral capsid proteins (Abdul-Rahman1 et al., 2013; Kerfeld et al., 2010; Yeates et al., 2010). An open question is whether CBs have the same rigidity as viruses. These data show that the particle stiffness and intrinsic rigidity, represented by the spring constant and Young's modulus of β-CBs, are both weaker in contrast to those of the human Herpes simplex virus type 1 (HSV-1) capsid and adenovirus, which have comparable dimensions (Ortega-Esteban et al., 2015; Roos et al., 2009; Sae-Ueng et al., 2014). Interestingly, β-CBs exhibit similar stiffness with the influenza virus which contains a lipid envelope (Li et al., 2014), whereas they are much softer compared to the icosahedral encapsulin (Snijder et al., 2016). Nevertheless, our results reveal the mechanical softness and flexibility of β-CBs in contrast to rigid virus capsids (Table4-1), likely ascribed to the specific assembly of multiple protein homologs in the complex CB shell architecture. Such unique mechanical signature of β-CBs might be essential to the functional plasticity of the metabolic machinery in response to environmental changes, and facilitate the metabolite passage, turnover of building blocks, recognition and regulation by other cellular components. The soft and flexible architecture could make it difficult to easily define the edges of β-CBs and

individual shell proteins by AFM imaging in solution even with gentle scanning force (100 pN), though β-CBs display regular polyhedral shape in TEM images. Applying more gentle scanning and sample fixation might be of help to obtain higher-resolution images.

Apart from the shell composition, a striking difference between the CB and viruses is the internal organisation. The viral genome is enclosed within the viral capsid, whereas the CB contains densely arranged enzymes inside the shell. It is unclear how the packing of Rubisco enzymes and protein–protein interactions within the β-CB have impacts on the overall architecture and mechanical properties of the shell and intact CB. It was shown that "empty" β-CB shells in the absence of Rubisco enzymes are only 20–30 nm in diameter (Cai et al., 2016). Further study is required to uncover what determines the assembly and intrinsic mechanics of β-CBs.

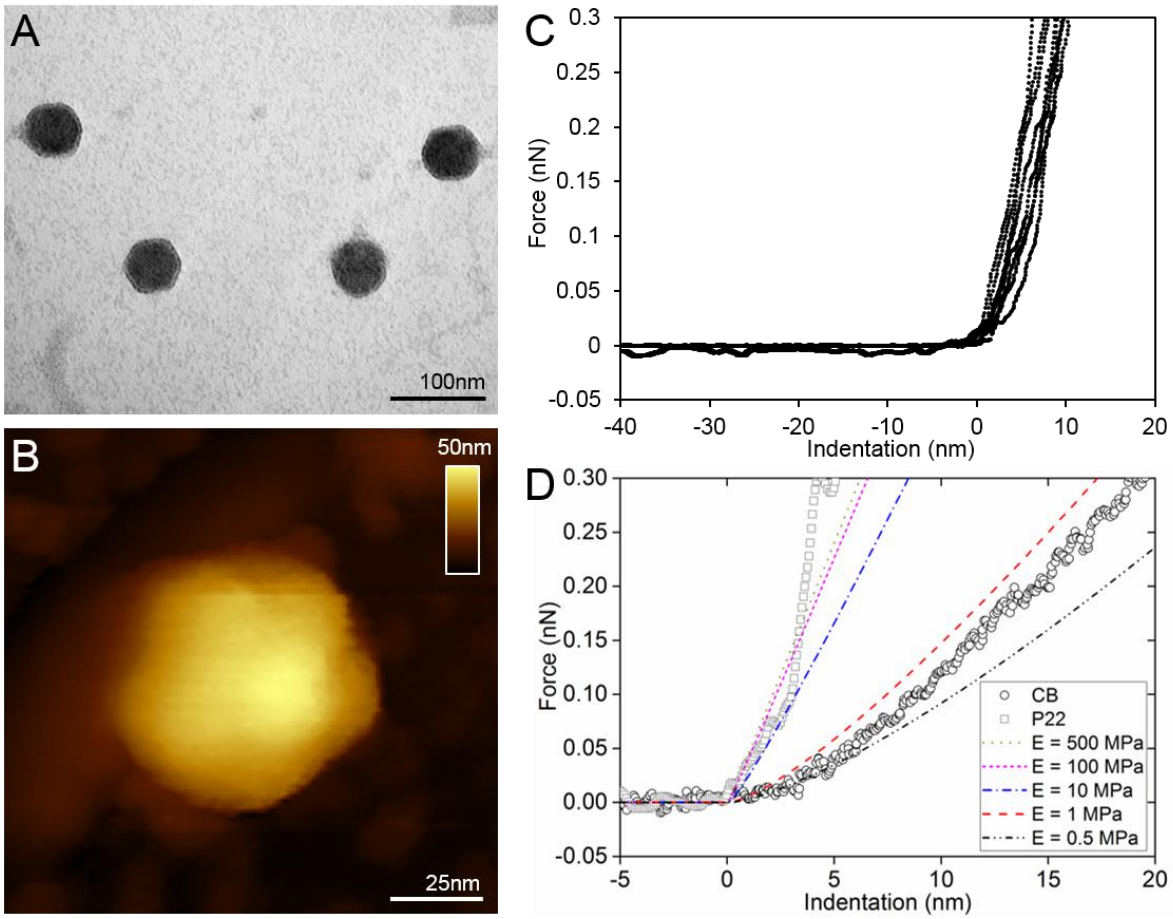


Figure 4-13 AFM NI, TEM, & AFM imaging of P22 VLP A) Representative TEM image of four isolated P22 VLP. B) Typical AFM image of an isolated P22 VLP, C) overlaid force curves captured from on top of several P22 VLP during AFM NI. D) A comparison of theoretical force curves for an infinite layer of material with 0.5 – 500 MPa young's modulus, and a typical curve for a CB and a P22 VLP.

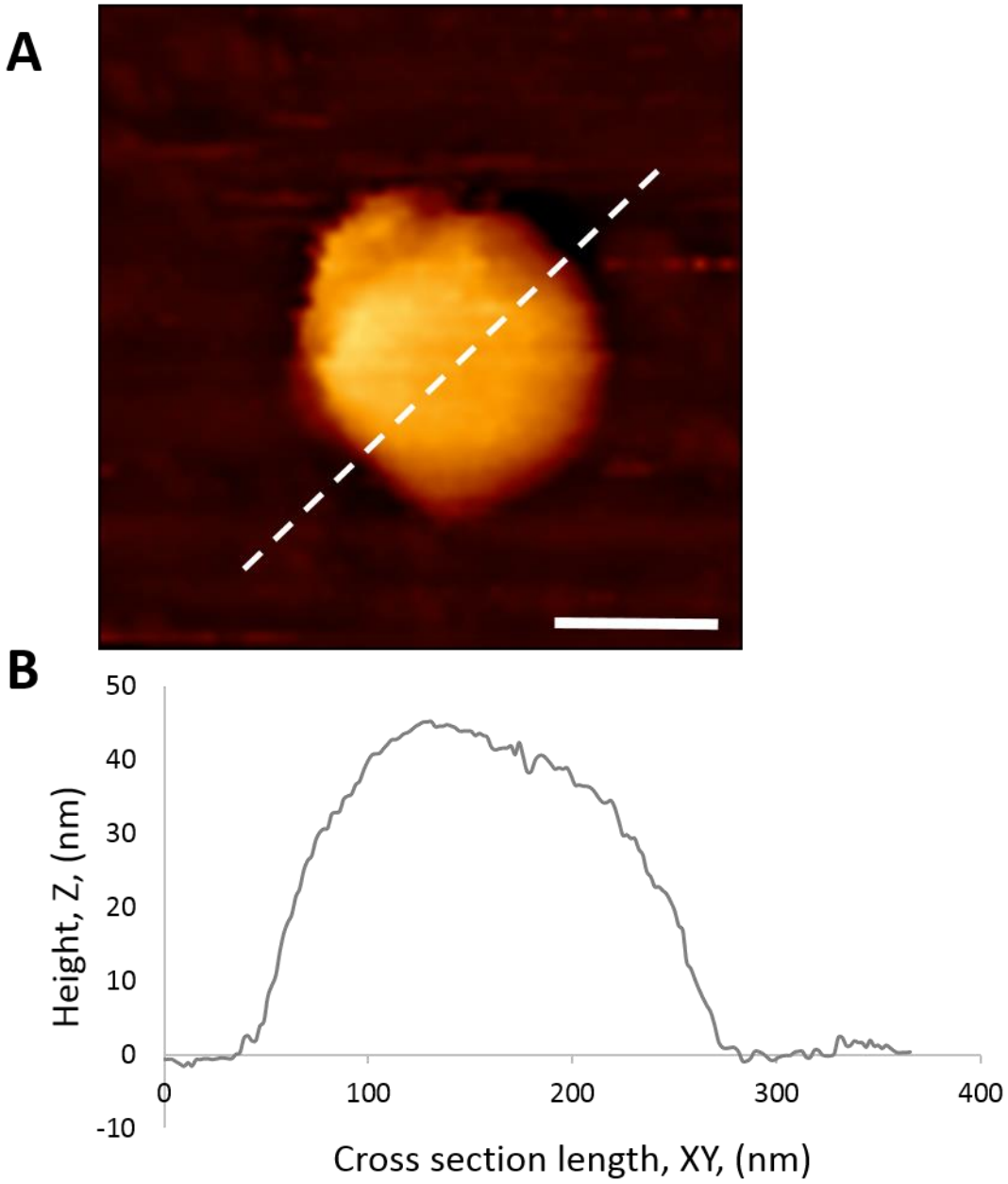


Figure 4-14 P22 VLP height determined by AFM. A) An AFM image of an isolated P22 VLP and B) the cross-section along the line indicated in A) indicating particle height of purified P22 VLP from *salmonella enterica*. The scale bar represents 100 nm.

4.3 Concluding remarks

CBs are the cornerstone metabolic modules for the high efficiency carbon fixation of cyanobacteria, and show great promise for synthetic engineering to improve the catalytic efficiency of enzymes in non-native hosts. In this work, I conducted the isolation of functional β-CBs (Figure 4-1) from the cyanobacterium Syn7942 and the direct visualization of the native organisation, topography and intrinsic mechanics of β-CBs using TEM (Figure 4-6 and 4-7), AFM (Figures 4-8 – 4-10) and proteomics (Table 4-3 and Table 4-4). I find that the intact β-CB contains three distinct structural domains, a single-layered shell that is approximately icosahedral, an inner layer of low density and densely packed paracrystalline arrays of interior Rubisco. I also characterized partial β-CB structures (Figure 4-3 and Figure 4-4) that consist of shell facets, shell-associated proteins including functional Rubisco enzymes (Figure 4-2), probably serving as the assembly intermediates of β-CBs. In addition to this, the native protein organisation of shell facets, as well as the topography and the intrinsic mechanics of native β-CBs were characterized by AFM for the first time. These results illustrate the soft nanomechanical properties of β-CBs (Figure 4-11 and 4-12) compared to icosahedral viruses (Table 4-1 and Figure 4-13), likely revealing the unique spatial organisation of CB proteins (Figure 4-6, I). The study provides new insights into the assembly, organisation and physical nature of functional β-CBs, which can be extended to α-CBs and other BMCs. Comprehensive understanding of the CB structure and mechanics will underpin the design and engineering of functional synthetic CBs, to enhance photosynthesis and develop new bio-nanoreactors and protein scaffolds using BMC proteins as nanoscale materials. It offers a powerful approach for assaying the functional organisations and material mechanics of natural and engineered biological systems. This approach could be used in the future to verify the correct assembly and structure of CBs expressed in plant chloroplasts, or extended to any BMC expressed heterologously elsewhere.

This page is intentionally blank and marks the end of chapter four.

*"**The Scientists of today think deeply instead of clearly.** One must be sane to think clearly, one can think deeply and be quite insane." **Nicola Tesla,** 1934.*

Chapter 5 The Comparison of Carboxysome Physical Properties, Structure and Function

5.1 Introduction

CBs can be classified into families based on the type of Rubisco they contain (Section **1.2.2** and Table 1-1). Long standing questions about CBs involve determining the differences between the α-CBs and the β-CBs. These differences may make one type of CB more suitable than the other for certain applications (Rae et al., 2013), which remains to be seen (Frank et al., 2013). Previous work has suggested these CBs may assemble differently and major differences such as their size and localisation within different bacteria (Chowdhury et al., 2015, 2014; Kerfeld et al., 2005; Rae et al., 2013; Sun et al., 2016). However, they are yet to be directly compared using a quantifiable structural biomarker. Therefore, it is not yet possible to know how these observed differences translate into the mechanical or structural properties of CBs.

In this study, I present the direct comparison of the structural and mechanical properties of α-CBs and β-CBs. Wild type (WT) CBs were purified from two model organisms, α-CBs from *H. neap* and β-CBs from Syn7942 and were studied using transmission TEM and in buffered near native conditions using AFM. The mechanical properties of each WT CB were studied using AFM NI as described for CcmK4 eGFP CBs in Chapter 4 (Faulkner et al., 2017). These properties allow us to determine the nature of the internal organisation of these CBs and directly compare them under similar conditions. This has given insight into how these two CBs differ.

5.2 Results and discussion

5.2.1 Purification of the CBs, confirming the purity and functionality

Both the *H. neap* and Syn7942 CBs were isolated using a similar procedure, ending in a sucrose gradient, as previously described in (Section **2.4**, Figure 4-1) (Faulkner et al., 2017). Successful separation of the CBs from other cellular components was confirmed using SDS-PAGE to determine CB protein composition, observation of CB by TEM in the absence of major contaminants, retained structural integrity, and retained Rubisco activity in the purified samples (Figure 5-1). The sucrose gradient steps and centrifuge speed differ for the α-CB and β-CB and were optimised independently, as they differ greatly (Rae et al., 2013) in both size and density (Figure 5-6, Table 4-4). In both cases the intact CBs were mostly observed in the 40% sucrose step (Figure 5-1, A) by TEM (Figure 5, but SDS-PAGE showed both α and β-carboxysomal proteins were located throughout the gradient. This is likely due to the inherent heterogeneity in size and presence of various CB fragments and/or assembly intermediates (Figure 5-1). WT Syn7942 β-CB proteins and Rubisco activity were detected in all fractions, including those fractions that contained partial β-CBs (Figure 5-1.C). Interestingly despite the similar carboxysomal protein and fragments of *H. neap* α-CBs observed in the other sucrose gradient fractions, only the 40% fraction contained significant Rubisco activity (Figure 5-1.F). Suggesting that, at least in our assay conditions, intact α-CBs are required for the *H. neap* 1A Rubisco to function effectively. Partial β-CB, or even free Syn7942 1B Rubisco octamers, can still function effectively but at a lower rate than when encapsulated in an intact β-CB. We cannot ignore that these α and β intact CB and their sub-complexes fractionate differently, and the composition of the fractions is not identical. Nor can we ignore that some α-CB sub-complexes may also be in the 40 % fraction. Either of these points could also explain, or be partially responsible for, the lack of ^{14}C fixation activity in the other sucrose fractions.

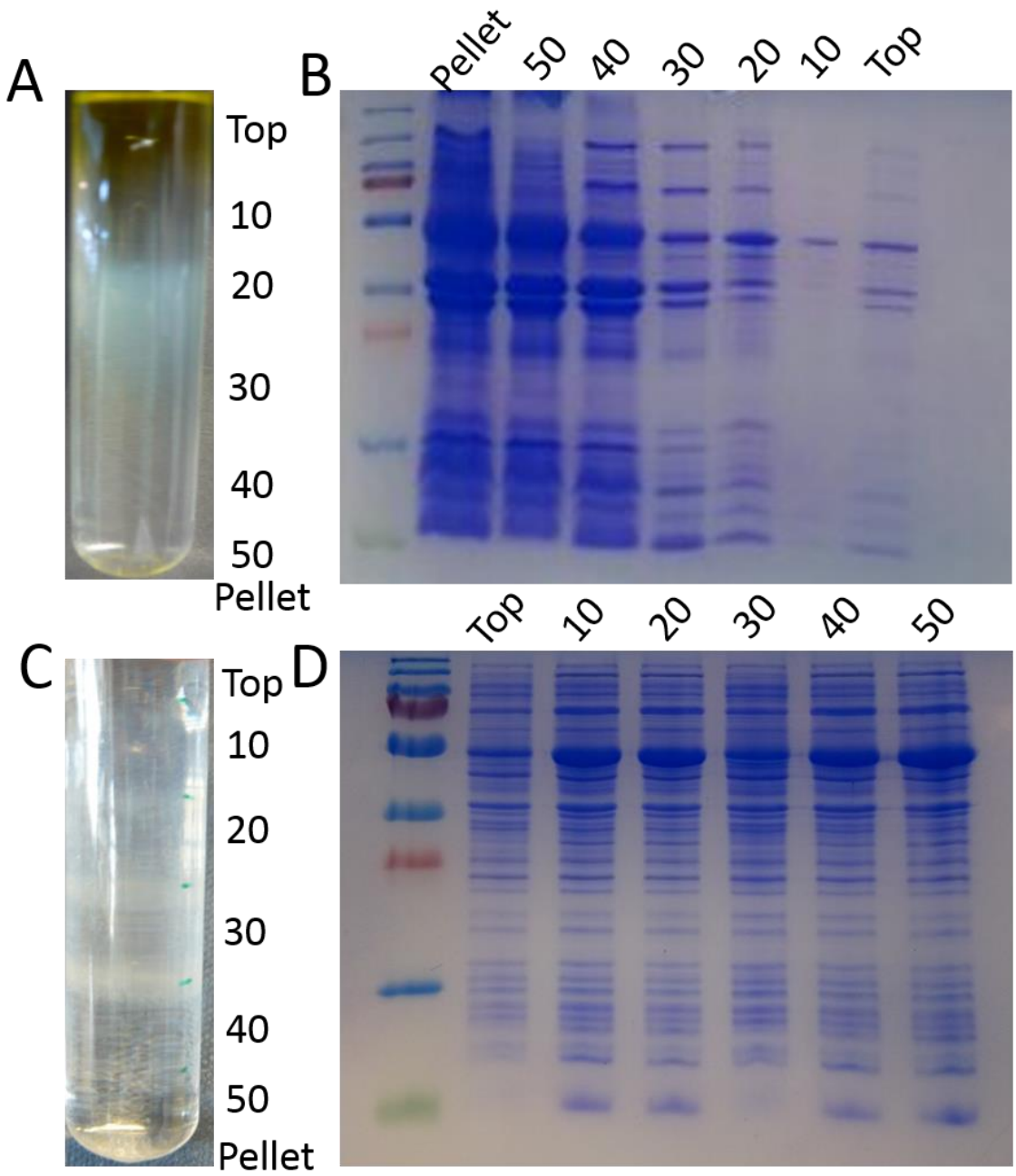


Figure 5-1 α- and β- CB fractionation by sucrose gradient. A) Photograph of the sucrose gradient after separating β-CBs from Syn7942. B) SDS-PAGE showing the protein content of the β-CB sucrose fractions, reminiscent of CcmK4 eGFP β-CB (Figure 4-1). C) Photograph of the sucrose gradient after separating α-CBs from *H. neap*. D) SDS-PAGE showing the protein content of the α-CB sucrose fractions.

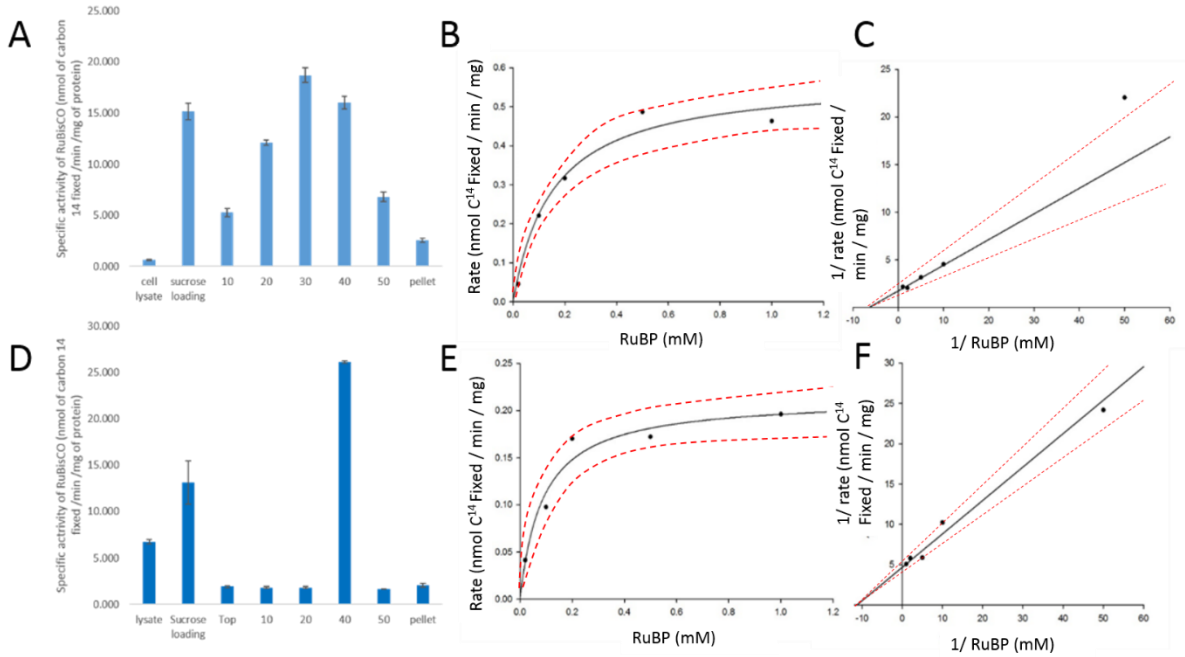


Figure 5-2 Comparing α- and β- CB functionality. A) The specific activity of Rubisco in β-CB sucrose fractions, as determined by ^{14}C radiometric assay. B) The MM curve of the Rubisco specific activity of β-CBs from the 40% sucrose fraction, in the presence of varied RUBP concentrations, and C) A Lineweaver-Burk plot (Lineweaver and Burk, 1934) of the same β-CB V_{max} = 0.57 ±0.31 nmol/min/mg K_m = 0.30 ±0.09 nmol/min/mg. D) The specific activity of Rubisco in α-CB sucrose fractions, as determined by ^{14}C radiometric assay. E) The MM curve of the Rubisco specific activity of α-CBs from the 40% sucrose fraction, in the presence of varied RUBP concentrations, and F) A Lineweaver-Burk plot of the same. α-CB V_{max} = 0.27 ±0.15 nmol/min/mg and K_m = 0.18 ±0.06 nmol/min/mg.

5.2.2 Comparing the mechanical properties of α- and β- CBs

AFM nano-indentation (NI) probed the mechanical properties of both the *H. neap* α and Syn7942 β-CBs. This technique has been widely used in the study of viruses and on CcmK4-eGFP β-CBs. Data shown are from an average of 5 individual force curves each measurement was taken on top of a CB as described in (section **2.6.6** & Faulkner et al., 2017). Using the same type of AFM cantilever, with the spring constant calibrated for each NI experiment, buffers and experimental conditions to probe both

CBs. CBs were located using a single overview scan using minimal imaging forces 50-100 pN, then 'framed' in the centre of a 400x400 nm scan again imaged only once with minimal force applied. During the indentations 300 pN was the maximum force applied, so that reproducible and accurate mechanical properties were recorded during only elastic deformation of the undamaged CB (Mateu, 2012).

Current theories in published BMC studies assume that BMC shells are structurally similar to viral capsids, and BMCs are often referred to as virus like (Howorka, 2011; Krupovic and Koonin, 2017; Roos et al., 2009; Yeates et al., 2011, 2010). Here the comparison of α- & β- CBs shows that *H. neap* α-CBs are significantly stiffer, p=0.00391 by T-test, in terms of both spring constant (K) 103.1 ±67.1 nN/µm and Young's modulus (E) 9.32 ±6.39 MPa than Syn7942 β-CBs K=12.69 ±7.68 and E=0.44 ±0.35 MPa (Table 1). These CBs were also compared to P22 VLP K=198.13 ±27.85 and E= 11.96 ±7.47 (Section **4.2.7** and Figure 4-14), and both types are significantly different to the P22, p=0.00273 by two-tailed T-test. Therefore, there are significant differences in the mechanical properties of α- and β- CBs. There are also significant differences between CB and P22 VLP. This suggests that the different types of BMC are not structurally similar amongst themselves, and the BMC shell cannot be assumed to be like a viral capsid.

Table 5-1 The properties of intact α- and β- CBs determined by AFM.

	Height (nm)	E_{hertz} (Mpa)	K (nN/µm)	Linear model (Mpa)	n
WT *H. neap*	90.70 ± 9.17	9.32 ± 6.39	103.10 ±67.10	253.25 ± 91.17	25
WT Syn7942	179.92 ± 76.50	0.44 ± 0.35	12.69±7.68	56.36 ±34.12	25

The mechanical properties calculated from CB during elastic deformation, from coincident force curves (Figure 5-3).

After several sets of successive nano-indentation measurements on the same CB the mechanical property values returned began to decrease (Table 5-2) and the nature of the force curves began to change (Figure 5-3). Elastic deformation where the extend and retract curves of each measurement coincide gave way to plastic deformation, where the retract curve no longer coincides with the extend curve. This phenomenon was observed after 2-3 sets of 5 measurements interspaced by images for the Syn7942 CB when applying 300 pN maximum force. This was not observed when applying 300 pN maximum force to the *H. neap* CBs but was observed in a similar fashion when applying a maximum of 500 pN for the *H. neap* CBs. Once the first non-coincident curves were observed a set of new mechanical properties was determined using the new plastic deformation curves on the 'damaged' CB (Figure 5-3).

If even higher maximum forces up to 1 nN were applied to either CB, no change from plastic deformation was observed for the Syn7942 CBs but additional unfolding events were observed for the *H. neap* CB (Figure 5-3, E).

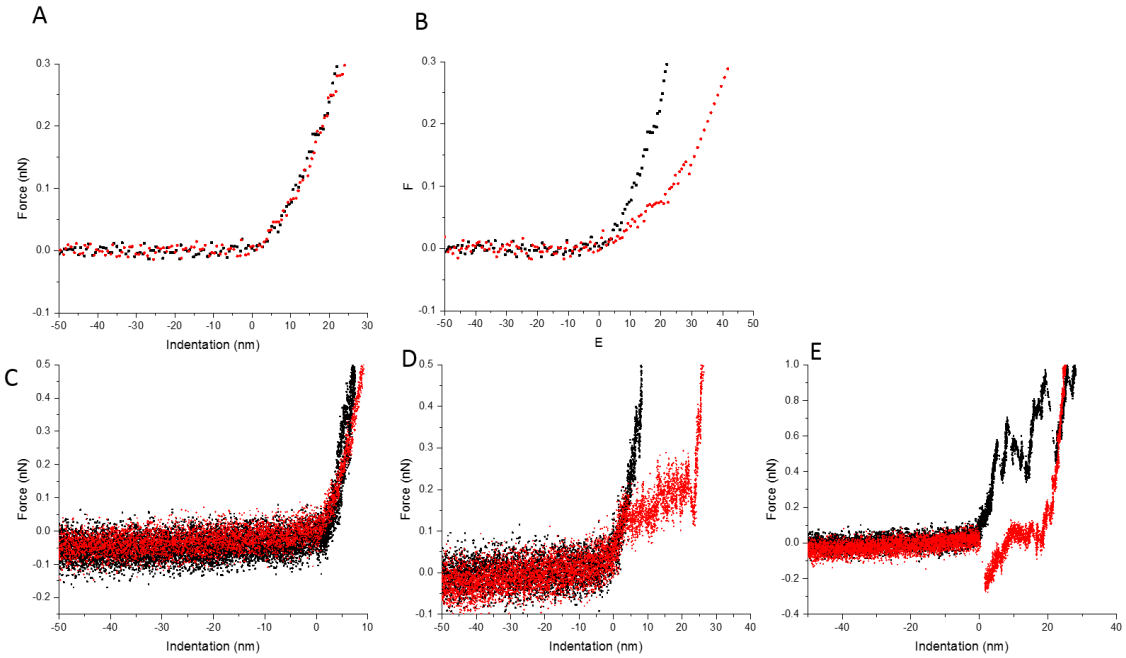


Figure 5-3 Determining which are intact, and damaged, CB by NI. A) A pair of extend (black) and retract (red) curves from an indentation on top of a WT Syn7942 β-CB. B) A subsequent pair of force curves on top of the same WT Syn7942 β-CB, where after successive indentations plastic deformation has now been caused and the CB is damaged so the extend and retract curves no longer coincide. C A pair of extend and retract curves from an indentation on top of a WT *H. neap* α-CB. D) A subsequent pair of extend and retract curves on top of the same WT *H.neap* α-CB. E) A pair of extend and retract curves for an indentation where 1nN was applied and unfolding events have been induced on the same WT *H. neap* α-CB from C & D.

The properties measured during elastic deformation can inform us on the physiologically relevant, natural properties of an intact CB (Table 5-1). Those same properties measured during plastic deformation, and the number of indentations and force of indentations required to 'damage' the CB inform us of the physical stress resistance and brittleness of the CB (Table 5-2). Being able to accurately measure both of these also allows us to use these properties as indicators of a CB's correct assembly during heterologous expression, or to see if a certain process damages a CB for example. Therefore,

the properties during both elastic and plastic deformation were investigated, for the same set of CBs. Interestingly the properties of damaged CB during plastic deformation are reproducible and comparable from CB to CB, suggesting that the damage caused is similar and is likely a simple "buckling" of the shell rather than the "shattering" type breaking events seen for viruses (Marchetti et al., 2016; Mateu, 2012; Roos et al., 2012, 2010). Buckled α-CB are also stiffer than buckled β-CB and retain higher values for both Young's modulus (Table 5-2). This was confirmed in the AFM images before and after indentations (Figure 5-5. B & D). Despite a similar buckling event taking place, the force required to induce it was much higher for α-CB (Figure 5-4 and Figure 5-5). No damage, only elastic deformation, was observed when applying 300 pN to α-CB, but the same force induced plastic deformation and damaged β-CB (Figure 5-4). At least 500 pN was required to induce comparable plastic deformation and buckling in α-CB. Thus suggesting α-CB are more robust than β-CB as *H. neap* α-CBs have a much more rigid and durable shell than Syn7942 β-CBs. This may be due to the differences in the luminal organisation (Figure 5-6 and Rae et al., 2013). The *H. neap* α-CBs require a stiffer, more robust and more capsid like shell, because the luminal contents are less densely packed, and less organised (Figure 5-6 and Rae et al., 2013), thus offer less support to the shell. Conversely, the more highly organised, para-crystalline, densely packed luminal contents of the Syn7942 β-CB may structurally support the shell and entire β-CB may act as a bulk material.

Table 5-2 The mechanical properties of damaged α- and β- CBs.

	E_{hertz} (MPa)		K (nN/μm)		n
WT *H. neap*	6.49	±1.17	30.17	±32.12	12
WT Syn7942	0.11	±0.05	3.93	±1.53	19

The mechanical properties calculated from CB during plastic deformation, from non-coincident force curves (Figure 5-3).

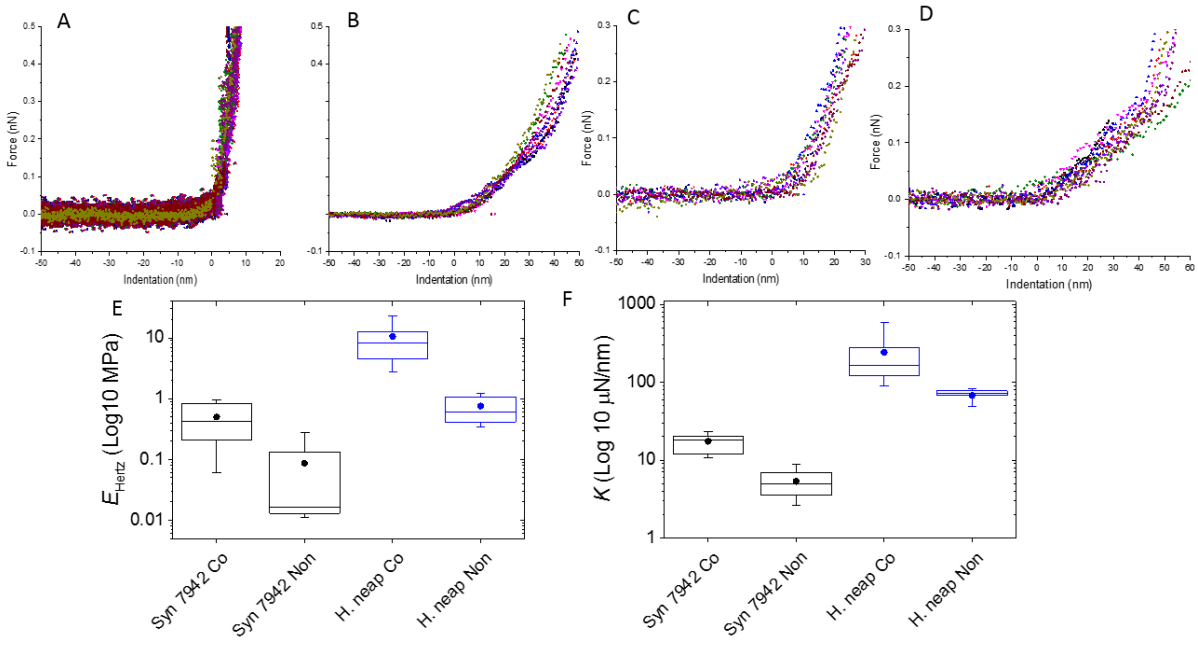


Figure 5-4 Comparing the mechanical properties of intact and damaged CBs. A) 10 overlaid typical extend curves from curves that coincide, recorded on top of 10 individual α-CB. B) 10 overlaid extend curves on top of the same CB as in A after the curves no longer coincide (Figure 5-3). C & D) The same overlaid sets of 10 force curves as shown in A & B but this time recorded on top of β-CB. E) A box plot showing Young's Modulus determined by the Hertz model from both coincident curves on 'intact' CBs and non-coincident curves on 'damaged' CBs. F) The stiffness determined from both coincident curves on 'intact' CBs and non-coincident curves on 'damaged' CBs. The box plots indicate the mean as a dot, Median as a line, upper and lower quartiles as the box, and 10 – 90 % of the data as the bars, n values as shown in (Table 5-1 and Figure 5-2).

The way in which the change in the mechanical properties occurs during plastic deformation is different between α- and β- CB, *H. neap* CBs lose apparent stiffness with a mean decrease of ~70%, 103.10 – 30.17 nN/µm in one major step after ~8-10 indentations (Figure 5-5, A). Whereas, Syn7942 CBs gradually begin to lose apparent stiffness after ~5-6 indentations with a similar ~70% proportional decrease overall after ~15 successive indentations, 12.69 – 3.93 nN/µm (Figure 5-5, A). The overall

change in Young's modulus is similar for both *H. neap* ~70%, 9.32 – 1.17 MPa, and Syn7942 ~70%, 0.44 – 0.11 MPa, interestingly, these changes are reproducible from plastically deformed 'damaged' CBs as shown by the typical force curves overlaid in (Figure 5-4, A-D), despite the difference in applied force. The change in height during plastic deformation is also different for the two types of CBs and is again reproducible, an example of how a typical CB of each type's topography responds to serial indentations is depicted in (Figure 5-5, C&E). From cross-section analysis after each batch of 5 indentations a decrease in the apparent height of the CBs was regularly observed, 15.45 ±4.31 nm for Syn7942 and 10.72 ±3.64 nm for *H. neap* CBs. A typical set of cross-section data for each CB is depicted in (Figure 5-5, C & E). These differences could again be explained by the relative internal organisation of the two types of CB.

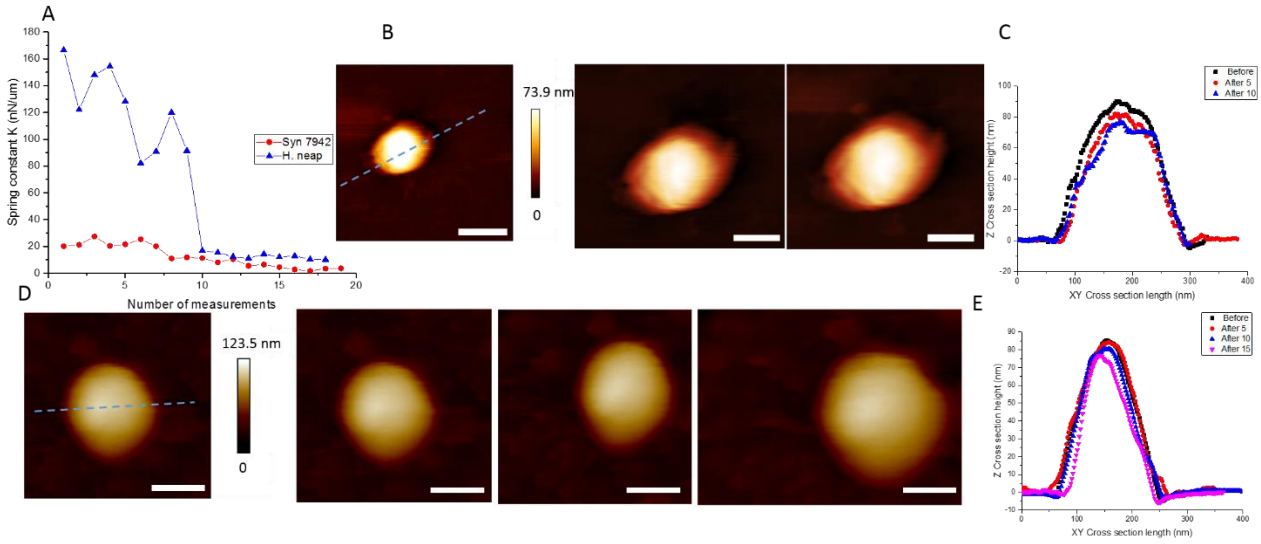


Figure 5-5 determining the resistance of CBs to mechanical fatigue. A) two series of spring constant (K) values measured by NI on top of either a WT Syn7942 CB (red) or a WT *H. neap* CB (blue). The K value is plotted against the number indentations made on top of the respective particle and decreased with successive measurements. B) Three AFM height topographs of the same Syn7942 CB taken before and after 5 then 10 NI measurements. C) Cross-sections taken from the three topography in B) along the dashed line indicated. D) Four AFM height topographs of the same *H. neap* CB taken before and after 5, 10, then 15 NI measurements. E) Cross-sections taken from the four topographs in D) along

the dashed line indicated. Both the height and spring constant appear to decrease with successive indentations as the CB are plastically deformed. Scale bars represent 100 nm.

5.2.3 Comparing the internal organisation of α- and β- CBs

The internal organisation of these CBs can also be observed using TEM, as shown in the previous study with CcmK4 eGFP β-CBs (Faulkner et al., 2017). *H. neap* α-CBs and Syn7942 β-CBs were imaged by TEM, under the same conditions, in this study; processed and rescaled images are presented in (Figure 5-6). In the majority of β-CBs the luminal para-crystalline organisation, shell and lower protein density 'gap' between the shell and luminal enzymes were all seen, reminiscent of CcmK4 eGFP (Figure 4-7). From all of the α-CBs observed by TEM, no such internal organisation or low density protein 'gap' was observed. The shell and an approximately homogeneously dense lumen are always observed. This corroborates the AFM observations and again suggests that the *H. neap* α-CBs are organised with a rigid shell, while the Syn7942 β-CBs have a more flexible shell, and have a bulk material like structure with a more interlinked shell and lumen.

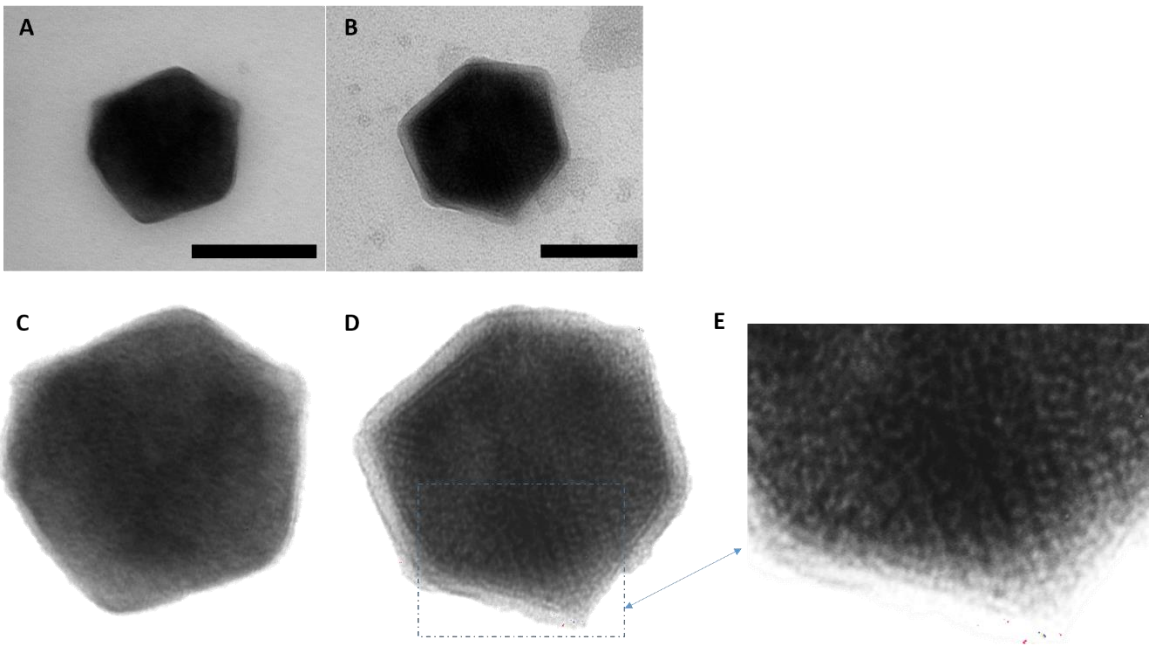


Figure 5-6 TEM imaging of α- and β- CBs. A) An original unprocessed TEM image of a single typical α-CB, B) a comparable unprocessed TEM image of a single β-CB, reminiscent of the CcmK4 eGFP-CB TEM images in (Figure 4-7). Scale bars represent 100 nm. C) A processed and contrast enhanced image from A). D) A processed and contrast enhanced image from B) where no luminal Rubisco can be seen. E) A zoomed in image from the blue hashed box in D) where the paracrystalline arrays of 1B Rubisco can be seen.

5.3 Concluding remarks

From the data presented here, one can conclude that α-CBs are physically different from β-CBs. The shell of α-CBs is stiffer than the shell of β-CBs. The luminal organisation of the α-CB and β-CB is different. The α-CB has no apparent pattern while the β-CB interior is para-crystalline, as found previously (Rae et al., 2013). Moreover, the organised luminal contents (Figure 5-6) of the β-CBs are at least partially responsible for the overall structural integrity (Figure 5-4 & Table 5-1). This organisational difference also suggests why partially formed or broken fragments of β-CBs maintain relatively high functionality (Figure 5-2 & Figure 4-1), whilst suggesting α-CBs may only have a significant function when they are completely intact (Figure 5-2), but further evidence is required to better support this. If the α-CB's contents are less strongly associated with the shell and each other than in the β-CB's, it is likely easier for them to disassemble from the CB complex and escape into solution if not completely encapsulated. Since our radiometric Rubisco assay uses $H^{14}CO_3^-$, protein aggregates with co-localised CA and Rubisco are required for the formation of radiolabelled 3-PGA which is measured to determine Rubisco activity. This observation suggests the α-CB is comprised of a shell encapsulating densely packed but loosely associated luminal contents, or at least loosely associated CA. The β-CB is composed of a shell strongly associated with the luminal contents which are also strongly associated with each other even when not fully encapsulated.

This study has shown clear evidence for structural, functional and physiologically relevant differences between two CBs from different species in terms of Rubisco activity and kinetics (Figure 5-2). These differences suggest that these BMC are fundamentally different and that they therefore likely have different preferred environments, where they function optimally. Further comparisons between other α-CB, and other β-CB, from different species are required to determine if any of the observations here are due to interspecies variation. This has implications when modifying, repurposing or heterologously expressing any BMC to create a new function or modulate activity. The BMC proteins may be modular

and they may be interchangeable with one another functionally. However, changing the protein composition or stoichiometry of a BMC may cause large scale organisational and mechanical changes, akin to those seen between the α- and β- CBs. These two BMCs appear to have developed a particular structure that must presumably be better suited to the intracellular environments of their respective host species. It, therefore, stands to reason that certain BMCs will function better in certain applications than other BMCs, either in the C3 plant chloroplast or encapsulating a different set of enzymes. Using the approach described in this study one can begin to answer these questions by conducting similar AFM investigations under varied conditions and by repeating these experiments using other purified BMCs. The mechanical properties of BMC can help us to unravel their secrets and unlock their potential.

This Page is intentionally blank and marks the end of chapter five.

*"**False facts are highly injurious to the progress of science**, for they often endure long; but false views, if supported by some evidence, do little harm, for every one takes a salutary pleasure in proving their falseness."* ***Charles Darwin**, (the Descent of Man), 1871.*

Chapter 6 Discussion and Concluding Remarks

6.1 Summary of findings

This study has elucidated new information that improves our understanding of BMCs and developed new methods which can be used in future studies of BMCs. Numerous aims were achieved, including the discovery of the dynamic nature of BMC-H shell protein self-assembly (Section **6.2**), the development of a method for the isolation of structurally intact functionally active CB from Syn7942 (Section **6.3**), and the first direct comparison of α and β-CB structure and function (Section **6.4**).

6.2 Dynamic self-assembly of bacterial microcompartment shell protein

Before this study, no dynamic movement of BMC shell proteins had been observed. From the HS-AFM data (Figure 3-7) it is clear that the BMC shell, or at least the BMC-H proteins in the shell, exist in a highly dynamic and mobile state. This has implications for future study of BMC shells when considering shell turnover and repair mechanisms. These data provide molecular insights into the formation, interactions, and dynamics of BMC shell facets. One of the open questions in this study is whether or not BMC shell protein assembly is dependent on environmental conditions. If BMC shell proteins assemble differenlty under varied pH conditions, for example, could be answered using our HS-AFM approach. The rate of dynamic events or the size of the patches could be pH dependent, and could be tested using HS-AFM in buffers of a different pH.

The design and construction of synthetic BMCs have attracted intense interest for the bioengineering of nanoreactors and molecular scaffolds (Frank et al., 2013; Kerfeld and Erbilgin, 2015). We show that preassembled shell hexamers form single layer sheets of uniform orientation (Figure 3-14), and

patches containing a mixture of orientations were not observed. This suggests that the BMC shell is comprised of proteins all in the same orientation. In addition to the arrangement of hexameric proteins, we find that individual hexamers can dissociate from and incorporate into assembled sheets (Figure 3-7), indicating an overall flexible intermolecular interaction. We also show that specific contacts at the interfaces of neighbouring proteins influence the dynamic features of shell proteins (Figure 3-8), and thereby the self-assembly of shell facets. Crucially for controlling future molecular scaffolds, we show the assembly characteristics of BMC-H proteins can be modulated by point mutations (Figure 3-16). Understanding the details of self-assembly of BMC shell proteins is a prerequisite for control and engineering of BMC-based architectures, with the aim of building designed nanoreactors and molecular scaffolds.

The HS-AFM method could also be used in future projects to observe the assembly of mixtures of BMC proteins. Chimeric BMC shell structures a can be formed by mixtures of the *Haliangium ocraceum* BMC-H, BMC-T and BMC-P proteins (Sutter et al., 2017), and other BMC shell protein mixtures have already been studied by HS-AFM (Garcia-Alles et al., 2017), it would be an interesting future study to investigate the stepwise formation of these partial BMC shells with HS-AFM.

6.3 Methods to isolate and study intact, functional β-carboxysomes from *Synechococcus*

In this work, we developed a method for and conducted the first isolation of functional β-CBs (Figure 4-1) from the cyanobacterium Syn7942 and the direct visualization of the native organisation (Figure 4-7), topography (Figure 4-9) and intrinsic mechanics (Figure 4-12) of β-CBs using TEM, AFM and proteomics. The method developed here allows for the isolation intact β-CB, where previous isolation methods have resulted in the loss of up to 70% of the shell proteins, or enrichment of β-CB in a relatively impure sample (Long et

al., 2007, 2005). This method can be used as the basis for future projects such as the study of β-CB structure in atomic detail by cryoEM. We find that the intact β-CB possess three distinct structural domains; a single-layered icosahedral shell, an inner layer and paracrystalline arrays of interior Rubisco (Figure 4-7). This ~2 nm wide inner layer (Figure 4-7) is likely comprised of mostly CcmM proteins, due to the ability of CcmM to bind both the shell proteins and Rubisco with opposing termini (Long et al., 2007), and its high abundance in purified β-CB (Table 4-3). Unfortunately the ratio of CcmM 58 to CcmM 35 isoforms could not be determined from our proteomic analysis due to the lack of unique peptides, so it is not possible to predict what the ratio of these may be in the inner layer, although it is likely more CcmM 58 than CcmM 35 as only the full length peptide can link the shell and Rubisco enzymes (Long et al., 2011, 2010). It is probable this layer contains other proteins such as CcmN or CcaA. CcmN may also function as a linker between the shell and Rubisco (Kinney et al., 2012) although interestingly it was not detected in these isolated β-CBs either by SDS-PAGE or by proteomic analyses. CcaA is relatively abundant as well and it has previously been suggested that CcaA forms a "Multiprotein Bicarbonate Dehydration Complex" with CcmM and CcmN (Cot et al., 2008). This sort of complex could be responsible for the 2.0 nm gap and could also explain why the outermost 2-3 layers of Rubisco closest to the shell are the best defined in our TEM images (Figure 4-7 and Figure 5-7). In order to truly understand the composition and organisation of the shell, inner layer, and lumen in molecular detail further structural study is required such as cryoEM or coarse-grained monte carlo simulations (Katsoulakis et al., 2003) of β-CB assembly. Both of these techniques have been successfully applied to the study of viral capsid structure (Grime et al., 2016; Jiang and Tang, 2017).

We also characterized partial β-CB structures that consist of shell facets, shell-associated proteins as well as Rubisco enzymes (Figure 4-2), with retained carbon fixation activity (Figure 4-1), probably serving as the assembly intermediates of β-CBs. In addition, we applied AFM to directly characterize the native protein organisation of shell facets and the topography and intrinsic mechanics of native β-CBs for the first time (Figure 4-4 and Figure 4-5). Our results illustrate the soft nanomechanical properties of β-CBs (Figure 4-12

and Table 4-5) compared to icosahedral P22 VLP (Figure 4-14), and other icosahedral viruses (Table 4-1), likely revealing the unique spatial organisation of carboxysomal building blocks. The study provides new insights into the assembly, organisation and physical nature of functional β-CBs, which has been extended to α-CBs (chapter 5) and can also be extended to other BMCs. Comprehensive understanding of the CB structure and mechanics will underpin the design and engineering of functional synthetic CBs, to enhance photosynthesis and develop new bio-nanoreactors and protein scaffolds using BMC proteins as nanoscale materials. It offers a powerful approach for assaying the functional organisation and material mechanics of natural and engineered biological systems.

6.4 The direct comparison of α and β-carboxysomes

This study directly compared the structure and function of two different CBs using AFM under the same conditions. In doing so, the mechanical properties, Young's modulus and spring constant of CBs under AFM-NI have been identified as potential indicators of the structural integrity of CBs (Figure 5-4). These parameters can be used to differentiate intact and damaged CBs (Figure 5-4 and Figure 5-5), and to compare CBs to other BMCs or synthetic to native CBs in future work. Using AFM-NI, α-CBs were found to be stiffer than β-CBs, and both α-CBs and β-CBs were less stiff than P22 VLP (Figure 4-14) or viruses from similar NI studies (Table 4-1). How these CBs compare to BMCs such as the Pdu or Eut has yet to be investigated and would be a good study for future work in this area. It would also be useful to compare the α- and β- CBs from this work to the α- and β- CBs from other bacteria in future work, to see how much of the variation observed here is due to the differences between two bacterial species and how much is due to the differences between the two types of CBs. Additional study is also required to determine the link between these mechanical properties and the enzymatic activity of BMCs. It could be developed as a biomarker for evaluating BMC-based systems generated via synthetic

biology. BMCs have already been heterologously expressed (Bonacci et al., 2012; Kerfeld, 2016; Yung et al., 2017). For future applications it is vital to know if these BMCs are assembly correctly and capable of maximum enzymatic activity. AFM NI could be used to compare synthetic BMCs with native BMCs to quantify the structural variation caused by the heterologous expression and evaluate the reproducibility of heterologous BMC biogenesis.

The α- and β- CB display different carbon fixation kinetics (Figure 5-2). The maximum turnover rate (Vmax) is higher for β-CB but the substrate affinity is higher for α-CB. These differences may have arisen due to differences in the cytosolic conditions within Syn7942 and *H. neap*, but the Lineweaver-Burk analysis (Lineweaver and Burk, 1934) indicates that the α-CB will function better under conditions where the available HCO_3^- concentration is low and the β-CB will function better when the available HCO_3^- concentration is high (Figure 5-2). Whether this difference is due to the variations in internal enzyme organisation (Figure 5-7, Rae et al., 2013) or another factor remains to be examined. In future work, it would be interesting to replace the Rubisco in these CBs with the alternative form and see if the carbon fixation kinetics or mechanical properties change accordingly. For example, does an α-CB containing form 1Bc Rubisco resemble a β-CB's luminal organisation and kinetics? It may require additional proteins such as Rubisco chaperones or Rubisco assembly factors. Would expressing proteins such as RbcX (Emlyn-Jones et al., 2006; Saschenbrecker et al., 2007) and Raf1 (Hauser et al., 2015; Kolesinski et al., 2017) alongside the α-CB make the form 1Ac Rubisco organise in paracrystalline arrays and increase Vmax?

6.5 Concluding remarks

There are still many open questions as to how BMC's self-assemble. In this work, HS-AFM elucidated some additional details behind the mechanisms of BMC-H proteins self-assembling into BMC-facet like sheets. This technique has proved to be a useful tool in the study of the dynamic processes of BMC self-assembly and could be used in a range of future studies. A combination of techniques were used to improve our understanding of the organisation, structure and function of β-CBs isolated by a newly developed method. This same methodology was subsequently applied to α-CBs, enabling the first direct comparison of these two BMC's. The mechanical properties of these CBs have been determined and have been identified as potential biomarkers of correct self-assembly of BMC. The mechanical properties can be used in future studies to validate heterologously expressed, synthetic or in vitro assembled. The enzyme kinetics of purified α and β CBs have been calculated. The difference between the enzymatic properties of these two CBs provides evidence that some particular BMCs' may be better suited to specific environments. This environmental preference may be an important consideration during future synthetic biology approaches to design a BMC.

*"**The greatest obstacle to discovery is not ignorance** -- **it is the illusion of knowledge**. Never tell people how to do things. Tell them what to do and they will surprise you with their ingenuity."*

***Daniel J. Boorstin**, 1984.*

Bibliography

Abdul-Rahman1, F., Petit1, E., Blanchard, J.L., 2013. The Distribution of Polyhedral Bacterial Microcompartments Suggests Frequent Horizontal Transfer and Operon Reassembly. J. Phylogenetics Evol. Biol. 1, 1–7. https //doi.org/10.4172/2329-9002.1000118

Adrian, M., Dubochet, J., Lepault, J., McDowall, A.W., 1984. Cryo-electron microscopy of viruses. Nature 308, 32–36. https //doi.org/10.1038/308032a0

Agirrezabala, X., Méndez-López, E., Lasso, G., Sánchez-Pina, M.A., Aranda, M., Valle, M., 2015. The near-atomic cryoEM structure of a flexible filamentous plant virus shows homology of its coat protein with nucleoproteins of animal viruses. eLife 4, e11795. https //doi.org/10.7554/eLife.11795

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. The Compartmentalization of Cells.

Alfreider, A., Schirmer, M., Vogt, C., 2012. Diversity and expression of different forms of RubisCO genes in polluted groundwater under different redox conditions. FEMS Microbiol. Ecol. 79, 649–660. https //doi.org/10.1111/j.1574-6941.2011.01246.x

Allemandou, F., Nussberger, J., Brunner, H.R., Brakch, N., 2003. Rapid Site-Directed Mutagenesis Using Two-PCR-Generated DNA Fragments Reproducing the Plasmid Template. J. Biomed. Biotechnol. 2003, 202–207. https //doi.org/10.1155/S1110724303209141

Allison, D.P., Mortensen, N.P., Sullivan, C.J., Doktycz, M.J., 2010. Atomic force microscopy of biological samples. Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol. 2, 618–634. https //doi.org/10.1002/wnan.104

Ando, T., 2014. High-speed AFM imaging. Curr. Opin. Struct. Biol. 28, 63–68. https //doi.org/10.1016/j.sbi.2014.07.011

Ando, T., Uchihashi, T., Kodera, N., 2013. High-speed AFM and applications to biomolecular systems. Annu. Rev. Biophys. 42, 393–414. https //doi.org/10.1146/annurev-biophys-083012-130324

Arteni, A.A., Liu, L.-N., Aartsma, T.J., Zhang, Y.-Z., Zhou, B.-C., Boekema, E.J., 2008. Structure and organisation of phycobilisomes on membranes of the red alga Porphyridium cruentum. Photosynth. Res. 95, 169–174. https //doi.org/10.1007/s11120-007-9264-z

Aussignargues, C., Paasch, B.C., Gonzalez-Esquer, R., Erbilgin, O., Kerfeld, C.A., 2015. Bacterial microcompartment assembly The key role of encapsulation peptides. Commun. Integr. Biol. 8, e1039755. https //doi.org/10.1080/19420889.2015.1039755

Aussignargues, C., Pandelia, M.-E., Sutter, M., Plegaria, J.S., Zarzycki, J., Turmo, A., Huang, J., Ducat, D.C., Hegg, E.L., Gibney, B.R., Kerfeld, C.A., 2016. Structure and Function of a Bacterial Microcompartment Shell Protein Engineered to Bind a [4Fe-4S] Cluster. J. Am. Chem. Soc. 138, 5262–5270. https //doi.org/10.1021/jacs.5b11734

Axen, S.D., Erbilgin, O., Kerfeld, C.A., 2014. A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. PLoS Comput. Biol. 10, e1003898. https //doi.org/10.1371/journal.pcbi.1003898

Badger, M.R., Bek, E.J., 2008. Multiple Rubisco forms in proteobacteria their functional significance in relation to CO2 acquisition by the CBB cycle. J. Exp. Bot. 59, 1525–1541. https //doi.org/10.1093/jxb/erm297

Baker, S.H., Jin, S., Aldrich, H.C., Howard, G.T., Shively, J.M., 1998. Insertion Mutation of the Form I cbbL Gene Encoding Ribulose Bisphosphate Carboxylase/Oxygenase (RuBisCO) in Thiobacillus neapolitanus Results in Expression of Form II RuBisCO, Loss of Carboxysomes, and an Increased CO2 Requirement for Growth. J. Bacteriol. 180, 4133–4139.

Bassham, J.A., Benson, A.A., Calvin, M., 1950. The path of carbon in photosynthesis. J. Biol. Chem. 185, 781–787.

Baumgart, M., Huber, I., Abdollahzadeh, I., Gensch, T., Frunzke, J., 2017. Heterologous expression of the Halothiobacillus neapolitanus carboxysomal gene cluster in Corynebacterium glutamicum. J. Biotechnol. https //doi.org/10.1016/j.jbiotec.2017.03.019

Bhardwaj, A., Olia, A.S., Cingolani, G., 2014. Architecture of viral genome-delivery molecular machines. Curr. Opin. Struct. Biol. 25, 1–8. https //doi.org/10.1016/j.sbi.2013.10.005

Bolhassani, A., Davoudi, N., Agi, E., 2017. Comparison of HCV core and coreE1E2 virus-like particles generated by stably transfected Leishmania tarentolae for stimulation of Th1 immune responses in mice. Curr. Drug Deliv.

Bonacci, W., Teng, P.K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P.A., Savage, D.F., 2012. Modularity of a carbon-fixing protein organelle. Proc. Natl. Acad. Sci. U. S. A. 109, 478–483. https //doi.org/10.1073/pnas.1108557109

Boyd, K.J., Bansal, P., Feng, J., May, E.R., 2015. Stability of Norwalk Virus Capsid Protein Interfaces Evaluated by in Silico Nanoindentation. Front. Bioeng. Biotechnol. 3, 103. https //doi.org/10.3389/fbioe.2015.00103

Bugbee, B.G., Salisbury, F.B., 1988. Exploring the limits of crop productivity. I. Photosynthetic efficiency of wheat in high irradiance environments. Plant Physiol. 88, 869–878.

Cai, F., Bernstein, S.L., Wilson, S.C., Kerfeld, C.A., 2016. Production and Characterization of Synthetic Carboxysome Shells with Incorporated Luminal Proteins. Plant Physiol. 170, 1868–1877. https //doi.org/10.1104/pp.15.01822

Cai, F., Dou, Z., Bernstein, S.L., Leverenz, R., Williams, E.B., Heinhorst, S., Shively, J., Cannon, G.C., Kerfeld, C.A., 2015a. Advances in Understanding Carboxysome Assembly in Prochlorococcus and Synechococcus Implicate CsoS2 as a Critical Component. Life Basel Switz. 5, 1141–1171. https //doi.org/10.3390/life5021141

Cai, F., Kerfeld, C.A., Sandh, G., 2012. Bioinformatic Identification and Structural Characterization of a New Carboxysome Shell Protein, in Functional Genomics and Evolution of Photosynthetic Systems, Advances in Photosynthesis and Respiration. Springer, Dordrecht, pp. 345–356.

Cai, F., Menon, B.B., Cannon, G.C., Curry, K.J., Shively, J.M., Heinhorst, S., 2009. The Pentameric Vertex Proteins Are Necessary for the Icosahedral Carboxysome Shell to Function as a CO2 Leakage Barrier. PLoS ONE 4. https //doi.org/10.1371/journal.pone.0007521

Cai, F., Sutter, M., Bernstein, S.L., Kinney, J.N., Kerfeld, C.A., 2015b. Engineering bacterial microcompartment shells chimeric shell proteins and chimeric carboxysome shells. ACS Synth. Biol. 4, 444–453. https //doi.org/10.1021/sb500226j

Cai, F., Sutter, M., Cameron, J.C., Stanley, D.N., Kinney, J.N., Kerfeld, C.A., 2013. The structure of CcmP, a tandem bacterial microcompartment domain protein from the β-carboxysome, forms a subcompartment within a microcompartment. J. Biol. Chem. 288, 16055–16063. https //doi.org/10.1074/jbc.M113.456897

Cameron, J.C., Wilson, S.C., Bernstein, S.L., Kerfeld, C.A., 2013. Biogenesis of a bacterial organelle the carboxysome assembly pathway. Cell 155, 1131–1140. https //doi.org/10.1016/j.cell.2013.10.044

Cannon, G.C., Bradburne, C.E., Aldrich, H.C., Baker, S.H., Heinhorst, S., Shively, J.M., 2001. Microcompartments in Prokaryotes Carboxysomes and Related Polyhedra. Appl. Environ. Microbiol. 67, 5351–5361. https //doi.org/10.1128/AEM.67.12.5351-5361.2001

Cannon, G.C., Shively, J.M., 1983. Characterization of a homogenous preparation of carboxysomes from Thiobacillus neapolitanus. Arch. Microbiol. 134, 52–59. https //doi.org/10.1007/BF00429407

Cardona, A., Schmid, B., Rueden, C., White, D.J., Frise, E., Arganda-Carreras, I., Tinevez, J.-Y., Schindelin, J., Eliceiri, K., Longair, M., Tomancak, P., Preibisch, S., Saalfeld, S., Pietzsch, T., Kaynig, V., Hartenstein, V., 2012. Fiji an open-source platform for biological-image analysis. Nat. Methods 9, 676. https //doi.org/10.1038/nmeth.2019

Carrasco, C., Carreira, A., Schaap, I.A.T., Serena, P.A., Gómez-Herrero, J., Mateu, M.G., de Pablo, P.J., 2006. DNA-mediated anisotropic mechanical reinforcement of a virus. Proc. Natl. Acad. Sci. U. S. A. 103, 13706–13711. https //doi.org/10.1073/pnas.0601881103

Casuso, I., Rico, F., Scheuring, S., 2011. High-speed atomic force microscopy Structure and dynamics of single proteins. Curr. Opin. Chem. Biol. 15, 704–709. https //doi.org/10.1016/j.cbpa.2011.05.007

Chang, K.-C., Chiang, Y.-W., Yang, C.-H., Liou, J.-W., 2012. Atomic force microscopy in biology and biomedicine. Tzu Chi Med. J. 24, 162–169. https //doi.org/10.1016/j.tcmj.2012.08.002

Chen, A.H., Robinson-Mosher, A., Savage, D.F., Silver, P.A., Polka, J.K., 2013. The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. PloS One 8, e76127. https //doi.org/10.1371/journal.pone.0076127

Cheng, S., Liu, Y., Crowley, C.S., Yeates, T.O., Bobik, T.A., 2008. Bacterial microcompartments their properties and paradoxes. Bioessays 30, 1084–1095. https //doi.org/10.1002/bies.20830

Chessher, A., Breitling, R., Takano, E., 2015. Bacterial Microcompartments Biomaterials for Synthetic Biology-Based Compartmentalization Strategies. ACS Biomater. Sci. Eng. 1, 345–351. https //doi.org/10.1021/acsbiomaterials.5b00059

Chowdhury, C., Chun, S., Pang, A., Sawaya, M.R., Sinha, S., Yeates, T.O., Bobik, T.A., 2015. Selective molecular transport through the protein shell of a bacterial microcompartment organelle. Proc. Natl. Acad. Sci. U. S. A. 112, 2990–2995. https //doi.org/10.1073/pnas.1423672112

Chowdhury, C., Sinha, S., Chun, S., Yeates, T.O., Bobik, T.A., 2014. Diverse Bacterial Microcompartment Organelles. Microbiol. Mol. Biol. Rev. MMBR 78, 438–468. https //doi.org/10.1128/MMBR.00009-14

Cot, S.S.-W., So, A.K.-C., Espie, G.S., 2008. A Multiprotein Bicarbonate Dehydration Complex Essential to Carboxysome Function in Cyanobacteria. J. Bacteriol. 190, 936–945. https //doi.org/10.1128/JB.01283-07

Crowley, C.S., Cascio, D., Sawaya, M.R., Kopstein, J.S., Bobik, T.A., Yeates, T.O., 2010. Structural insight into the mechanisms of transport across the Salmonella enterica Pdu microcompartment shell. J. Biol. Chem. 285, 37838–37846. https //doi.org/10.1074/jbc.M110.160580

Crowley, C.S., Sawaya, M.R., Bobik, T.A., Yeates, T.O., 2008. Structure of the PduU shell protein from the Pdu microcompartment of Salmonella. Struct. Lond. Engl. 1993 16, 1324–1332. https //doi.org/10.1016/j.str.2008.05.013

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. 97, 6640–6645. https //doi.org/10.1073/pnas.120163297

Diekmann, Y., Pereira-Leal, J.B., 2013. Evolution of intracellular compartmentalization. Biochem. J. 449, 319–331. https //doi.org/10.1042/BJ20120957

Dou, Z., Heinhorst, S., Williams, E.B., Murin, C.D., Shively, J.M., Cannon, G.C., 2008. CO2 fixation kinetics of Halothiobacillus neapolitanus mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO2. J. Biol. Chem. 283, 10377–10384. https //doi.org/10.1074/jbc.M709285200

Drews, G., Niklowitz, W., 1956. Beiträge zur Cytologie der Blaualgen. Arch. Für Mikrobiol. 24, 147–162. https //doi.org/10.1007/BF00408629

Dümmler, A., Lawrence, A.-M., de Marco, A., 2005. Simplified screening for the detection of soluble fusion constructs expressed in E. coli using a modular set of vectors. Microb. Cell Factories 4, 34. https //doi.org/10.1186/1475-2859-4-34

Eghiaian, F., Rico, F., Colom, A., Casuso, I., Scheuring, S., 2014. High-speed atomic force microscopy imaging and force spectroscopy. FEBS Lett. 588, 3631–3638. https //doi.org/10.1016/j.febslet.2014.06.028

Emlyn-Jones, D., Woodger, F.J., Price, G.D., Whitney, S.M., 2006. RbcX Can Function as a Rubisco Chaperonin, But is Non-Essential in Synechococcus PCC7942. Plant Cell Physiol. 47, 1630–1640. https //doi.org/10.1093/pcp/pcl028

English, R.S., Jin, S., Shively, J.M., 1995. Use of Electroporation To Generate a Thiobacillus neapolitanus Carboxysome Mutant. Appl. Environ. Microbiol. 61, 3256–3260.

Fan, C., Bobik, T.A., 2011. The N-Terminal Region of the Medium Subunit (PduD) Packages Adenosylcobalamin-Dependent Diol Dehydratase (PduCDE) into the Pdu Microcompartment ▿ . J. Bacteriol. 193, 5623–5628. https //doi.org/10.1128/JB.05661-11

Fan, C., Cheng, S., Sinha, S., Bobik, T.A., 2012. Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. Proc. Natl. Acad. Sci. U. S. A. 109, 14995–15000. https //doi.org/10.1073/pnas.1207516109

Faulkner, M., Rodriguez-Ramos, J., Dykes, G.F., Owen, S.V., Casella, S., Simpson, D.M., Beynon, R.J., Liu, L.-N., 2017. Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes. Nanoscale. https //doi.org/10.1039/C7NR02524F

Frank, S., Lawrence, A.D., Prentice, M.B., Warren, M.J., 2013. Bacterial microcompartments moving into a synthetic biological world. J. Biotechnol., From Gene to Product 163, 273–279. https //doi.org/10.1016/j.jbiotec.2012.09.002

Frenkel, A., Gaffron, H., Battley, E.H., 1950. Photosynthesis and photoreduction by the blue green alga, Synechococcus elongatus, Näg. Biol. Bull. 99, 157–162. https //doi.org/10.2307/1538735

Frey, R., Mantri, S., Rocca, M., Hilvert, D., 2016. Bottom-up Construction of a Primordial Carboxysome Mimic. J. Am. Chem. Soc. 138, 10072–10075. https //doi.org/10.1021/jacs.6b04744

Froger, A., Hall, J.E., 2007. Transformation of Plasmid DNA into E. coli Using the Heat Shock Method. J. Vis. Exp. JoVE. https //doi.org/10.3791/253

Galmés, J., Kapralov, M.V., Andralojc, P.J., Conesa, M.À., Keys, A.J., Parry, M. a. J., Flexas, J., 2014. Expanding knowledge of the Rubisco kinetics variability in plant species environmental and evolutionary trends. Plant Cell Environ. 37, 1989–2001. https //doi.org/10.1111/pce.12335

Gantt, E., Conti, S.F., 1969. Ultrastructure of Blue-Green Algae. J. Bacteriol. 97, 1486–1493.

Garcia-Alles, L.F., Lesniewska, E., Root, K., Aubry, N., Pocholle, N., Mendoza, C.I., Bourillot, E., Barylyuk, K., Pompon, D., Zenobi, R., Reguera, D., Truan, G., 2017. Spontaneous non-canonical assembly of CcmK hexameric components from β-carboxysome shells of cyanobacteria. PLOS ONE 12, e0185109. https //doi.org/10.1371/journal.pone.0185109

Golden, S.S., 1988. Mutagenesis of cyanobacteria by classical and gene-transfer-based methods. Methods Enzymol. 167, 714–727.

Gonzalez-Esquer, C.R., Shubitowski, T.B., Kerfeld, C.A., 2015. Streamlined Construction of the Cyanobacterial CO2-Fixing Organelle via Protein Domain Fusions for Use in Plant Synthetic Biology. Plant Cell 27, 2637–2644. https //doi.org/10.1105/tpc.15.00329

Gowik, U., Westhoff, P., 2011. The Path from C3 to C4 Photosynthesis. Plant Physiol. 155, 56–63. https //doi.org/10.1104/pp.110.165308

Grime, J.M.A., Dama, J.F., Ganser-Pornillos, B.K., Woodward, C.L., Jensen, G.J., Yeager, M., Voth, G.A., 2016. Coarse-grained simulation reveals key features of HIV-1 capsid self-assembly. Nat. Commun. 7, 11568. https //doi.org/10.1038/ncomms11568

Gust, B., Chandra, G., Jakimowicz, D., Yuqing, T., Bruton, C.J., Chater, K.F., 2004. Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces. Adv. Appl. Microbiol. 54, 107–128. https //doi.org/10.1016/S0065-2164(04)54004-2

Hanson, M.R., Lin, M.T., Carmo-Silva, A.E., Parry, M.A.J., 2016. Towards engineering carboxysomes into C3 plants. Plant J. Cell Mol. Biol. https //doi.org/10.1111/tpj.13139

Harris, G.C., Königer, M., 1997. The “high” concentrations of enzymes within the chloroplast. Photosynth. Res. 54, 5–23. https //doi.org/10.1023/A 1005895213775

Hauser, T., Bhat, J.Y., Miličić, G., Wendler, P., Hartl, F.U., Bracher, A., Hayer-Hartl, M., 2015. Structure and mechanism of the Rubisco-assembly chaperone Raf1. Nat. Struct. Mol. Biol. 22, 720–728. https //doi.org/10.1038/nsmb.3062

Heldt, D., Frank, S., Seyedarabi, A., Ladikis, D., Parsons, J.B., Warren, M.J., Pickersgill, R.W., 2009. Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol utilization in Clostridium kluyveri. Biochem. J. 423, 199–207. https //doi.org/10.1042/BJ20090780

Hernando-Pérez, M., Zeng, C., Delalande, L., Tsvetkova, I.B., Bousquet, A., Tayachi-Pigeonnat, M., Temam, R., Dragnea, B., 2016. Nanoindentation of Isometric Viruses on Deterministically Corrugated Substrates. J. Phys. Chem. B 120, 340–347. https //doi.org/10.1021/acs.jpcb.5b08362

Holthuijzen, Y.A., Breemen, J.F.L. van, Konings, W.N., Bruggen, E.F.J. van, 1986. Electron microscopic studies of carboxysomes of <Emphasis Type="Italic">Thiobacillus neapolitanus</Emphasis>. Arch. Microbiol. 144, 258–262. https //doi.org/10.1007/BF00410959

Hood, R.D., Higgins, S.A., Flamholz, A., Nichols, R.J., Savage, D.F., 2016. The stringent response regulates adaptation to darkness in the cyanobacterium Synechococcus elongatus. Proc. Natl. Acad. Sci. U. S. A. 113, E4867–4876. https //doi.org/10.1073/pnas.1524915113

Hopkinson, B.M., Young, J.N., Tansik, A.L., Binder, B.J., 2014. The minimal CO2-concentrating mechanism of Prochlorococcus spp. MED4 is effective and efficient. Plant Physiol. 166, 2205–2217. https //doi.org/10.1104/pp.114.247049

Horcas, I., Fernández, R., Gómez-Rodríguez, J.M., Colchero, J., Gómez-Herrero, J., Baro, A.M., 2007. WSXM A software for scanning probe microscopy and a tool for nanotechnology. Rev. Sci. Instrum. 78, 013705. https //doi.org/10.1063/1.2432410

Howorka, S., 2011. Molecular Assembly in Natural and Engineered Systems. Academic Press.

Huber, I., Palmer, D.J., Ludwig, K.N., Brown, I.R., Warren, M.J., Frunzke, J., 2017. Construction of Recombinant Pdu Metabolosome Shells for Small Molecule Production in Corynebacterium glutamicum. ACS Synth. Biol. 6, 2145–2156. https //doi.org/10.1021/acssynbio.7b00167

Hutchinson, M., Johnstone, K.I., White, D., 1967. Taxonomy of anaerobic thiobacilli. J. Gen. Microbiol. 47, 17–23. https //doi.org/10.1099/00221287-47-1-17

Hutchinson, M., Johnstone, K.I., White, D., 1965. The taxonomy of certain thiobacilli. J. Gen. Microbiol. 41, 357–366. https //doi.org/10.1099/00221287-41-3-357

Iancu, C.V., Ding, H.J., Morris, D.M., Dias, D.P., Gonzales, A.D., Martino, A., Jensen, G.J., 2007. The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography. J. Mol. Biol. 372, 764–773. https //doi.org/10.1016/j.jmb.2007.06.059

Iancu, C.V., Morris, D.M., Dou, Z., Heinhorst, S., Cannon, G.C., Jensen, G.J., 2010. Organisation, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells. J. Mol. Biol. 396, 105–117. https //doi.org/10.1016/j.jmb.2009.11.019

Imamura, M., Uchihashi, T., Ando, T., Leifert, A., Simon, U., Malay, A.D., Heddle, J.G., 2015. Probing structural dynamics of an artificial protein cage using high-speed atomic force microscopy. Nano Lett. 15, 1331–1335. https //doi.org/10.1021/nl5045617

Ivanovska, I.L., de Pablo, P.J., Ibarra, B., Sgalari, G., MacKintosh, F.C., Carrascosa, J.L., Schmidt, C.F., Wuite, G.J.L., 2004. Bacteriophage capsids Tough nanoshells with complex elastic properties. Proc. Natl. Acad. Sci. U. S. A. 101, 7600–7605. https //doi.org/10.1073/pnas.0308198101

Ivanovska, I., Wuite, G., Jönsson, B., Evilevitch, A., 2007. Internal DNA pressure modifies stability of WT phage. Proc. Natl. Acad. Sci. U. S. A. 104, 9603–9608. https //doi.org/10.1073/pnas.0703166104

Jensen, P.E., Leister, D., 2014. Cyanobacteria as an Experimental Platform for Modifying Bacterial and Plant Photosynthesis. Front. Bioeng. Biotechnol. 2. https //doi.org/10.3389/fbioe.2014.00007

Jiang, W., Tang, L., 2017. Atomic cryo-EM structures of viruses. Curr. Opin. Struct. Biol., Cryo electron microscopy exciting advances in CryoEM herald a new era in structural biology • Biophysical methods behind the scenes of the cryo-EM revolution 46, 122–129. https //doi.org/10.1016/j.sbi.2017.07.002

Johnson, K.L., Johnson, K.L., 1987. Contact Mechanics. Cambridge University Press.

Kaneko, Y., Danev, R., Nagayama, K., Nakamoto, H., 2006. Intact carboxysomes in a cyanobacterial cell visualized by hilbert differential contrast transmission electron microscopy. J. Bacteriol. 188, 805–808. https //doi.org/10.1128/JB.188.2.805-808.2006

Katsoulakis, M.A., Majda, A.J., Vlachos, D.G., 2003. Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems. J. Comput. Phys. 186, 250–278. https //doi.org/10.1016/S0021-9991(03)00051-2

Kelly, D.P., Wood, A.P., 2000. Reclassification of some species of Thiobacillus to the newly designated genera Acidithiobacillus gen. nov., Halothiobacillus gen. nov. and Thermithiobacillus gen. nov. Int. J. Syst. Evol. Microbiol. 50, 2, 511–516. https //doi.org/10.1099/00207713-50-2-511

Kent, S.S., Tomany, M.J., 1984. Kinetic Variance of Ribulose-1,5-bisphosphate Carboxylase/Oxygenase Isolated from Diverse Taxonomic Sources. Plant Physiol. 75, 645–650.

Kerfeld, C.A., 2016. Rewiring Escherichia coli for carbon-dioxide fixation. Nat. Biotechnol. 34, 1035–1036. https //doi.org/10.1038/nbt.3693

Kerfeld, C.A., Erbilgin, O., 2015. Bacterial microcompartments and the modular construction of microbial metabolism. Trends Microbiol. 23, 22–34. https //doi.org/10.1016/j.tim.2014.10.003

Kerfeld, C.A., Heinhorst, S., Cannon, G.C., 2010. Bacterial microcompartments. Annu. Rev. Microbiol. 64, 391–408. https //doi.org/10.1146/annurev.micro.112408.134211

Kerfeld, C.A., Melnicki, M.R., 2016. Assembly, function and evolution of cyanobacterial carboxysomes. Curr. Opin. Plant Biol. 31, 66–75. https //doi.org/10.1016/j.pbi.2016.03.009

Kerfeld, C.A., Sawaya, M.R., Tanaka, S., Nguyen, C.V., Phillips, M., Beeby, M., Yeates, T.O., 2005. Protein structures forming the shell of primitive bacterial organelles. Science 309, 936–938. https //doi.org/10.1126/science.1113397

Kim, E.Y., Slininger, M.F., Tullman-Ercek, D., 2014. The effects of time, temperature, and pH on the stability of PDU bacterial microcompartments. Protein Sci. 23, 1434–1441. https //doi.org/10.1002/pro.2527

Kinney, J.N., Salmeen, A., Cai, F., Kerfeld, C.A., 2012. Elucidating Essential Role of Conserved Carboxysomal Protein CcmN Reveals Common Feature of Bacterial Microcompartment Assembly. J. Biol. Chem. 287, 17729–17736. https //doi.org/10.1074/jbc.M112.355305

Kiracofe, D., Melcher, J., Raman, A., Balasubramaniam, S., Johnson, S., Hu, S., 2014. VEDA Virtual Environment for Dynamic AFM. Rev Sci Instrum. 2008 Jun;79(6):061301. doi: 10.1063/1.2938864.

Kirst, H., Formighieri, C., Melis, A., 2014. Maximizing photosynthetic efficiency and culture productivity in cyanobacteria upon minimising the phycobilisome light-harvesting antenna size. Biochim. Biophys. Acta 1837, 1653–1664. https //doi.org/10.1016/j.bbabio.2014.07.009

Klein, M.G., Zwart, P., Bagby, S.C., Cai, F., Chisholm, S.W., Heinhorst, S., Cannon, G.C., Kerfeld, C.A., 2009. Identification and Structural Analysis of a Novel Carboxysome Shell Protein with Implications for Metabolite Transport. J. Mol. Biol. 392, 319–333. https //doi.org/10.1016/j.jmb.2009.03.056

Kolesinski, P., Rydzy, M., Szczepaniak, A., 2017. Is RAF1 protein from Synechocystis sp. PCC 6803 really needed in the cyanobacterial Rubisco assembly process? Photosynth. Res. 132, 135–148. https //doi.org/10.1007/s11120-017-0336-4

Krupovic, M., Koonin, E.V., 2017. Cellular origin of the viral capsid-like bacterial microcompartments. Biol. Direct 12, 25. https //doi.org/10.1186/s13062-017-0197-y

Kuan, D., Duff, S., Posarac, D., Bi, X., 2015. Growth optimization of Synechococcus elongatus PCC7942 in lab flasks and a 2-D photobioreactor. Can. J. Chem. Eng. 93, 640–647. https //doi.org/10.1002/cjce.22154

Larsson, A.M., Hasse, D., Valegård, K., Andersson, I., 2017. Crystal structures of β-carboxysome shell protein CcmP ligand binding correlates with the closed or open central pore. J. Exp. Bot. 1;68, 3857–3867. https //doi.org/10.1093/jxb/erx070

Lassila, J.K., Bernstein, S.L., Kinney, J.N., Axen, S.D., Kerfeld, C.A., 2014. Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function. J. Mol. Biol. 426, 2217–2228. https //doi.org/10.1016/j.jmb.2014.02.025

Lawrence, A.D., Frank, S., Newnham, S., Lee, M.J., Brown, I.R., Xue, W.-F., Rowe, M.L., Mulvihill, D.P., Prentice, M.B., Howard, M.J., Warren, M.J., 2014. Solution structure of a bacterial microcompartment targeting peptide and its application in the construction of an ethanol bioreactor. ACS Synth. Biol. 3, 454–465. https //doi.org/10.1021/sb4001118

Lee, K.K., Gui, L., 2016. Dissecting Virus Infectious Cycles by Cryo-Electron Microscopy. PLOS Pathog. 12, e1005625. https //doi.org/10.1371/journal.ppat.1005625

Lee, M.J., Brown, I.R., Juodeikis, R., Frank, S., Warren, M.J., 2016. Employing bacterial microcompartment technology to engineer a shell-free enzyme-aggregate for enhanced 1,2-propanediol production in Escherichia coli. Metab. Eng. 36, 48–56. https //doi.org/10.1016/j.ymben.2016.02.007

Lee, M.J., Mantell, J., Hodgson, L., Alibhai, D., Fletcher, J.M., Brown, I.R., Frank, S., Xue, W.-F., Verkade, P., Woolfson, D.N., Warren, M.J., 2017. Engineered synthetic scaffolds for organising proteins within the bacterial cytoplasm. Nat. Chem. Biol. 14(2):142-147 https //doi.org/10.1038/nchembio.2535

Liang, M., Frank, S., Lünsdorf, H., Warren, M.J., Prentice, M.B., 2017. Bacterial microcompartment-directed polyphosphate kinase promotes stable polyphosphate accumulation in E. coli. Biotechnol. J. 12. https //doi.org/10.1002/biot.201600415

Lineweaver, H., Burk, D., 1934a. The Determination of Enzyme Dissociation Constants. J. Am. Chem. Soc. 56, 658–666. https //doi.org/10.1021/ja01318a036

Lin, M.T., Occhialini, A., Andralojc, P.J., Devonshire, J., Hines, K.M., Parry, M.A.J., Hanson, M.R., 2014a. β-Carboxysomal proteins assemble into highly organised structures in Nicotiana chloroplasts. Plant J. Cell Mol. Biol. 79, 1–12. https //doi.org/10.1111/tpj.12536

Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A.J., Hanson, M.R., 2014b. A faster Rubisco with potential to increase photosynthesis in crops. Nature 513, 547–550. https //doi.org/10.1038/nature13776

Li, S., Sieben, C., Ludwig, K., Höfer, C.T., Chiantia, S., Herrmann, A., Eghiaian, F., Schaap, I.A.T., 2014. pH-Controlled Two-Step Uncoating of Influenza Virus. Biophys. J. 106, 1447–1456. https //doi.org/10.1016/j.bpj.2014.02.018

Liu, L.-N., Aartsma, T.J., Thomas, J.-C., Lamers, G.E.M., Zhou, B.-C., Zhang, Y.-Z., 2008. Watching the native supramolecular architecture of photosynthetic membrane in red algae topography of phycobilisomes and their crowding, diverse distribution patterns. J. Biol. Chem. 283, 34946–34953. https //doi.org/10.1074/jbc.M805114200

Liu, L.-N., Scheuring, S., 2013. Investigation of photosynthetic membrane structure using atomic force microscopy. Trends Plant Sci. 18, 277–286. https //doi.org/10.1016/j.tplants.2013.03.001

Llauró, A., Luque, D., Edwards, E., Trus, B.L., Avera, J., Reguera, D., Douglas, T., Pablo, P.J. de, Castón, J.R., 2016a. Cargo–shell and cargo–cargo couplings govern the mechanics of artificially loaded virus-derived cages. Nanoscale 8, 9328–9336. https //doi.org/10.1039/C6NR01007E

Llauró, A., Schwarz, B., Koliyatt, R., de Pablo, P.J., Douglas, T., 2016b. Tuning Viral Capsid Nanoparticle Stability with Symmetrical Morphogenesis. ACS Nano 10, 8465–8473. https //doi.org/10.1021/acsnano.6b03441

Long, B.M., Badger, M.R., Whitney, S.M., Price, G.D., 2007. Analysis of carboxysomes from Synechococcus PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. J. Biol. Chem. 282, 29323–29335. https //doi.org/10.1074/jbc.M703896200

Long, B.M., Price, G.D., Badger, M.R., 2005. Proteomic assessment of an established technique for carboxysome enrichment from *Synechococcus* PCC7942. Can. J. Bot. 83, 746–757. https //doi.org/10.1139/b05-058

Long, B.M., Rae, B.D., Badger, M.R., Price, G.D., 2011. Over-expression of the β-carboxysomal CcmM protein in Synechococcus PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content. Photosynth. Res. 109, 33–45. https //doi.org/10.1007/s11120-011-9659-8

Long, B.M., Tucker, L., Badger, M.R., Price, G.D., 2010. Functional Cyanobacterial β-Carboxysomes Have an Absolute Requirement for Both Long and Short Forms of the CcmM Protein1[W][OA]. Plant Physiol. 153, 285–293. https //doi.org/10.1104/pp.110.154948

Long, S.P., Zhu, X.-G., Naidu, S.L., Ort, D.R., 2006. Can improvement in photosynthesis increase crop yields? Plant Cell Environ. 29, 315–330. https //doi.org/10.1111/j.1365-3040.2005.01493.x

Mahalik, J.P., Brown, K.A., Cheng, X., Fuentes-Cabrera, M., 2016. Theoretical Study of the Initial Stages of Self-Assembly of a Carboxysome's Facet. ACS Nano 10, 5751–5758. https //doi.org/10.1021/acsnano.5b07805

Mangan, N.M., Brenner, M.P., 2013. Systems analysis of the CO2 concentrating mechanism in cyanobacteria. eLife 3. https //doi.org/10.7554/eLife.02043

Mangan, N.M., Flamholz, A., Hood, R.D., Milo, R., Savage, D.F., 2016. pH determines the energetic efficiency of the cyanobacterial CO2 concentrating mechanism. Proc. Natl. Acad. Sci. U. S. A. 113, E5354–5362. https //doi.org/10.1073/pnas.1525145113

Marchetti, M., Wuite, G., Roos, W.H., 2016. Atomic force microscopy observation and characterization of single virions and virus-like particles by nano-indentation. Curr. Opin. Virol. 18, 82–88. https //doi.org/10.1016/j.coviro.2016.05.002

Martin, W., 2010. Evolutionary origins of metabolic compartmentalization in eukaryotes. Philos. Trans. R. Soc. B Biol. Sci. 365, 847–855. https //doi.org/10.1098/rstb.2009.0252

Mateu, M.G., 2012. Mechanical properties of viruses analyzed by atomic force microscopy A virological perspective. Virus Res. 168, 1–22. https //doi.org/10.1016/j.virusres.2012.06.008

McEwen, J.T., Machado, I.M.P., Connor, M.R., Atsumi, S., 2013. Engineering Synechococcus elongatus PCC 7942 for Continuous Growth under Diurnal Conditions. Appl. Environ. Microbiol. 79, 1668–1675. https //doi.org/10.1128/AEM.03326-12

McNevin, D., von Caemmerer, S., Farquhar, G., 2006. Determining RuBisCO activation kinetics and other rate and equilibrium constants by simultaneous multiple non-linear regression of a kinetic model. J. Exp. Bot. 57, 3883–3900. https //doi.org/10.1093/jxb/erl156

Menon, B.B., Dou, Z., Heinhorst, S., Shively, J.M., Cannon, G.C., 2008. Halothiobacillus neapolitanus Carboxysomes Sequester Heterologous and Chimeric RubisCO Species. PLoS ONE 3. https //doi.org/10.1371/journal.pone.0003570

Michelet, L., Zaffagnini, M., Morisse, S., Sparla, F., Pérez-Pérez, M.E., Francia, F., Danon, A., Marchand, C.H., Fermani, S., Trost, P., Lemaire, S.D., 2013. Redox regulation of the Calvin–Benson cycle something old, something new. Front. Plant Sci. 4. https //doi.org/10.3389/fpls.2013.00470

Michel, J.P., Ivanovska, I.L., Gibbons, M.M., Klug, W.S., Knobler, C.M., Wuite, G.J.L., Schmidt, C.F., 2006. Nanoindentation studies of full and empty viral capsids and the effects of capsid protein mutations on elasticity and strength. Proc. Natl. Acad. Sci. U. S. A. 103, 6184–6189. https //doi.org/10.1073/pnas.0601744103

Mitragotri, S., Lahann, J., 2009. Physical approaches to biomaterial design. Nat. Mater. 8, 15–23. https //doi.org/10.1038/nmat2344

Mosberg, J.A., Lajoie, M.J., Church, G.M., 2010. Lambda Red Recombineering in Escherichia coli Occurs Through a Fully Single-Stranded Intermediate. Genetics 186, 791–799. https //doi.org/10.1534/genetics.110.120782

Nečas, D., Klapetek, P., 2011. Gwyddion an open-source software for SPM data analysis. Open Phys. 10, 181–188. https //doi.org/10.2478/s11534-011-0096-2

Newman, J., Branden, C.I., Jones, T.A., 1993. Structure determination and refinement of ribulose 1,5-bisphosphate carboxylase/oxygenase from Synechococcus PCC6301. Acta CrystallogrSectD 49, 548–560. https //doi.org/10.2210/pdb1rbl/pdb

Occhialini, A., Lin, M.T., Andralojc, P.J., Hanson, M.R., Parry, M.A.J., 2016. Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO2. Plant J. Cell Mol. Biol. 85, 148–160. https //doi.org/10.1111/tpj.13098

Ooms, M.D., Dinh, C.T., Sargent, E.H., Sinton, D., 2016. Photon management for augmented photosynthesis. Nat. Commun. 7, 12699. https //doi.org/10.1038/ncomms12699

Ortega-Esteban, A., Condezo, G.N., Pérez-Berná, A.J., Chillón, M., Flint, S.J., Reguera, D., San Martín, C., de Pablo, P.J., 2015. Mechanics of Viral Chromatin Reveals the Pressurization of Human Adenovirus. ACS Nano 9, 10826–10833. https //doi.org/10.1021/acsnano.5b03417

Orus, M.I., Rodriguez, M.L., Martinez, F., Marco, E., 1995. Biogenesis and Ultrastructure of Carboxysomes from Wild Type and Mutants of Synechococcus sp. Strain PCC 7942. Plant Physiol. 107, 1159–1166.

Pang, A., Frank, S., Brown, I., Warren, M.J., Pickersgill, R.W., 2014. Structural insights into higher-order assembly and function of the bacterial microcompartment protein PduA. J. Biol. Chem. jbc.M114.569285. https //doi.org/10.1074/jbc.M114.569285

Pang, A., Warren, M.J., Pickersgill, R.W., 2011. Structure of PduT, a trimeric bacterial microcompartment protein with a 4Fe–4S cluster-binding site. Acta Crystallogr. D Biol. Crystallogr. 67, 91–96. https //doi.org/10.1107/S0907444910050201

Park, J., Chun, S., Bobik, T.A., Houk, K.N., Yeates, T.O., 2017. Molecular Dynamics Simulations of Selective Metabolite Transport across the Propanediol Bacterial Microcompartment Shell. J. Phys. Chem. B 121, 8149–8154. https //doi.org/10.1021/acs.jpcb.7b07232

Parry, B.R., Surovtsev, I.V., Cabeen, M.T., O'Hern, C.S., Dufresne, E.R., Jacobs-Wagner, C., 2014. The bacterial cytoplasm has glass-like properties and is fluidized by metabolic activity. Cell 156, 183–194. https //doi.org/10.1016/j.cell.2013.11.028

Parsons, J.B., Frank, S., Bhella, D., Liang, M., Prentice, M.B., Mulvihill, D.P., Warren, M.J., 2010. Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. Mol. Cell 38, 305–315. https //doi.org/10.1016/j.molcel.2010.04.008

Perlmutter, J.D., Mohajerani, F., Hagan, M.F., 2016. Many-molecule encapsulation by an icosahedral shell. eLife 5. https //doi.org/10.7554/eLife.14078

Peterhansel, C., Horst, I., Niessen, M., Blume, C., Kebeish, R., Kürkcüoglu, S., Kreuzaler, F., 2010. Photorespiration. Arab. Book Am. Soc. Plant Biol. 8. https //doi.org/10.1199/tab.0130

Plegaria, J.S., Kerfeld, C.A., 2018. Engineering nanoreactors using bacterial microcompartment architectures. Curr. Opin. Biotechnol., Systems biology • Nanobiotechnology 51, 1–7. https //doi.org/10.1016/j.copbio.2017.09.005

Prasad, B.V.V., Schmid, M.F., 2012. Principles of Virus Structural Organisation. Adv. Exp. Med. Biol. 726, 17–47. https //doi.org/10.1007/978-1-4614-0980-9_3

Price, G.D., Badger, M.R., 1989. Isolation and Characterization of High CO2-Requiring-Mutants of the Cyanobacterium Synechococcus PCC7942 1. Plant Physiol. 91, 514–525.

Price, G.D., Pengelly, J.J.L., Forster, B., Du, J., Whitney, S.M., von Caemmerer, S., Badger, M.R., Howitt, S.M., Evans, J.R., 2013. The cyanobacterial CCM as a source of genes for improving photosynthetic CO2 fixation in crop species. J. Exp. Bot. 64, 753–768. https //doi.org/10.1093/jxb/ers257

Quin, M.B., Perdue, S.A., Hsu, S.-Y., Schmidt-Dannert, C., 2016. Encapsulation of multiple cargo proteins within recombinant Eut nanocompartments. Appl. Microbiol. Biotechnol. 100, 9187–9200. https //doi.org/10.1007/s00253-016-7737-8

Rae, B.D., Long, B.M., Badger, M.R., Price, G.D., 2013. Functions, compositions, and evolution of the two types of carboxysomes polyhedral microcompartments that facilitate CO2 fixation in cyanobacteria and some proteobacteria. Microbiol. Mol. Biol. Rev. MMBR 77, 357–379. https //doi.org/10.1128/MMBR.00061-12

Rae, B.D., Long, B.M., Badger, M.R., Price, G.D., 2012. Structural determinants of the outer shell of β-carboxysomes in Synechococcus elongatus PCC 7942 roles for CcmK2, K3-K4, CcmO, and CcmL. PloS One 7, e43871. https //doi.org/10.1371/journal.pone.0043871

Rae, B.D., Long, B.M., Whitehead, L.F., Förster, B., Badger, M.R., Price, G.D., 2013b. Cyanobacterial carboxysomes microcompartments that facilitate CO2 fixation. J. Mol. Microbiol. Biotechnol. 23, 300–307. https //doi.org/10.1159/000351342

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., Stanier, R.Y., 1979. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. Microbiology 111, 1–61. https //doi.org/10.1099/00221287-111-1-1

Roberts, E.W., Cai, F., Kerfeld, C.A., Cannon, G.C., Heinhorst, S., 2012. Isolation and Characterization of the Prochlorococcus Carboxysome Reveal the Presence of the Novel Shell Protein CsoS1D. J. Bacteriol. 194, 787–795. https //doi.org/10.1128/JB.06444-11

Roos, W.H., 2011. How to perform a nanoindentation experiment on a virus. Methods Mol. Biol. Clifton NJ 783, 251–264. https //doi.org/10.1007/978-1-61779-282-3_14

Roos, W.H., Bruinsma, R., Wuite, G.J.L., 2010. Physical virology. Nat. Phys. 6, 733–743. https //doi.org/10.1038/nphys1797

Roos, W.H., Gertsman, I., May, E.R., Brooks, C.L., Johnson, J.E., Wuite, G.J.L., 2012. Mechanics of bacteriophage maturation. Proc. Natl. Acad. Sci. U. S. A. 109, 2342–2347. https //doi.org/10.1073/pnas.1109590109

Roos, W.H., Radtke, K., Kniesmeijer, E., Geertsema, H., Sodeik, B., Wuite, G.J.L., 2009. Scaffold expulsion and genome packaging trigger stabilization of herpes simplex virus capsids. Proc. Natl. Acad. Sci. U. S. A. 106, 9673–9678. https //doi.org/10.1073/pnas.0901514106

Sae-Ueng, U., Li, D., Zuo, X., Huffman, J.B., Homa, F.L., Rau, D., Evilevitch, A., 2014. Solid-to-fluid DNA transition inside HSV-1 capsid close to the temperature of infection. Nat. Chem. Biol. 10, 861–867. https //doi.org/10.1038/nchembio.1628

Samborska, B., Kimber, M.S., 2012. A Dodecameric CcmK2 Structure Suggests β-Carboxysomal Shell Facets Have a Double-Layered Organisation. Structure 20, 1353–1362. https //doi.org/10.1016/j.str.2012.05.013

Sargent, F., Davidson, F.A., Kelly, C.L., Binny, R., Christodoulides, N., Gibson, D., Johansson, E., Kozyrska, K., Lado, L.L., MacCallum, J., Montague, R., Ortmann, B., Owen, R., Coulthurst, S.J., Dupuy, L., Prescott, A.R., Palmer, T., 2013. A synthetic system for expression of components of a bacterial microcompartment. Microbiology 159, 2427–2436. https //doi.org/10.1099/mic.0.069922-0

Sarnaik, A., Pandit, R., Lali, A., 2017. Growth engineering of Synechococcus elongatus PCC 7942 for mixotrophy under natural light conditions for improved feedstock production. Biotechnol. Prog. 33, 1182–1192. https //doi.org/10.1002/btpr.2490

Saschenbrecker, S., Bracher, A., Rao, K.V., Rao, B.V., Hartl, F.U., Hayer-Hartl, M., 2007. Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. Cell 129, 1189–1200. https //doi.org/10.1016/j.cell.2007.04.025

Savage, D.F., Afonso, B., Chen, A.H., Silver, P.A., 2010. Spatially ordered dynamics of the bacterial carbon fixation machinery. Science 327, 1258–1261. https //doi.org/10.1126/science.1186090

Schillers, H., Rianna, C., Schäpe, J., Luque, T., Doschke, H., Wälte, M., Uriarte, J.J., Campillo, N., Michanetzis, G.P.A., Bobrowska, J., Dumitru, A., Herruzo, E.T., Bovio, S., Parot, P., Galluzzi, M., Podestà, A., Puricelli, L., Scheuring, S., Missirlis, Y., Garcia, R., Odorico, M., Teulon, J.-M., Lafont, F., Lekka, M., Rico, F., Rigato, A., Pellequer, J.-L., Oberleithner, H., Navajas, D., Radmacher, M., 2017. Standardized Nanomechanical Atomic Force Microscopy Procedure (SNAP) for Measuring Soft and Biological Samples. Sci. Rep. 7, 5117. https //doi.org/10.1038/s41598-017-05383-0

Schmid, M.F., Paredes, A.M., Khant, H.A., Soyer, F., Aldrich, H.C., Chiu, W., Shively, J.M., 2006. Structure of Halothiobacillus neapolitanus Carboxysomes by Cryo-Electron Tomography. J. Mol. Biol. 364, 526–535. https //doi.org/10.1016/j.jmb.2006.09.024

Shih, P.M., Occhialini, A., Cameron, J.C., Andralojc, P.J., Parry, M.A.J., Kerfeld, C.A., 2016. Biochemical characterization of predicted Precambrian RuBisCO. Nat. Commun. 7. https //doi.org/10.1038/ncomms10382

Shively, J.M., Ball, F., Brown, D.H., Saunders, R.E., 1973a. Functional organelles in prokaryotes polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus. Science 182, 584–586.

Shively, J.M., Ball, F.L., Kline, B.W., 1973b. Electron Microscopy of the Carboxysomes (Polyhedral Bodies) of Thiobacillus neapolitanus. J. Bacteriol. 116, 1405–1411.

Silva, J.C., Gorenstein, M.V., Li, G.-Z., Vissers, J.P.C., Geromanos, S.J., 2006. Absolute quantification of proteins by LCMSE a virtue of parallel MS acquisition. Mol. Cell. Proteomics MCP 5, 144–156. https //doi.org/10.1074/mcp.M500230-MCP200

Sinha, S., Cheng, S., Sung, Y.W., McNamara, D.E., Sawaya, M.R., Yeates, T.O., Bobik, T.A., 2014. Alanine scanning mutagenesis identifies an asparagine-arginine-lysine triad essential to assembly of the shell of the Pdu microcompartment. J. Mol. Biol. 426, 2328–2345. https //doi.org/10.1016/j.jmb.2014.04.012

Sirohi, D., Chen, Z., Sun, L., Klose, T., Pierson, T.C., Rossmann, M.G., Kuhn, R.J., 2016. The 3.8 Å resolution cryo-EM structure of Zika virus. Science aaf5316. https //doi.org/10.1126/science.aaf5316

Slattery, R.A., Ort, D.R., 2015. Photosynthetic Energy Conversion Efficiency Setting a Baseline for Gauging Future Improvements in Important Food and Biofuel Crops1. Plant Physiol. 168, 383–392. https //doi.org/10.1104/pp.15.00066

Snijder, J., Kononova, O., Barbu, I.M., Uetrecht, C., Rurup, W.F., Burnley, R.J., Koay, M.S.T., Cornelissen, J.J.L.M., Roos, W.H., Barsegov, V., Wuite, G.J.L., Heck, A.J.R., 2016. Assembly and Mechanical Properties of the Cargo-Free and Cargo-Loaded Bacterial Nanocompartment Encapsulin. Biomacromolecules. https //doi.org/10.1021/acs.biomac.6b00469

Snijder, J., Reddy, V.S., May, E.R., Roos, W.H., Nemerow, G.R., Wuite, G.J.L., 2013. Integrin and Defensin Modulate the Mechanical Properties of Adenovirus. J. Virol. 87, 2756–2766. https //doi.org/10.1128/JVI.02516-12

Sommer, M., Cai, F., Melnicki, M., Kerfeld, C.A., 2017. β-Carboxysome bioinformatics identification and evolution of new bacterial microcompartment protein gene classes and core locus constraints. J. Exp. Bot. 68, 3841–3855. https //doi.org/10.1093/jxb/erx115

Spreitzer, R.J., 2003. Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. Arch. Biochem. Biophys. 414, 141–149.

Sun, Y., Casella, S., Fang, Y., Huang, F., Faulkner, M., Barrett, S., Liu, L.-N., 2016. Light Modulates the Biosynthesis and Organisation of Cyanobacterial Carbon Fixation Machinery through Photosynthetic Electron Flow. Plant Physiol. 171, 530–541. https //doi.org/10.1104/pp.16.00107

Sutter, M., Faulkner, M., Aussignargues, C., Paasch, B.C., Barrett, S., Kerfeld, C.A., Liu, L.-N., 2016. Visualization of Bacterial Microcompartment Facet Assembly Using High-Speed Atomic Force Microscopy. Nano Lett. 16, 1590–1595. https //doi.org/10.1021/acs.nanolett.5b04259

Sutter, M., Greber, B., Aussignargues, C., Kerfeld, C.A., 2017. Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell. Science 356, 1293–1297. https //doi.org/10.1126/science.aan3289

Sutter, M., Roberts, E.W., Gonzalez, R.C., Bates, C., Dawoud, S., Landry, K., Cannon, G.C., Heinhorst, S., Kerfeld, C.A., 2015. Structural Characterization of a Newly Identified Component of α-Carboxysomes The AAA+ Domain Protein CsoCbbQ. Sci. Rep. 5, 16243. https //doi.org/10.1038/srep16243

Sutter, M., Wilson, S.C., Deutsch, S., Kerfeld, C.A., 2013. Two new high-resolution crystal structures of carboxysome pentamer proteins reveal high structural conservation of CcmL orthologs among distantly related cyanobacterial species. Photosynth. Res. 118, 9–16. https //doi.org/10.1007/s11120-013-9909-z

Tabita, F.R., Satagopan, S., Hanson, T.E., Kreel, N.E., Scott, S.S., 2008. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. J. Exp. Bot. 59, 1515–1524. https //doi.org/10.1093/jxb/erm361

Tanaka, S., Kerfeld, C.A., Sawaya, M.R., Cai, F., Heinhorst, S., Cannon, G.C., Yeates, T.O., 2008. Atomic-level models of the bacterial carboxysome shell. Science 319, 1083–1086. https //doi.org/10.1126/science.1151458

Tanaka, S., Sawaya, M.R., Phillips, M., Yeates, T.O., 2009. Insights from multiple structures of the shell proteins from the beta-carboxysome. Protein Sci. Publ. Protein Soc. 18, 108–120. https //doi.org/10.1002/pro.14

Tanaka, S., Sawaya, M.R., Yeates, T.O., 2010. Structure and mechanisms of a protein-based organelle in Escherichia coli. Science 327, 81–84. https //doi.org/10.1126/science.1179513

Taylor, T.C., Backlund, A., Bjorhall, K., Spreitzer, R.J., Andersson, I., 2001. First Crystal Structure of Rubisco from a Green Alga,Chlamydomonas reinhardtii. J. Biol. Chem. 276, 48159–48164. https //doi.org/10.1074/jbc.M107765200

Thompson, M.C., Yeates, T.O., 2014. A challenging interpretation of a hexagonally layered protein structure. Acta Crystallogr. D Biol. Crystallogr. 70, 203–208. https //doi.org/10.1107/S139900471302422X

Tsai, Y., Sawaya, M.R., Cannon, G.C., Cai, F., Williams, E.B., Heinhorst, S., Kerfeld, C.A., Yeates, T.O., 2007. Structural Analysis of CsoS1A and the Protein Shell of the Halothiobacillus neapolitanus Carboxysome. PLOS Biol. 5, e144. https //doi.org/10.1371/journal.pbio.0050144

Tsai, Y., Sawaya, M.R., Yeates, T.O., 2009. Analysis of lattice-translocation disorder in the layered hexagonal structure of carboxysome shell protein CsoS1C. Acta Crystallogr. D Biol. Crystallogr. 65, 980–988. https //doi.org/10.1107/S0907444909025153

Uchihashi, T., Watanabe, H., Fukuda, S., Shibata, M., Ando, T., 2016. Functional extension of high-speed AFM for wider biological applications. Ultramicroscopy 160, 182–196. https //doi.org/10.1016/j.ultramic.2015.10.017

Uetrecht, C., Versluis, C., Watts, N.R., Roos, W.H., Wuite, G.J.L., Wingfield, P.T., Steven, A.C., Heck, A.J.R., 2008. High-resolution mass spectrometry of viral assemblies Molecular composition and stability of dimorphic hepatitis B virus capsids. Proc. Natl. Acad. Sci. U. S. A. 105, 9216–9220. https //doi.org/10.1073/pnas.0800406105

Vahabi, S., Nazemi Salman, B., Javanmard, A., 2013. Atomic Force Microscopy Application in Biological Research A Review Study. Iran. J. Med. Sci. 38, 76–83.

Vinckier, A., Semenza, G., 1998. Measuring elasticity of biological materials by atomic force microscopy. FEBS Lett. 430, 12–16. https //doi.org/10.1016/S0014-5793(98)00592-4

Whitehead, L., Long, B.M., Price, G.D., Badger, M.R., 2014. Comparing the in vivo function of α-carboxysomes and β-carboxysomes in two model cyanobacteria. Plant Physiol. 165, 398–411. https //doi.org/10.1104/pp.114.237941

Wildman, S.G., Bonner, J., 1947. The proteins of green leaves; isolation, enzymatic properties and auxin content of spinach cytoplasmic proteins. Arch. Biochem. 14, 381–413.

Yamori, W., Suzuki, K., Noguchi, K., Nakai, M., Terashima, I., 2006. Effects of Rubisco kinetics and Rubisco activation state on the temperature dependence of the photosynthetic rate in spinach leaves from contrasting growth temperatures. Plant Cell Environ. 29, 1659–1670. https //doi.org/10.1111/j.1365-3040.2006.01550.x

Yang, C.C., Chang, K.W., Liu, C.H., Shen, C.R., Yao, D.J., 2015. Using gradient micro-fluidics chips to optimize BG-11 medium for the growth of cyanobacteria Synechococcus elongatus PCC7942, in 10th IEEE International Conference on Nano/Micro Engineered and Molecular Systems. Presented at the 10th IEEE International Conference on Nano/Micro Engineered and Molecular Systems, pp. 136–140. https //doi.org/10.1109/NEMS.2015.7147393

Yeates, T.O., Crowley, C.S., Tanaka, S., 2010. Bacterial Microcompartment Organelles Protein Shell Structure and Evolution. Annu. Rev. Biophys. 39, 185–205. https //doi.org/10.1146/annurev.biophys.093008.131418

Yeates, T.O., Thompson, M.C., Bobik, T.A., 2011. The protein shells of bacterial microcompartment organelles. Curr. Opin. Struct. Biol. 21, 223–231. https //doi.org/10.1016/j.sbi.2011.01.006

Yeates, T.O., Tsai, Y., Tanaka, S., Sawaya, M.R., Kerfeld, C.A., 2007. Self-assembly in the carboxysome a viral capsid-like protein shell in bacterial cells. Biochem. Soc. Trans. 35, 508–511. https //doi.org/10.1042/BST0350508

Young, E.J., Burton, R., Mahalik, J.P., Sumpter, B.G., Fuentes-Cabrera, M., Kerfeld, C.A., Ducat, D.C., 2017. Engineering the Bacterial Microcompartment Domain for Molecular Scaffolding Applications. Front. Microbiol. 8, 1441. https //doi.org/10.3389/fmicb.2017.01441

Yung, M.C., Bourguet, F.A., Carpenter, T.S., Coleman, M.A., 2017. Re-directing bacterial microcompartment systems to enhance recombinant expression of lysis protein E from bacteriophage φX174 in Escherichia coli. Microb. Cell Factories 16. https //doi.org/10.1186/s12934-017-0685-x

Zhu, X.-G., Long, S.P., Ort, D.R., 2010. Improving photosynthetic efficiency for greater yield. Annu. Rev. Plant Biol. 61, 235–261. https //doi.org/10.1146/annurev-arplant-042809-112206

Appendix A Full list of primers and vectors used for the work described in this thesis

Primers

HOCH_5815 – chapter 3

5'- TGACACCATGGTATGGCTGA -3' – HOCH_5815 restriction cloning forward

5'- GGACACTCGAGCTACGCGCT -3' – HOCH_5815 restriction cloning reverse

5'- GGCCGCCGCCGTCGAACTGATCGGCTACGAGAAGA -3' - K28A INFUSION HOCH_5815 forward

5'- GTTCGACGCCGGCGGCCTTGACCATGGCGTCGGCG -3' - K28A INFUSION HOCH_5815 reverse

5'- CATCCCGGCCCCGCACGTCAACGTCGACGCGGCGC -3' – R78A INFUSION HOCH_5815 forward

5' – CGTGCGGGCCCGGGATGACGTGCACGGCGACGACT -3' – R78A INFUSION HOCH_5815 reverse

CcmK4 eGFP – chapter 4

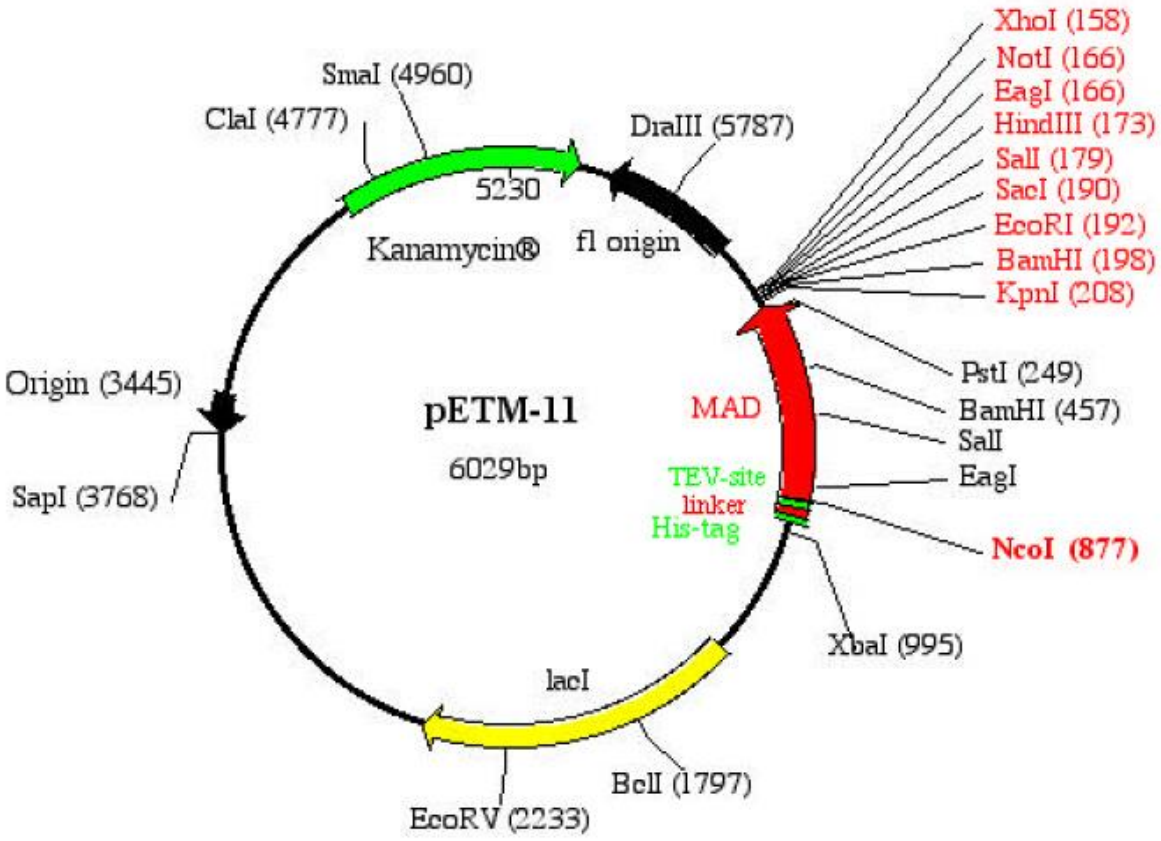
5'- TATCGACTCTCTGCCGAAGGAACTGGTAGTGGCCGCCGTCTGCCGGGCCCGGAGCTGCC -3' CcmK4 eGFP cloning forward

5'- CCCTCAGCCCAAATCAACCCTTTTAATCAGTCGCTACCCATTCCGGGGATCCGTCGACC-3' CcmK4 eGFP cloning reverse

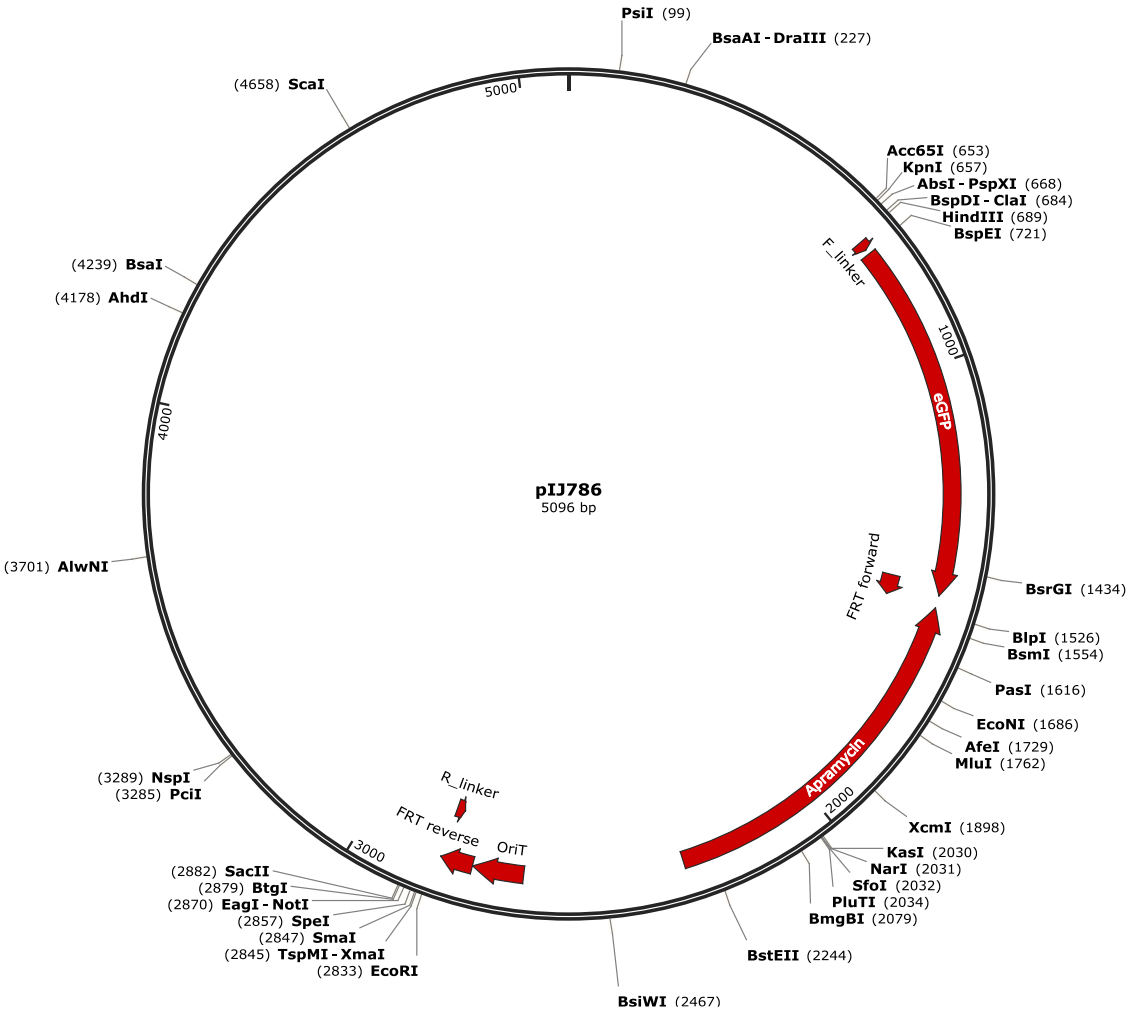
5' – GCCCAGAAGTGGAACAATA -3' CcmK4 segregation forward

5' - TGAGGTCCTTCCCTATCAAA -3' CcmK4 segregation reverse

Vectors



The map of pETM-11, as described by Dummler et al, used to express HOCH_5815 in this work by restriction cloning with Nco1 and Xho1 (Dümmler et al., 2005). This vector was subsequently transformed into *E. coli* by heat shock (Froger and Hall, 2007).

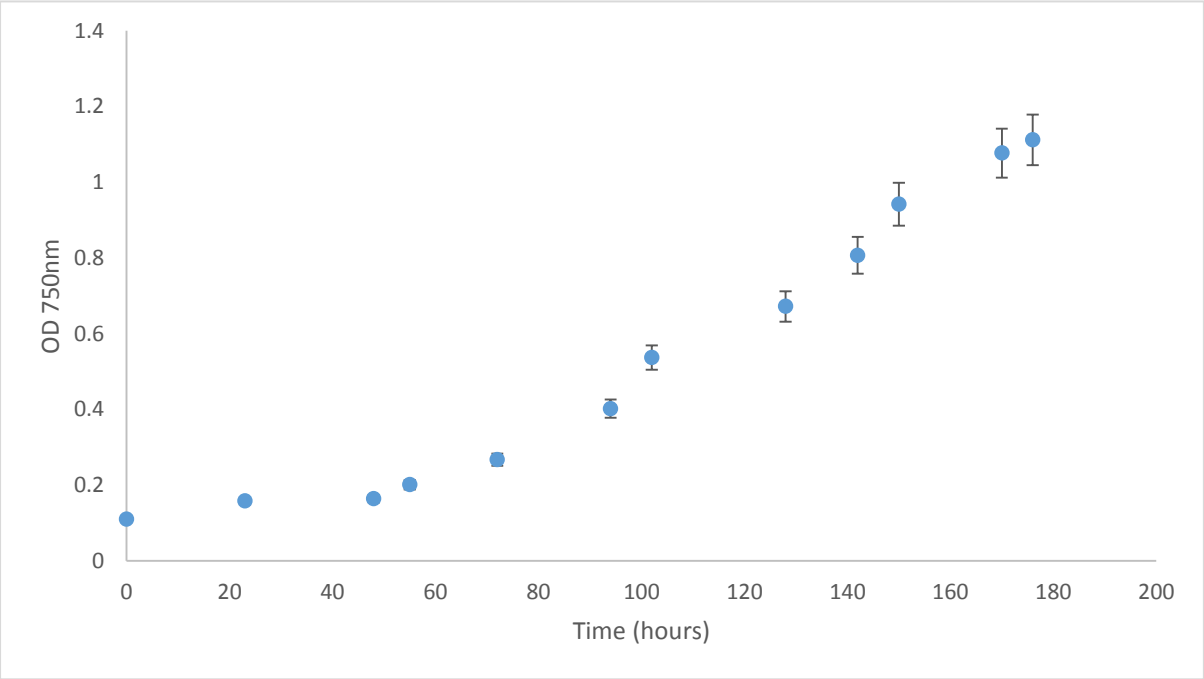


The pIJ786 vector map. Synpcc7942_0285 (*ccmK4*), was cloned using the pGEM-T Easy cloning system (Promega). An eGFP fusion was created by inserting the eGFP apramycin region amplified from the plasmid pIJ786 to the C terminus of *ccmK4* (Sun et al., 2016), using the Redirect strategy (Gust et al., 2004, 2002). The plasmid was verified by PCR and sequencing. The plasmids were then transformed into *Synechococcus* cells following the method described earlier (Golden, 1988).

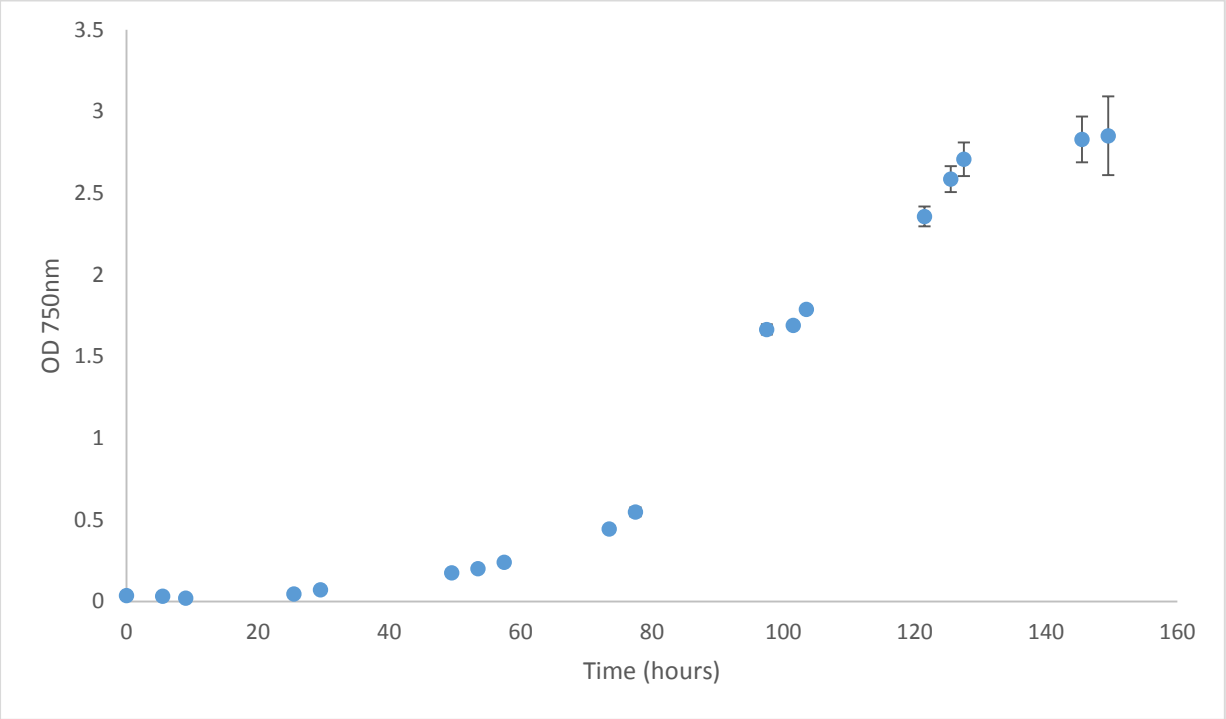
Appendix B Growth curves for the bacterial culture conditions described in this work

For all these growth curves the data represent mean of n=3 biological repeats and error bars represent 1 standard deviation.

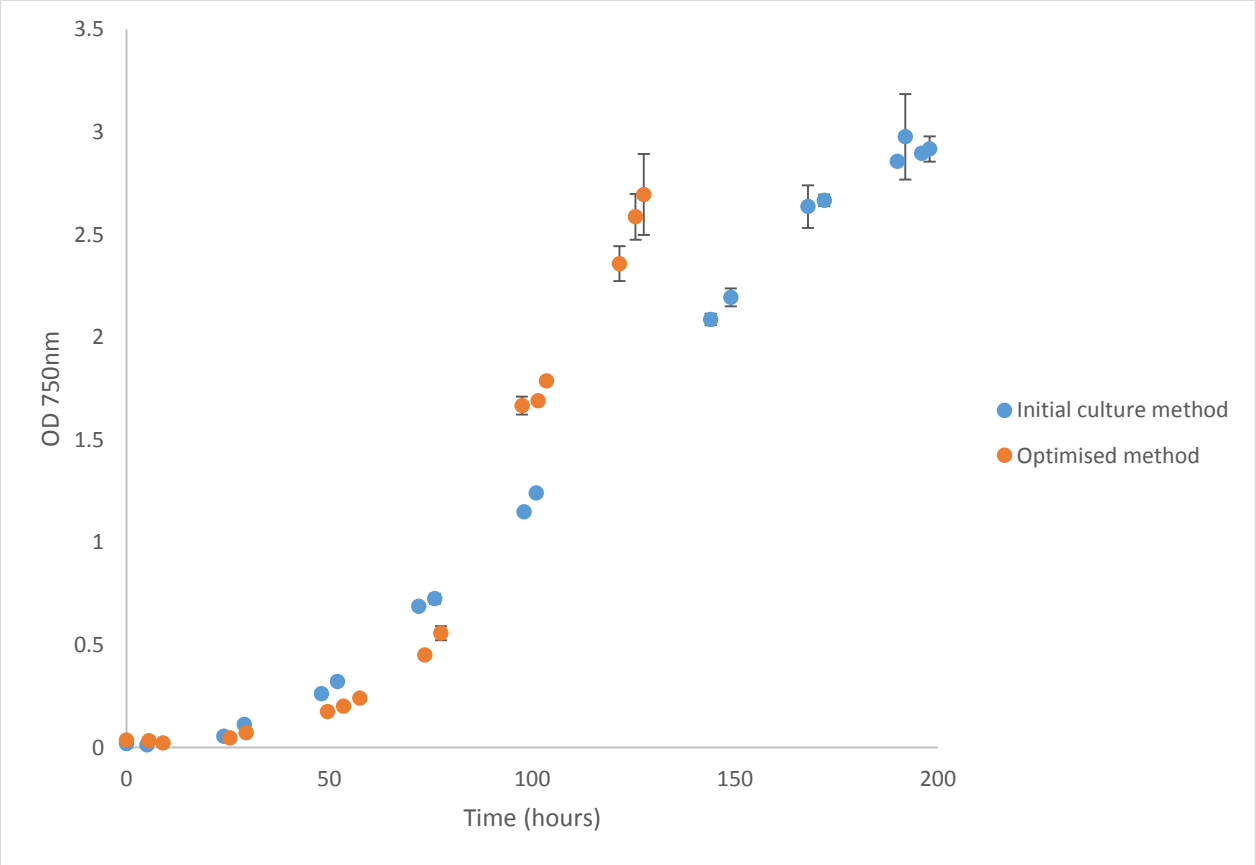
Syn 7942 standard cultivar batch culture



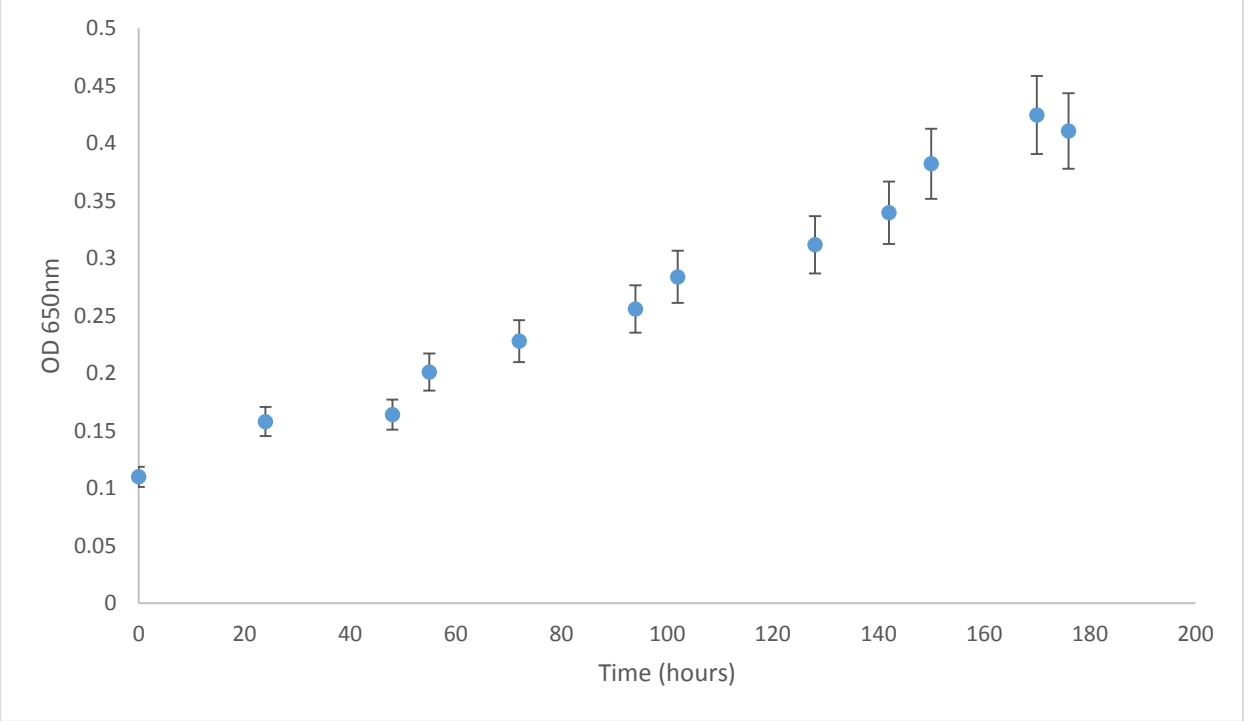
Syn 7942 5 L conical flask



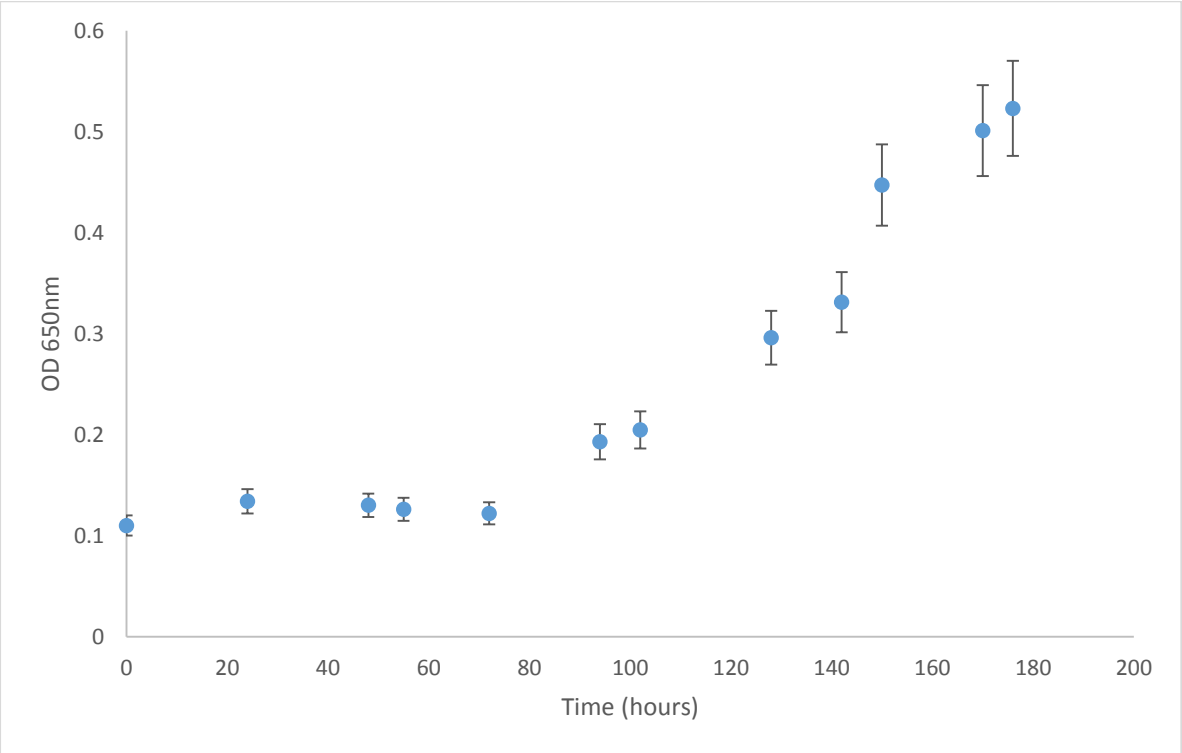
Syn 7942 photobioreactor



H. neap standard cultivar batch culture



H. neap bioreactor culture (before chemostatic dilution)



Co-author contributions made during my PhD and other data not directly relevant to the main body of this thesis

In order to disclose all of the work conducted during my PhD in full detail, I have included the following appendix. This appendix describes the work already published in co-authored manuscripts and some additional work not included in the main body of the text.

Co-authored publications

Yaqi Sun, Selene Casella, Yi Fang, Fang Huang, **Matthew Faulkner**, Steve Barrett, Lu-Ning Liu. *"Light Modulates the Biosynthesis and Organization of Cyanobacterial Carbon Fixation Machinery through Photosynthetic Electron Flow."* Plant Physiol. May 2016.171(1):530-41.

I contributed the transformation, genetic confirmation and culture maintenance of the CcmK4 eGFP *Syn7942* strain used in this study. I was also responsible for collecting part of RuBisCO radiometric assay data to ascertain the carbon fixation efficiency changes, and helped to develop the procedure for accurately calculating the specific activity.

Lu-Ning Liu, **Matthew Faulkner**, Xuan Liu, Fang Huang, Alistair C Darby, Neil Hall. "*Revised Genome Sequence of the Purple Photosynthetic Bacterium Blastochloris viridis."* Genome Announc. Jan 2016. 21;4(1).

I cultivated the *Blastochloris viridis* cells and isolated the chromosomal DNA used to sequence the genome of this strain. I also conducted part of the genome annotation.

"Engineering and modulating functional cyanobacterial CO_2-fixing organelles"

Accepted 15/04/18. https://www.frontiersin.org/articles/10.3389/fpls.2018.00739/abstract

Front. Plant Sci. | doi: 10.3389/fpls.2018.00739

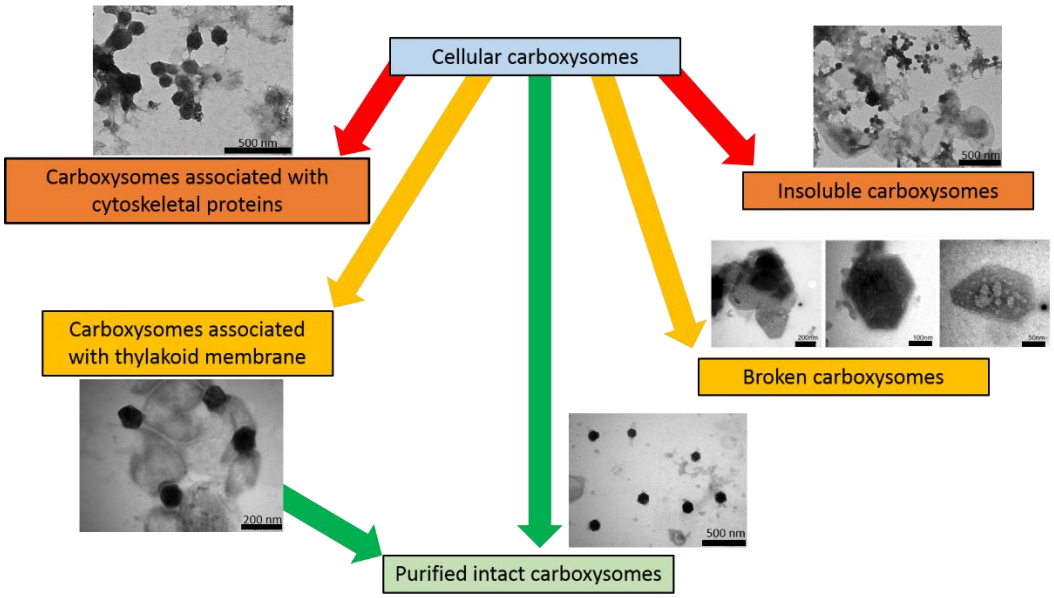
This article details the functional modulation of a cyanobacterial carboxysome heterologously expressed in *E. coli*. To which I have contributed AFM imaging, combined AFM and confocal imaging, image processing, RuBisCO enzyme kinetic measurements determined by radiometric assay and data analysis.

Additional data

Optimising the β-carboxysome purification

Purifying a high yield of intact CBs from *Syn7942* required a great deal of optimisation. This was an iterative process that took dozens of attempts over approximately 1 year. The most suitable time point during growth and the best growth conditions had to be determined first. After identifying the method of growth resulting in the most suitable compromise between the number of CB per cell, yield of cells, fewest interactions between CBs and cytoskeleton, fewest interactions between CBs and thylakoid membrane, before the onset stationary phase, elongation of cells and formation of storage polysaccharides; I had to optimise each step of the purification process.

The following flow diagram shows TEM images of the 40 % fractions from some of the other purifications I attempted during the optimisation process:



During the optimisation process I observed that using stationary phase, or cultures in which cells have started to elongate, results in "insoluble" CBs which sediment during purification. These CBs are associated with a large amount of material which appears in TEM images to be cytoskeletal protein, thylakoid membrane and possibly storage polysaccharides. Using cultures before mid-exponential phase resulted in CBs associated with appears to be cytoskeletal protein by TEM. Using cultures in mid to late exponential phase grown for 5 days under my optimised conditions resulted in CBs associated with what appeared to be thylakoid membrane. Given that the membrane can be emulsified by detergent treatment weak enough to leave the protein intact I was able to purify the intact β-CB from these samples.

The cell lysis also required a significant amount of optimisation. Lysis had to be strong enough that the majority of the cells were sufficiently broken and thus released their CBs into solution, but not so strong as to break the majority of the CBs. I tested chemical lysis and mechanical lysis by French press, glass beads and sonication. Sonication was found to be most effective at breaking Syn7942 but it also resulted in a high proportion of broken CBs. In order to reduce the duration and amplitude of sonication required to lyse >90 % of the Syn7942 I included a mild chemical pre-treatment. In doing so I was able to still able to release most of the CBs into solution and reduced the number of them broken by the sonication.