Optical Control over Monomeric and Multimeric Protein Hybrids

Optical Control over Monomeric and Multimeric Protein Hybrids

Rindia Maharani Putri

Rindia Maharani Putri

2017

OPTICAL CONTROL OVER MONOMERIC AND MULTIMERIC PROTEIN HYBRIDS

Rindia Maharani Putri

Members of the committee:

Chairman:	Prof. dr. H.H.J. ten Kate	(University of Twente)
Promotors:	Prof. dr. J.J.L.M. Cornelissen	(University of Twente)
	Prof. dr. N.H. Katsonis	(University of Twente)
Members:	Prof. dr. G.A. Woolley	(University of Toronto)
	Prof. dr. J.R. Castón	(CSIC Madrid)
	Dr. M.S.T. Koay	(Sanofi-Aventis Deutschland GmbH)
	Prof. dr. M.M.A.E. Claessens	(University of Twente)
	Prof. dr. J. Huskens	(University of Twente)

The research described in this thesis was performed within the laboratories of the Biomolecular Nanotechnology (BNT) group, the MESA+ Institute for Nanotechnology, and the Department of Science and Technology (TNW) of the University of Twente. This research was supported by the Netherlands Organization for Scientific Research (NWO), the European Research Council (ERC) and Indonesia Endowment Fund for Education (LPDP).

UNIVERSITY OF TWENTE.

Optical control over monomeric and multimeric protein hybrids

Copyright © 2017, Rindia Maharani Putri, Enschede, The Netherlands.

All rights reserved. No part of this thesis may be reproduced or transmitted in any form, by any means, electronic or mechanical without prior written permission of the author.

ISBN: 978-90-365-4365-1

DOI: 10.3990/1.9789036543651

Cover art: Pichamon Graphic House and Rindia Maharani Putri
(also featured in the inside back cover of ChemPhysChem 12/2016)

Printed by: Gildeprint – The Netherlands

OPTICAL CONTROL OVER MONOMERIC AND MULTIMERIC PROTEIN HYBRIDS

DISSERTATION

to obtain

the degree of doctor at the University of Twente,

on the authority of rector magnificus

Prof. dr. T. T. M. Palstra,

on account of the decision of the graduation committee,

to be publicly defended

on Friday September 8, 2017 at 12.45 h

by

Rindia Maharani Putri

Born on May 28, 1991

in Padang, Indonesia

This dissertation has been approved by:

Promoters: Prof. dr. J.J.L.M. Cornelissen

Prof. dr. N.H. Katsonis

To my family

Table of contents

Chapter 1: General introduction	1
1.1 Introduction	1
1.2 Aim and outline	2
1.3 References	3
Chapter 2: Strategies to modulate allosteric regulation and self-assembly of proteins using light	5
2.1 Introduction	6
2.2 Photo-activated allosteric regulation	8
2.2.1 Inclusion of photo-active protein domains	8
2.2.2 Inclusion of synthetic photo-switches	12
2.3 Photo-responsive protein assemblies of naturally occurring structures	14
2.3.1 Photo-activated dimerization	15
2.3.2 Photo-responsive cage-like and filament proteins	19
2.4 Photo-responsive artificial protein-based systems and higher-order structures	22
2.5 Perspectives	24
2.6 References	25
Chapter 3: Programming allostery in the human serum albumin with light	31
3.1 Introduction	32
3.2 Results and discussion	33
3.2.1 Design and characterization of a photo-switchable hybrid	33
3.2.2 Optical control over ligand binding to subdomain IB	37
3.2.3 Molecular dynamic simulations	41
3.3 Conclusions	44
3.4 Acknowledgements	44

3.5	Materials and methods	45
3.5.1	General	45
3.5.2	Synthesis of the photo-switch	45
3.5.3	Synthesis and characterization of the hybrid system	48
3.5.4	Photo-responsive ligand binding analyses	48
3.5.5	PFGSE-NMR experiments	50
3.5.6	Determination of effective quenching constants	50
3.5.7	Analyses of protein global structure	50
3.5.8	Data and schematic representation	51
3.6	References	51

Chapter 4: Bacterial encapsulins as *in vitro* and *in vivo* protein-based nanoplateforms **55**

4.1	Introduction	56
4.2	Results and discussion	57
4.2.1	Characterization of encapsulin stability and structure	57
4.2.2	<i>In vitro</i> study of encapsulin as nanoreactors on surface	65
4.2.3	<i>In vivo</i> study of encapsulin as an agent for cellular infection	68
4.3	Conclusions	70
4.4	Acknowledgements	70
4.5	Materials and methods	71
4.5.1	General	71
4.5.2	Cryo-EM and image processing	71
4.5.3	Characterization of encapsulin stability	72
4.5.4	Immobilization of encapsulin and catalytic assay	73
4.5.5	Cell experiments	74
4.6	References	74

Chapter 5: Labeling encapsulin with light-switchable fluorophores **77**

5.1	Introduction	78
5.2	Results and discussion	79

5.2.1	Covalent attachment of spiropyran to encapsulin	79
5.2.2	Photo-switchable fluorescence	82
5.2.3	Structural integrity of encapsulin particles	84
5.3	Conclusions	86
5.4	Acknowledgements	86
5.5	Materials and methods	86
5.5.1	General	86
5.5.2	Synthesis of spiropyran-succinimide	86
5.5.3	Covalent attachment of spiropyran and encapsulin	89
5.5.4	Photo-switchable fluorescence analysis	89
5.5.5	Characterization of encapsulin structure	90
5.6	References	90

Chapter 6: Light-fueled assembly of encapsulin and chaperone into hybrid superstructures **93**

6.1	Introduction	94
6.2	Results and discussion	95
6.2.1	Light-fueled assembly of hybrid superstructures	95
6.2.2	Structural changes of proteins upon assembly	101
6.2.3	Emergence of light-dependent order by inclusion of photoacid in the superstructures	104
6.3	Conclusions	107
6.4	Acknowledgements	107
6.5	Materials and methods	108
6.5.1	General	108
6.5.2	Reversible formation of hybrid superstructures	108
6.5.3	Characterization of hybrid superstructures	109
6.5.4	Characterization of structural changes of proteins	109
6.5.5	Analysis of photoacid inclusion in the superstructures	109
6.6	References	110

Summary **113**

Samenvatting **115**

Acknowledgements	117
About the author	121

Chapter 1

General introduction

1.1 Introduction

From naturally occurring to semi-synthetic systems, proteins dictate a plethora of essential life processes, including catalysis,^[1] structural support,^[2] cell signaling^[3] as well as molecular transport and delivery.^[4] Modulating the functioning of these proteins using an external trigger would thus offer an effective means to remote control a range of complex biological processes and networks of regulations. In this regard, light appears as an attractive trigger, as it is precise, compatible with a wide range of condensed phases, clean, relatively non-destructive towards proteins, and therefore it is envisioned that it can be adapted for *in vivo* and *in vitro* applications. Moreover, light enables reversible and selective spatiotemporal control, which in turn allows an on/off and dose-dependent regulation over complex molecular processes.

Recent years have witnessed an increasing attempt to modulate biological processes by photo-engineering some key proteins involved in cellular regulations. The photo-modulation of biological processes has been reported,^[5-9] including enzyme catalysis,^[10] protein folding,^[11] the opening of channel proteins^[12] as well as protein dimerization.^[13] Among biological regulations involving proteins, allostery and self-assembly of proteins are highly dynamic, tunable processes^[14, 15] that are essential to the functioning of cellular networks, yet remain not entirely understood and therefore they are challenging to modulate in a controlled manner. In this thesis, we address the challenge to understand the mechanisms of allostery and hierarchical self-assembly of proteins and gain control over such cooperative and dynamic actions to obtain a measurable output from individual switching events. We use pre-existing biological systems as cooperative molecular media to amplify the motion of molecular switches to generate collective behaviors of hybrid systems.

1.2 Aim and outline

We engineer an allosteric transport protein^[16] and a bacterial nanocompartment encapsulin^[17] into dynamic building blocks for light-responsive hybrid (semi-synthetic) systems. Naturally, neither of these protein systems is responsive to light. Upon chemical modification with artificial molecular photo-switches into their structures, their properties and dynamics become reversibly controllable using irradiation with light.

The strategy involves modifying the proteins with photo-responsive spiropyran-based switches,^[18] which can be grafted covalently on the proteins (Chapter 3 and 5) or dispersed in the medium that contains them (Chapter 6). In a reversible manner, the photo-switches change in conformation and polarity as a response to light irradiation, therefore affecting the proteins to which they are associated with. This strategy enables control over the dynamics of allostery in transport protein (Chapter 3) and the chemical and self-assembly properties of encapsulins (Chapter 5 and 6, respectively).

Chapter 2 describes recent strategies towards the modulation of protein functions using light, with a special focus on allostery and self-assembly.

Chapter 3 describes a light-programmed allostery in the transport protein Human Serum Albumin (HSA), that is engineered with a spiropyran photo-switch. The photo-switch is incorporated into a specific binding site of HSA, and once it switches from closed, non-charged spiropyran form to open, zwitter-ionic merocyanine form, we can demonstrate that it is a neighboring binding site that responds to the environmental changes, mediated by allostery.

In **Chapter 4**, the structural and functional basis of multimeric encapsulin from the bacteria *B. linens* is presented. This chapter describes structural characterization of the cage-like particles using cryo-electron microscopy (cryo-EM) reconstruction and analyses of particle stability in front of environmental changes (i.e., pH, ionic strength and addition of organic solvent). Furthermore, the functionality of encapsulin as nanoplateforms is demonstrated both *in vitro* and *in vivo*.

Chapter 5 presents the labeling of *B. linens* encapsulin with photo-switchable spiropyran-based fluorophores. Following structure- and function-based analyses in Chapter

4, the covalent attachment of multiple spiropyran molecules at encapsulin exterior is described in this chapter as well as the on/off photo-induced fluorescence in cycles.

Chapter 6 describes a light-fueled dynamic assembly of *T. maritima* encapsulin and *E. coli* chaperone protein into transient giant superstructures. The assembly is mediated by a spiropyran-based photoacid that plays two roles in the assembly event: 1) as a photo-active building block for the assembly and 2) to reversibly and gradually control the pH of the system to form ordered superstructures. We demonstrate that the resulting superstructures depend on continuous irradiation to sustain their assembled state.

1.3 References

- [1] S. J. Benkovic, S. Hammes-Schiffer, *Science* **2003**, *301*, 1196-1202.
- [2] A. Akhmanova, M. O. Steinmetz, *Nat. Rev. Mol. Cell Bio.* **2015**, *16*, 711-726.
- [3] C. J. Miller, L. A. Davidson, *Nat. Rev. Genet.* **2013**, *14*, 733-744.
- [4] A. H. Futerman, *Nature* **2007**, *449*, 35-37.
- [5] A. Gautier, C. Gauron, M. Volovitch, D. Bensimon, L. Jullien, S. Vriz, *Nat. Chem. Biol.* **2014**, *10*, 533-541.
- [6] W. Szymanski, J. M. Beierle, H. A. Kistemaker, W. A. Velema, B. L. Feringa, *Chem. Rev.* **2013**, *113*, 6114-6178.
- [7] K. E. Brechun, K. M. Arndt, G. A. Woolley, *Curr. Opin. Struct. Biol.* **2016**, *45*, 53-58.
- [8] G. Guglielmi, H. J. Falk, S. De Renzis, *Trends Cell Biol.* **2016**, *26*, 864-874.
- [9] M. Baker, *Nat. Methods* **2012**, *9*, 443-446.
- [10] C. W. Riggsbee, A. Deiters, *Trends Biotechnol.* **2010**, *28*, 468-475.
- [11] U. T. Bornscheuer, *Nature* **2016**, *540*, 345-346.
- [12] A. Kocer, M. Walko, W. Meijberg, B. L. Feringa, *Science* **2005**, *309*, 755-758.
- [13] C. L. Tucker, J. D. Vrana, M. J. Kennedy, *Curr. Protoc. Cell Biol.* **2014**, *64*, 17 16 11-20.
- [14] S. R. Tzeng, C. G. Kalodimos, *Nature* **2009**, *462*, 368-372.
- [15] J. J. McManus, P. Charbonneau, E. Zaccarelli, N. Asherie, *Curr. Opin. Colloid Interface Sci.* **2016**, *22*, 73-79.
- [16] X. M. He, D. C. Carter, *Nature* **1992**, *358*, 209-215.

- [17] M. Sutter, D. Boehringer, S. Gutmann, S. Guenther, D. Prangishvili, M. J. Loessner, K. O. Stetter, E. Weber-Ban, N. Ban, *Nat. Struct. Mol. Biol.* **2008**, *15*, 939-947.
- [18] R. Klajn, *Chem. Soc. Rev.* **2014**, *43*, 148-184.

Chapter 2

Strategies to modulate allosteric regulation and self-assembly of proteins using light

Living cells and life processes rely on networks of interactions and cellular dynamics that are orchestrated by protein assemblies. The early quest was to understand how the assemblies are spontaneously put together and controlled by Nature, leading to extensive research on protein assembly and regulation. Among the mechanisms for regulation exemplified by Nature, allostery and self-assembly of proteins are particularly prominent as they not only enable a dynamic and tunable control of naturally occurring systems, but also inspire the design and generation of artificial protein-based hybrids. Using light to stimulate and externally interfere with such mechanisms would open up the possibility to control complex biological processes as well as developing smart materials based on protein hybrids. Here, we provide a concise review of recent advances on the strategies developed and used to interfere with the allostery and self-assembly mechanisms in naturally occurring systems and artificially designed systems.

2.1 Introduction

Supramolecular structures in Nature that are formed by protein assemblies play crucial roles in cells, for instance as structural components and cellular machinery.^[1] The skeleton of the cell (cytoskeleton) is composed by self-assembled protein filaments that maintain the structure and the shape of the cell. These filaments also provide platforms for vesicles and organelles mobility inside the cell. Besides the structural support, major metabolic pathways in cells are carried out by dozens of protein assemblies, for instance, ribosome (the translation machinery) and spliceosome (the RNA-splicing machinery) are made up by multiple protein subunits interacting with each other building complex catalytic protein clusters.^[2-4] Consequently, the cell itself can be viewed as a crowded and confined factory built up by interlinking, highly-organized protein assemblies.^[5,6]

Nature regulates the activity of protein assemblies using various mechanisms, including allosteric and cooperative regulation,^[7] tunable interactions and assembly of protein subunits^[8] as well as signal amplification and feedback loops in cascade systems.^[9] Allostery and self-assembly of proteins are highly dynamic, tunable processes and modulating such mechanisms using external triggers would enable interfering with complex systems in cellular networks. Light, in particular, is an attractive means to control protein assemblies as it enables a precise spatiotemporal control and applicable in physiological conditions. The use of light to modulate protein functionality in general has attracted attention in recent years, leading to the success of photo-modulating various biological processes,^[10-15] such as enzyme catalysis,^[16] protein folding/unfolding,^[17] fluctuations in the activity of channel proteins^[18] as well as the dynamic process of protein dimerization.^[19]

Naturally occurring, non-photoresponsive proteins can be engineered into photo-switchable systems using chemical and/or genetic approaches (Figure 2.1). Chemical approaches are based on covalent incorporation of synthetic photo-switchable molecules in the structure of the proteins or in the ligands/substrates that bind the proteins. Azobenzene^[20,21] and spiropyran^[22] switches are the most common examples of photo-switches incorporated into protein or ligand structures. Upon UV light irradiation, azobenzene switches from its *trans* form to *cis* form, while spiropyran photo-isomerization involves a ring opening into merocyanine (Figure 2.1).^[21,22] The changes in the conformation

and dipole moment of the switches upon irradiation affect the biochemical properties of the proteins or ligands to which they are attached, enabling photo-control of the overall system. Moreover, the light-induced switching is a reversible process (either thermally or upon irradiation with visible light), thus allowing a reversible control over protein functionality.^[21,22] Another chemical approach commonly used is covalent incorporation of bulky photo-caging molecules (for instance, anchored to the ligands) that are cleaved once irradiated and hence result in photo-activation.^[23]

Figure 2.1 Chemical and genetic strategies to develop photo-switchable proteins. Chemical approach relies on incorporation of synthetic photo-responsive switches such as a) azobenzene and b) spiropyran, whereas genetic approach involves the insertion of c) naturally occurring photo-active domain or d) photo-responsive unnatural amino acid (UAA) via genetic engineering.

On the other hand, genetic approaches (Figure 2.1) involve genetic engineering that includes construction of chimeric proteins with an insertion of photo-active domains (such

as the light-oxygen-voltage-sensing protein domain, LOV)^[24] or point mutation to incorporate photo-responsive unnatural amino acids^[25] (photo-caged amino acids or azobenzene-modified amino acids). Here, we address both chemical/synthetic and genetic strategies to modulate the phenomena of allosteric regulation and the assembly of proteins in naturally occurring systems and artificial (designed) systems.

2.2 Photo-activated allosteric regulation

Nature assembles proteins into dimeric and multimeric forms to enhance cellular efficiency, for instance mutations that are introduced in the genetic level will be amplified in the assembled structure symmetrically.^[8] Multimeric proteins often allow regulation via cooperativity between protein subunits. The cooperative effect between subunits in a multimeric protein is commonly known as a form of allostery. In general, allosteric regulation enables communication between distinct sites in a protein and hence allows indirect control over one site by influencing another.^[7] As the phenomenon is central to life yet not well understood and largely unresolved, allostery is referred to as the ‘second secret of life’ (the first one is the genetic code).^[26] Using allostery, Nature gains effective control over various biological processes, ranging from molecular transport to enzyme catalysis.

To modulate allosteric regulation using light, photo-responsive switches can be introduced either in the vicinity of allosteric binding sites or included in the chemical structure of allosteric effectors (i.e., small molecules that bind to allosteric sites). In addition to modifying naturally occurring allosteric proteins, photo-responsive allosteric proteins can be generated from non-allosteric proteins by inserting photo-active protein domains (i.e., the LOV) in a pre-determined position in the structure.

2.2.1 Inclusion of photo-active protein domains

Naturally, light is capable of stimulating certain processes in living cells due to the presence of photo-active moieties embedded in the photo-active protein domains. A superfamily of protein domains across the kingdoms of life that plays a fundamental role as sensors for environmental stimuli in general is called the Per-ARNT-Sim (PAS) domains.^[24] PAS domains are responsible for signaling and adapting the cells to environmental changes caused by external stimulations such as light irradiation, redox potential, oxygen and energy

level. The mechanism of action of PAS domains usually involves cofactor binding and mediating interactions between proteins.^[24]

Figure 2.2 Modulation of allosteric regulation via photo-active protein domain. a) Structure of LOV domain displaying the blue light-responsive FMN chromophore (red spheres) and the flanking J α helix cap. Reprinted with permission from ref [28]. Copyright © 2003, American Association for the Advancement of Science. In plant phototropin, the unfolding of J α helix (blue cylinder) allows activation of the neighboring kinase domain (pink box) upon irradiation. The LOV is suggested to act either as a kinase inhibitor in the dark state or kinase activator in the lit state. Reprinted with permission from ref [31]. Copyright © 2004, American Chemical Society. b) Schematic representation of LOV in algal aureochrome showing the A' α helix that blocks the dimerization site of LOV (in grey). Upon irradiation, the A' α helix unfolds to allow the dimerization and promote the binding of double-stranded (ds) DNA to DNA-binding domain (highlighted in red). Reprinted with permission from ref [33]. Copyright © 2016, Oxford University Press.

An example of PAS domains is the blue light-responsive domain found in photoreceptor proteins in plants (i.e., the phototropins), which is referred to as the light-oxygen-voltage-sensing (LOV) domain.^[27] LOV domain is highly conserved and its response to light originates from the presence of a chromophore, flavin mononucleotide cofactor (FMN), in

the interior of LOV/PAS core (Figure 2.2a).^[28] The non-covalently bound FMN cofactor acts as a blue light absorbing moiety ($\lambda < 500$ nm) that reacts reversibly with a cysteine residue in vicinity (4.2 Å away), resulting in a formation of a covalent protein-flavin adduct.^[29,30] In plant phototropins, formation of the adduct is coupled allosterically with a neighboring catalytic domain with a kinase activity, altogether enabling the photo-activation of the kinase.^[28,31] Furthermore, the formation of the adduct is unfavorable energetically and hence is spontaneously reversible in the absence of blue light (from seconds to hours, varies from one system to another).^[29,32]

Structural analysis of a phototropin from oat (*Avena sativa*) reveals that the formation of protein-flavin adduct results in a destabilizing conformational change that selectively unfolds a helical segment outside of the PAS/LOV core called the J α helix cap^[28,31] (Figure 2.2a), which is suggested to play a role in the kinase activation. To confirm that the unfolding of J α helix segment is the cause of kinase activation, several point mutations are made along the structure of J α helix that induce the unfolding event to occur independently of photo-responsive LOV core domain.^[31] The investigation proves that mutation-induced unfolding of J α helix results in kinase activation without requiring any illumination, highlighting the role of J α helix in bridging the allostery between the LOV core and the neighboring kinase as illustrated in Figure 2.2a.^[31]

An allosteric regulation is also observed in the dynamic of blue light receptors of algae, called the aureochrome.^[33,34] In the aureochrome, LOV domain is coupled to a DNA-binding domain (called the basic region leucine zipper or bZIP domain) instead of a catalytic kinase domain. The structural arrangement of the LOV domain (the sensory module) and the bZIP domain (the effector) is inversed compared to plant phototropins. A linker region that contains a helical segment (called the A' α helix) connects the LOV core and the DNA-binding bZIP domain (Figure 2.2b). Structural analysis of aureochrome1a from alga *Phaeodactylum tricornutum* reveals the occurrence of photo-activated homodimerization of LOV domain that is regulated by the unfolding of the linking A' α helix.^[33] A' α helix is found to cover the dimerization site of LOV in the dark state and its unfolding upon light irradiation allows the dimerization to occur (Figure 2.2b). On the other end, the LOV core of aureochrome is connected to a flanking J α helix that displays similar behavior with J α helix from phototropins (i.e., unfolds upon blue light irradiation). Subsequent unfolding events of

$J\alpha$ helix and $A'\alpha$ helix are found to be allosterically coupled as the latter only occurs as the consequence of the former.^[34] Both helical segments have to unfold to allow the dimerization of the aureochrome to occur. The formation of the dimer state consequently increases the affinity of aureochrome for DNA binding via its bZIP domain, altogether enabling a photo-activated DNA binding event. The increase in DNA binding is due to conformational changes in the bZIP region as a result of the dimerization (Figure 2.2b).^[33]

The role of flavin-binding photo-active domains in modulating allosterically coupled effector domains is also observed in prokaryotes. The blue light using flavin (BLUF) domains in bacteria is responsible for the allosteric photo-activation, for instance in the modulation of adenylate cyclase from cyanobacterium *Oscillatoria acuminata*^[35] and phosphodiesterase from *Klebsiella pneumoniae*.^[36] Although prokaryotic BLUF domains typically bind FAD instead of FMN (as in LOV domains), thus resulting in different photochemistry, they are highly comparable in terms of photo-activated signal propagation through structural changes at the interface of connected domains.^[37] Furthermore, owing to the small size of both BLUF and LOV domains (around 100 – 140 amino acids), they can be genetically fused to non-photo-responsive proteins to design artificial systems that allow allosterically regulated photo-activation.

In a rationally designed chimeric protein, LOV domain from *A. sativa* is fused to dihydrofolate reductase (DHFR) from *E. coli* resulting in a photo-dependent reductase activity.^[38] The photo-activation of DHFR is highly dependent on to which site of DHFR the LOV is genetically inserted. To select a site with higher probability to result in the coupling of fused LOV and DHFR, a statistical coupling analysis (SCA) is performed. SCA quantitatively pinpoints the long-range interaction and conserved communication between distinct sites in a protein, revealing the possible influence of different surface sites to the catalytic site of DHFR. Two surface sites of DHFR are selected, which are the coupled site A (i.e., showing communication with the catalytic site) and the non-coupled site B. The LOV domain is inserted to selected sites of DHFR via its N- and C-terminal helical extensions. The results show that all chimeric proteins connected via non-coupled site B do not result in any light activation, while one of the chimeric proteins connected via coupled site A shows an increase of activity upon blue light illumination up to 2-fold. Although the increase is

modest, this study presents the proof-of-concept that photo-responsive allosteric proteins can be rationally designed and built by fusing a photo-active domain to a non-allosteric protein.

The success of fusion approach heavily relies on whether the photo-active domain is genetically inserted to the “right” region of the active protein that would result in maximum perturbation and hence a measurable impact. This is corroborated by a rational design of a chimeric protein of LOV and a viral K⁺ channel pore (Kcv).^[39] In this study, LOV is fused to various regions of Kcv that are important for channel gating in order to generate a light-gated K⁺ channel. One variant in which the LOV is fused to the N-terminal of Kcv displays a modest increase in conductance (ion flux) when irradiated with blue light ($\lambda = 455$ nm).

A similar concept with a different strategy is used in a rational design of a light-activated DNA binding allosteric protein.^[40] Instead of directly inserting the LOV domain to allosterically coupled sites, a rigid α -helical domain linker is employed as the “allosteric lever arm” to connect the LOV domain through its J α helix extension to truncated N-terminal of *trp* repressor from *E. coli*. As the two connected domains structurally share a helical arm in between, it is assumed that the photo-activated unfolding of J α helix of LOV will cause an energetically unfavorable bending of the connecting arm that will be relieved through a global shift in the conformation, affecting the functionality of the *trp* repressor. A series of constructs to insert LOV domain to successive truncations of N-terminal helix of *trp* repressor results in 12 protein fusions. Among the twelve, one fusion generated from connecting the C-terminal of J α helix to the middle of N-terminal helix of *trp* repressor shows a significant difference in DNA binding affinities upon illumination at $\lambda = 470$ nm (dark state is 5.6-fold higher than lit state).

2.2.2 Inclusion of synthetic photo-switches

Previous section elaborates how genetically fusing photo-active domains has been proven successful to transform non-allosteric proteins into light-modulated allosteric proteins. However, the fusion approach is not suitable in the cases where an on/off control is desired and might be rather cumbersome to execute. A complementary approach is to incorporate smaller, synthetic photo-switches into the protein structure or in the structure of the allosteric effector. A reversible attachment of an agonist (i.e., an activating effector) to the allosteric site of a protein via an azobenzene linker is demonstrated in the case of allosteric ionotropic

glutamate receptor (iGluR), an ion-channel involved in the central nervous system (Figure 2.3a).^[41] Upon irradiation with UV light ($\lambda = 380$ nm), the azobenzene switches from *trans* to *cis* conformation, eventually allowing the agonist to reach the allosteric site, whereas the spatial arrangement of *trans* conformation does not allow such event to occur (Figure 2.3a). As the photo-isomerization of azobenzene is reversible (by irradiation at $\lambda = 500$ nm), the channel gating measured by whole cell patch clamping can be turned on and off upon illumination at different wavelengths. In addition to chemical attachment of azobenzene moiety, genetic-code expansion can also be used to incorporate photo-active unnatural amino acid (UAA) in the protein structure, for example the photo-cross-linker azido-phenylalanine (Figure 2.3b).^[42]

Figure 2.3 Modulation of allosteric regulation using synthetic photo-switches. a) Allosterically regulated opening and closing of an ion-channel with an agonist (orange) attached to the protein via an azobenzene linker (red and black). Reprinted with permission from ref [41]. Copyright © 2006, Macmillan Publishers Limited, part of Springer Nature. b) Incorporation of photo-active unnatural amino acid (azido-phenylalanine) into a protein structure. Reprinted with permission from ref [42]. Copyright © 2014, National Academy of Sciences. c) Incorporation of azobenzene moiety in an allosteric effector, resulting in either *trans* and *cis*-alloswitch upon irradiation at different wavelengths. Reprinted with permission from ref [43]. Copyright © 2014, Macmillan Publishers Limited, part of Springer Nature.

Furthermore, azobenzene switches can be incorporated into the chemical structure of the allosteric effectors (drugs), resulting in an on/off photo-modulation of natural allosteric proteins. In photopharmacology, this strategy is favored to avoid side effects as allosteric drugs usually display higher selectivity in binding to receptors. The allosteric effectors are modified with azobenzene in such a way so that one photo-isomer significantly favors binding (on/active state) compared to the other (off/inactive state), owing to steric difference between the isomers. The approach has been demonstrated to modulate G-protein-coupled receptors (GPCRs) using light,^[43] a class of transmembrane proteins responsible for signal transduction and cellular response to stimuli in eukaryotic cells. Effectors that display structural homology to Ar–N=N–Ar moiety of azobenzene are ideal target for modification (Figure 2.3c). The N=N group specifically can be introduced, for instance, to replace the amide bonds in the effectors. Remarkably, this strategy has been demonstrated to control the motility of living cells of *X. tropicalis* tadpoles as a response to light stimulation.^[43]

2.3 Photo-responsive protein assemblies of naturally occurring structures

Self-assembly of protein subunits into dimers and multimers is crucial for their functioning. A plethora of proteins that are responsible for essential, yet complex life processes such as protein-synthesizing ribosomes and microtubules are active and functional in their assembled forms.^[2] Consequently, manipulating the self-assembly phenomenon using external triggers would enable interfering with vital biological processes embedded in the interlinking cellular networks. Comparable to modulating allosteric regulation with light, self-assembly of proteins can be activated/deactivated upon illumination by introducing photo-responsive moieties either in the structure of the protein subunits or the regulatory molecules that chemically affect the process (e.g., inhibitors and activators of the assembly event). Likewise, the rational design and the position to which the switches are chemically or genetically introduced play a vital role in determining whether a strategy could result in the desired effects. Modification with photo-switches has to introduce a significant change in the structural dynamics of proteins to overcome the native behavior of assembly, whereas

if the structural modification leads to excessive structural perturbations, the protein subunits might lose their ability to assemble altogether.

2.3.1 Photo-activated dimerization

The simplest example of protein-based assemblies is the homodimers. Inside the cells, homodimers (and further, homooligomers) are formed to reach structures with higher stability as well as to minimize the genome size needed to synthesize functional proteins. The formation of a dimer is more preferred instead of a multi-subunit single-chain protein due to the ability to dissociate under certain conditions, which could be crucial in regard to biological functions.^[44] Although individually the intermolecular forces that support the dimeric structure can be classified as relatively weak interactions, the collaborative networks of interactions forming the dimers are among the strongest interactions found in Nature. These collaborative networks mainly consist of non-covalent interactions, while stronger individual interactions (i.e., covalent bonds as in disulfide bridges) are only present in a small number at the dimeric interfaces.^[45]

Despite the collectively strong forces that support protein dimers, interfering with dimerization events using light can be achieved with the help of a class of activating molecules called the dimerizers. Naturally, some of dimerization events require the presence of natural products such as rapamycin and abscisic acid; such processes are referred to as the chemically induced dimerization (CID).^[46] Induced dimerization is employed by Nature to control and direct the localization of proteins to specific cellular compartments and the activation of signaling pathways. Optical control over CID can be achieved by modifying the dimerizers with photo-caged molecules that are removed upon illumination. In the caged state, the dimerizers cannot access the active site of the protein due to steric hindrance caused by the caging moieties. Upon illumination, the steric hindrance is eliminated as the caging moieties are irreversibly released (typically with UV light), hence resulting in photo-activation. The time scales of such processes range from seconds to minutes. Photo-caging dimerizers are relevant for photopharmacology as they enable dose-dependent photo-activation of signaling pathways and cellular events.

Photo-caging molecules commonly used are the photo-cleavable nitrobenzyl derivatives (Figure 2.4a).^[23] As an example, the immunosuppressive drug rapamycin can be modified

into a photo-caged rapamycin dimer using nitrobenzyl derivatives to control the dimerization of modular proteins FKBP12 and FRB, which regulate various processes including transcription, protein localization and enzyme catalysis, using UV light ($\lambda = 365$ nm).^[47] Similar photo-caging approach using photo-removable 4,5-dimethoxy-2-nitrobenzyl is demonstrated for a plant hormone, abscisic acid, to control the dimerization of the so-called ABI and PYL proteins using UV light.^[48] Such photo-caged systems enable photo-modulation of regulatory processes in living cells such as signal transduction, translocation of proteins and cytoskeletal remodeling. The concept of photo-caging has been expanded to other dimerizers, including the phytohormone gibberellic acid^[49] and antibiotic trimethoprim.^[50, 51] Moreover, the caging group can be used as a photo-cleavable linker that acts as a bridge that connects two different proteins together. This concept is demonstrated by covalently linking a HaloTag ligand chloroalkane and a SNAP-tag ligand O6-benzylguanine together with a photo-labile methyl-6-nitroveratryl group in between (Figure 2.4a).^[52] Upon irradiation with UV light, the linker is cleaved, liberating the ligands and hence the dimer into two separate moieties, further allowing photo-modulation over protein relocation. Although it is efficient to control cellular events, photo-caging approach typically does not allow an inherent reversible control and therefore is not suitable if an on/off system is desired. Such processes can be reversed for instance by adding a competitor binding.^[46,53]

An option for reversible optical dimerizers is already exemplified by Nature in the form of blue light-responsive LOV system. As elaborated in the previous section on allosteric regulation, in some examples such as in algal aureochrome, the photo-switching of LOV activates homodimerization that further enables binding event.^[33,54] Similar mechanism is observed in the marine bacterium *E. litoralis* in the activation of a transcription factor called EL222.^[55,56] Naturally occurring LOV further inspires the design of optical dimerizer tags called the tunable light-inducible dimerization tags or TULIPs.^[57] TULIPs harness the interaction of LOV from plant phototropin (*A. sativa*) and an engineered PDZ domain as binding partners (i.e., a protein domain in the signaling proteins across the kingdoms of life).

Figure 2.4 Strategies to photo-modulate dimerization events. a) A photo-cleavable synthetic dimerizer consisting of a photo-labile nitroveratryl group (blue) bridging two ligands: one ligand is a substrate for a protein tag called SNAP-tag (red) and the other is a substrate for a protein tag called HaloTag (green). Upon irradiation with UV light, the dimer disassociates as the linker (blue) is photo-cleaved. Reprinted with permission from ref [52]. Copyright © 2014, Wiley-VCH Verlag GmbH & Co. b) Light-inducible dimerization tag based on LOV-PDZ binding partners that are responsive to blue light. Reprinted with permission from ref [57]. Copyright © 2012, Macmillan Publishers Limited, part of Springer Nature. c) Plant phytochrome B (PhyB)-PIF binding partners that are responsive to red/far red light. Reprinted with permission from ref [61]. Copyright © 2009, Macmillan Publishers Limited, part of Springer Nature.

To generate light-inducible tags, the LOV domain (~125 residues) is fused through the J α helix to a specific peptide sequence that binds to PDZ (~194 residues). In the dark state, the J α helix physically blocks the peptide from binding to PDZ domain (Figure 2.4b).^[57] Upon irradiation with blue light, the J α helix unfolds and exposes the peptide to allow the interaction with PDZ domain, hence resulting in photo-activated dimerization of PDZ and LOV. The tags have been employed to control protein localization using laser in living yeast. However, the presence of PDZ domain and PDZ-binding peptide in the system raises concern that the cross-talk with endogenous signaling pathways might occur. Therefore, to improve orthogonality, PDZ as the binding partner can be replaced with a small adaptor protein called SspB from *E. coli*.^[58] In this case, a peptide tag that binds SspB called the SsrA peptide from *E. coli* is fused to LOV domain. Coupled with point mutations that stabilize the J α helix in the dark (optimized from computational library screening), the tag leads to changes in binding affinity for SspB over 50-fold upon blue light illumination. The photo-induced binding occurs within seconds, while the reversion process occurs in the dark within minutes. In addition to LOV-based systems, other light-inducible dimerizers are developed from plant *Arabidopsis thaliana*, consisting of blue light-responsive flavoprotein cryptochrome 2 and its interacting transcription factor CIB1^[59] as well as UV light-responsive photoreceptor UVR8 with its binding partner COP1.^[60]

As an alternative to the blue light/UV light-responsive systems, dimerization tags that are activated by red light irradiation ($\lambda \geq 650$ nm) are derived from *Arabidopsis thaliana* phytochrome B (PhyB) and its interacting factors (PIF3 and PIF6).^[61-63] Structurally, PhyB binds a chromophore namely phycocyanobilin (PCB) and is not capable of PIF binding in the dark (Figure 2.4c). Upon red light illumination ($\lambda = 650$ nm), the chromophore undergoes a conformational change that results in the structural change of PhyB, further allowing binding to PIF (Figure 2.4c).^[61] PhyB-PIF binding is reversible upon illumination with far-red light ($\lambda = 750$ nm), which has been used to control protein localization in eukaryotic cells such as yeast and mammalian cells. A set of comparative studies using yeast transcriptional assay further reveals that PhyB/PIF3 dimerization tag allows a higher level of activation and a lower background than TULIPs and PhyB/PIF6 systems.^[64] This system therefore holds promise for photopharmacology, also because the reversible activation using red light wavelength enables better tissue penetration than blue or UV light.^[65] Moreover, TULIPs

system raises concern of toxicity when used for regulation of yeast transcription (observed from cellular growth defects).^[64]

2.3.2 Photo-responsive cage-like and filament proteins

Expanding the concept of red/far-red responsive PhyB/PIF interaction beyond dimerization of monomeric proteins enables a photo-control of higher-ordered protein assemblies, for instance the viral protein cages. An adeno-associated virus is engineered with PIF6 motifs at its exterior to enable light-dependent interaction with nuclear localization sequence that is tagged with PhyB (Figure 2.5a).^[66] Illumination with either red or far-red light results in the changes of translocation of viruses into the host nucleus, allowing control of gene delivery upon irradiation with different wavelengths (Figure 2.5a). Another approach to modulate protein cages functionality using light is demonstrated by engineering metal-storing ferritin cages to incorporate light-responsive manganese (Mn)-carbonyl moieties in the interior of the cage.^[67] Such metal complexes inherently dissociate and release carbon monoxide upon irradiation with visible light,^[68] allowing the generation of CO-releasing ferritin that is modulated with light. Furthermore, naturally occurring photo-active moieties can be used to trigger oligomerization and clustering of protein molecules. Such control is achieved by incorporation of plant *A. thaliana* cryptochrome 2 that naturally assembles into the so-called photobodies in plant cells when irradiated with blue light.^[69,70] Genetically incorporating cryptochrome 2 into signaling proteins such as RhoA (i.e., a GTPase that mediates cellular tension and cytoskeletal contraction) enables photo-activation of signaling pathways in mammalian cells.^[70]

Nevertheless, incorporation of bulky photo-active domains is rather difficult to result in a precise control over the assembly process of large and highly defined structures, such as cage-like proteins and filaments. Attachment of synthetic photo-switches due to their smaller size offers a safer approach in order to avoid excessive perturbations that might disfavor assembly altogether. Owing to the complexity of higher-ordered structures and their natural tendency to favor assembly, to be able to use light to interfere with such assemblies is a challenging task. Photo-switches need to be carefully introduced into the structure of the target protein to result in sufficient structural perturbations that allow photo-modulation, without compromising the ability to assemble. Furthermore, highly-defined assemblies are

usually robust and modifications of the exterior of such structures, for instance bacterial encapsulin nanocages^[71] and filamentous M13 bacteriophage,^[72] with photo-switchable molecules (i.e., spiropyran and azobenzene, respectively) have been shown to not compromise the stability of the assembled structures upon photo-switching.

A remarkable photo-control over the assembly behavior of a cage-like protein is demonstrated in the case of homo-oligomeric cage-like group II chaperonin from *M. maripaludis*.^[73] This chaperonin naturally exists in two states: open and closed, depending on the binding and hydrolysis of nucleotide (ATP) molecules. To control the opening/closing of the chaperonin with light, azobenzene molecules are covalently placed at the interface of the protein subunits, acting as a spacer crosslinking cysteine residues between the subunits (Figure 2.5b).^[73] Light-responsive isomerization of the *trans* and *cis* forms of azobenzene crosslinkers results in a distance change between the subunits (i.e., 18 Å in the *trans* state and 5 – 12 Å in *cis* state). Using computational screening, the cysteine residues to which the azobenzene are attached are genetically introduced to match the distance of the *cis* and *trans* forms of azobenzene. Upon irradiation with UV light ($\lambda = 365$ nm), the spacers are in their *cis* form, resulting in the chaperonin cage favoring the closed state. Upon subsequent blue light irradiation ($\lambda = 450$ nm), the spacers convert to their extended *trans* form, allowing the opening of the chaperonin cage (Figure 2.5b).

Revealed by *in silico* investigation, the success of this approach relies on two important aspects: first, the energy landscape, which cooperatively favors either of the two dominant states (i.e., either open or closed)^[74] and second, the strong allosteric coupling between the orientation of the subunits (affected by the distance change from the photo-active azobenzene spacers) with the structural rearrangements of the nucleotide-binding pocket.^[75] The allosteric coupling between the two events is mediated by a rigid-body rocking and rotation of the subunits due to light-triggered distance change (Figure 2.5b).^[74,75] The movement subsequently results in the destabilization of protein-nucleotide interaction in the nucleotide-binding pocket, leading to the opening of the cage.

A complementary approach to modification of protein structures is to incorporate photo-switchable moieties into regulatory molecules that chemically affect the assembly process (e.g., activators and inhibitors). This approach has been successfully used to activate or inhibit the assembly of cellular filaments, the microtubules.

Figure 2.5 Strategies to generate photo-responsive cage-like and filament proteins. a) PhyB-PIF binding partner that is responsive to red/far red light is used to control cellular localization of adeno-associated virus using light. Reprinted with permission from ref [66]. Copyright © 2015, American Chemical Society. b) Opening and closing of a protein cage is modulated with light enabled by azobenzene spacers between protein subunits. Reprinted with permission from ref [73] and ref [74]. Copyright © 2013, Macmillan Publishers Limited, part of Springer Nature. c) Photo-responsive analogue, photostatin (PST-1), of a potent inhibitor of assembly of microtubule filaments

(combretastatin A-4). PST-1 is only active in its *cis* form, hence allowing photo-regulated inhibition of microtubule assembly. Reprinted with permission from ref [78]. Copyright © 2015, Elsevier Inc.

In the case of photo-activation of microtubule assembly, the nucleotide substrate GTP that is required for self-assembly is caged with a photo-cleavable nitrobenzyl derivative.^[76] Due to the steric hindrance of the caged form of GTP, the rate of microtubule assembly is significantly reduced in the absence of illumination. Upon UV flash photolysis, the photo-labile moiety is cleaved off and the GTP is released within milliseconds, resulting in the activation of the microtubule assembly.

On the other hand, a class of active molecules called the combretastatin A-4 (CA4) (Figure 2.5c) is identified as a potent inhibitor of microtubule assembly in its *cis* form, but not in its *trans* form.^[77] Replacing the C=C bond in the CA4 structure with N=N bond to mimic azobenzene (Figure 2.5c) results in a photo-active derivative of CA4 called the photostatins (PST).^[78] PST switches from its inactive *trans* to active *cis* form upon irradiation at $\lambda = 390 - 430$ nm, which is reversible upon irradiation at $\lambda = 500 - 530$ nm or by thermal relaxation (in the dark, half-life = 6 min), allowing a fully reversible control over the drug cytotoxicity. The activity of the *trans* form in the dark is about 250-fold less compared to the *cis* form upon irradiation with blue light. This remarkable difference has been demonstrated to photo-modulate the microtubule dynamic assembly in live cells of *C. elegans* embryo.^[78]

2.4 Photo-responsive artificial protein-based systems and higher-order structures

Synthetic photo-switches can be used to bridge proteins together to form artificial extended nanostructures in which the assembly event is modulated with light. Here, we focus on three examples of artificial higher-order structures mediated by three different photo-switches: azobenzene, spiropyran and a photo-labile group. In the first design, azobenzene is incorporated in the bridging fragment of a nanowire-like assembly of stable protein one (SP1), a ring-shaped protein from plant *Populus tremula* (Figure 2.6a).^[79,80] The stacking of SP1 to form nanowires is structurally mediated by a globular poly(amido-amine) or PAMAM dendrimer via electrostatic interaction.^[80] Incorporation of azobenzene in the PAMAM structure results in the bending of the overall structure upon UV light irradiation when the

azobenzene switches from *trans* to *cis* conformation (Figure 2.6a). The curved nanowires can be converted back to the straight conformation upon visible light irradiation. In such design, the rigidity of the bridging moiety is important to introduce a significant change in the interspace angle between the stacking proteins, hence cooperatively resulting in the bending of the entire assembly.

Figure 2.6 Strategies to generate photo-responsive higher-order structures. a) Reversible formation of curved nanowire of SP1 protein stacks (blue) upon irradiation enabled by incorporation of azobenzene in the PAMAM-based bridging fragment (red). Reprinted with permission from ref [80]. Copyright © 2016, The Royal Society of Chemistry. b) Reversible assembly of tubular structure of GroEL protein stacks (blue) upon irradiation enabled by incorporation of spiropyran/merocyanine as the bridging fragment. Reprinted with permission from ref [83]. Copyright © 2013, American Chemical Society. c) Photo-triggered dissociation of an array of CCMV protein cages upon irradiation enabled by incorporation of a photo-labile group in the dendrimer linkers. Reprinted with permission from ref [85]. Copyright © 2010, Macmillan Publishers Limited, part of Springer Nature.

In the second design, spiropyran/merocyanine switch is used to bridge a tubular chaperone protein GroEL to stack into a filamentous structure (up to 2.5 μm long). The

opening of the cylindrical cavity of GroEL is genetically modified with 14 cysteine residues to enable attachment of photo-switches bearing a maleimide moiety.^[81, 82] Upon UV light irradiation ($\lambda = 280$ nm) at room temperature, the spiropyran opens into its merocyanine form, which enables electrostatic interaction with magnesium ions (Figure 2.6b).^[83] Mg^{2+} ions mediate the interaction between the stacking GroEL proteins via merocyanine- Mg^{2+} -merocyanine bridges. As the closed spiropyran cannot interact with Mg^{2+} , the assembled structure collapses into shorter fragments upon irradiation with visible light ($\lambda > 400$ nm) when the merocyanine form is converted back to spiropyran form.^[83] Notably, the light-induced reversibility is achieved only when a radical scavenger dithiothreitol (DTT) is present in the system.^[82,83] The authors postulated that radical species might have been generated upon photoexcitation that could covalently cross-link the stacking GroEL together.^[83] Therefore, addition of a radical scavenger like DTT is necessary to prevent any covalent cross-linking and ensure the reversibility of the system. Moreover, the cylindrical cavity of the stacking GroEL is able to confine cargo molecules such as drugs, enabling the design of photo-responsive nanocarriers.^[84]

In the third design, a photo-responsive array of a viral capsid protein namely Cowpea Chlorotic Mottle Virus (CCMV) is generated via electrostatic interaction between negatively-charged exterior of the capsid and positively-charged dendrimers.^[85] A photo-labile nitrobenzyl linker is embedded in the structure of the bridging dendrimers (Figure 2.6c). Upon UV light irradiation ($\lambda = 365$ nm), the photo-labile moiety is cleaved, leading to the collapse of the bridging structures. Consequently, the CCMV array dissociates into individual particles upon UV light irradiation. A similar design can also be applied to other protein cages, such as the magnetoferritin.^[85]

2.5 Perspectives

Strategies on modulating naturally occurring proteins with photo-switchable moieties, either chemically or genetically introduced, would enable interfering with complex regulations that orchestrate cellular events, particularly in relation to photopharmacology.^[86] Even when a strategy does not result in a complete on/off photo-control but rather a modest difference between the dark and lit-state, still a significant effect can be expected as the target

proteins are usually involved in interlinking networks of cellular regulations. A fluctuation triggered with light at one point in a cascade system would result in changes in the overall dynamics of the process. Among the presented strategies, generating a library of photo-responsive analogue drugs is a promising means to gain effective and reversible control over a specific system. In terms of reversibility, the drugs ideally should undergo thermal deactivation spontaneously, hence is only active upon irradiation at the desired site. Nevertheless, the challenge remains in the stability of the designed drugs (especially against spontaneous hydrolysis), orthogonality and specificity of action, as well as the means of delivery of the designed drugs.

Moreover, most of the presented systems are active upon UV light illumination, which is not the ideal option to apply in physiological processes due to the weak tissue penetration. Red and far-red light with longer wavelengths are more suited to carry such task.^[65] Nature has presented an example of red light-responsive system in plant phytochrome B and its interacting factors.^[69] As a complementary approach, synthetic chemists have made attempts to design and develop photo-switches that are responsive to red light illumination, such as the azo-derivatives.^[87-90] In the future, the application of such strategies based on red light-responsive systems shall be expanded, particularly for *in vivo* studies. In a complementary manner, artificially designed extended nanostructures are envisioned to stimulate the development of photo-tunable arrays and carriers^[91] and/or molecular labels for functional assemblies and supramolecular platforms.^[92-94]

2.6 References

- [1] A. Typas, V. Sourjik, *Nat. Rev. Microbiol.* **2015**, *13*, 559-572.
- [2] R. J. Arnold, J. P. Reilly, *Anal. Biochem.* **1999**, *269*, 105-112.
- [3] E. Brody, J. Abelson, *Science* **1985**, *228*, 963-967.
- [4] C. L. Will, R. Luhrmann, *Curr. Opin. Cell Biol.* **1997**, *9*, 320-328.
- [5] B. Alberts, *Cell* **1998**, *92*, 291-294.
- [6] T. Pawson, P. Nash, *Science* **2003**, *300*, 445-452.
- [7] H. N. Motlagh, J. O. Wrabl, J. Li, V. J. Hilser, *Nature* **2014**, *508*, 331-339.
- [8] M. H. Ali, B. Imperiali, *Bioorg. Med. Chem.* **2005**, *13*, 5013-5020.

- [9] B. N. Kholodenko, *Nat. Rev. Mol. Cell Bio.* **2006**, 7, 165-176.
- [10] A. Gautier, C. Gauron, M. Volovitch, D. Bensimon, L. Jullien, S. Vriza, *Nat. Chem. Biol.* **2014**, 10, 533-541.
- [11] W. Szymanski, J. M. Beierle, H. A. V. Kistemaker, W. A. Velema, B. L. Feringa, *Chem. Rev.* **2013**, 113, 6114-6178.
- [12] K. E. Brechun, K. M. Arndt, G. A. Woolley, *Curr. Opin. Struct. Biol.* **2016**, 45, 53-58.
- [13] G. Guglielmi, H. J. Falk, S. De Renzis, *Trends Cell Biol.* **2016**, 26, 864-874.
- [14] M. Baker, *Nat. Methods* **2012**, 9, 443-446.
- [15] C. W. Riggsbee, A. Deiters, *Trends Biotechnol.* **2010**, 28, 468-475.
- [16] U. T. Bornscheuer, *Nature* **2016**, 540, 345-346.
- [17] F. Z. Zhang, A. Zarrine-Afsar, M. S. Al-Abdul-Wahid, R. S. Prosser, A. R. Davidson, G. A. Woolley, *J. Am. Chem. Soc.* **2009**, 131, 2283-2289.
- [18] A. Kocer, M. Walko, W. Meijberg, B. L. Feringa, *Science* **2005**, 309, 755-758.
- [19] C. L. Tucker, J. D. Vrana, M. J. Kennedy, *Curr. Protoc. Cell Biol.* **2014**, 64, 17.16-17.20.
- [20] R. J. Mart, R. K. Allemann, *Chem. Commun.* **2016**, 52, 12262-12277.
- [21] A. A. Beharry, G. A. Woolley, *Chem. Soc. Rev.* **2011**, 40, 4422-4437.
- [22] R. Klajn, *Chem. Soc. Rev.* **2014**, 43, 148-184.
- [23] G. Mayer, A. Heckel, *Angew. Chem. Int. Ed.* **2006**, 45, 4900-4921.
- [24] B. L. Taylor, I. B. Zhulin, *Microbiol. Mol. Biol. Rev.* **1999**, 63, 479-506.
- [25] A. S. Baker, A. Deiters, *ACS Chem. Biol.* **2014**, 9, 1398-1407.
- [26] A. W. Fenton, *Trends Biochem. Sci.* **2008**, 33, 420-425.
- [27] J. M. Christie, M. Salomon, K. Nozue, M. Wada, W. R. Briggs, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 8779-8783.
- [28] S. M. Harper, L. C. Neil, K. H. Gardner, *Science* **2003**, 301, 1541-1544.
- [29] S. Crosson, K. Moffat, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 2995-3000.
- [30] M. Salomon, W. Eisenreich, H. Durr, E. Schleicher, E. Knieb, V. Massey, W. Rudiger, F. Muller, A. Bacher, G. Richter, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 12357-12361.
- [31] S. M. Harper, J. M. Christie, K. H. Gardner, *Biochemistry* **2004**, 43, 16184-16192.
- [32] F. Circolone, J. Granzin, K. Jentzsch, T. Drepper, K. E. Jaeger, D. Willbold, U. Krauss, R. Batra-Safferling, *J. Mol. Biol.* **2012**, 417, 362-374.

- [33] A. Banerjee, E. Herman, M. Serif, M. Maestre-Reyna, S. Hepp, R. Pokorny, P. G. Kroth, L. O. Essen, T. Kottke, *Nucleic Acids Res.* **2016**, *44*, 5957-5970.
- [34] E. Herman, T. Kottke, *Biochemistry* **2015**, *54*, 1484-1492.
- [35] M. Ohki, K. Sugiyama, F. Kawai, H. Tanaka, Y. Nihei, S. Unzai, M. Takebe, S. Matsunaga, S. Adachi, N. Shibayama, Z. W. Zhou, R. Koyama, Y. Ikegaya, T. Takahashi, J. R. H. Tame, M. Iseki, S. Y. Park, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 6659-6664.
- [36] A. Winkler, A. Udvarhelyi, E. Hartmann, J. Reinstein, A. Menzel, R. L. Shoeman, I. Schlichting, *J. Mol. Biol.* **2014**, *426*, 853-868.
- [37] J. M. Christie, J. Gawthorne, G. Young, N. J. Fraser, A. J. Roe, *Mol. Plant* **2012**, *5*, 533-544.
- [38] J. Lee, M. Natarajan, V. C. Nashine, M. Socolich, T. Vo, W. P. Russ, S. J. Benkovic, R. Ranganathan, *Science* **2008**, *322*, 438-442.
- [39] C. Cosentino, L. Alberio, S. Gazzarrini, M. Aquila, E. Romano, S. Cermenati, P. Zuccolini, J. Petersen, M. Beltrame, J. L. Van Etten, J. M. Christie, G. Thiel, A. Moroni, *Science* **2015**, *348*, 707-710.
- [40] D. Strickland, K. Moffat, T. R. Sosnick, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10709-10714.
- [41] M. Volgraf, P. Gorostiza, R. Numano, R. H. Kramer, E. Y. Isacoff, D. Trauner, *Nat. Chem. Biol.* **2006**, *2*, 47-52.
- [42] S. J. Zhu, M. Riou, C. A. Yao, S. Carvalho, P. C. Rodriguez, O. Bensaude, P. Paoletti, S. X. Ye, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6081-6086.
- [43] I. Pittolo, X. Gomez-Santacana, K. Eckelt, X. Rovira, J. Dalton, C. Goudet, J. P. Pin, A. Llobet, J. Giraldo, A. Llebaria, P. Gorostiza, *Nat. Chem. Biol.* **2014**, *10*, 813-815.
- [44] N. J. Marianayagam, M. Sunde, J. M. Matthews, *Trends Biochem. Sci.* **2004**, *29*, 618-625.
- [45] S. Jones, J. M. Thornton, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13-20.
- [46] S. Voss, L. Klewer, Y. W. Wu, *Curr. Opin. Chem. Biol.* **2015**, *28*, 194-201.
- [47] K. A. Brown, Y. Zou, D. Shirvanyants, J. Zhang, S. Samanta, P. K. Mantravadi, N. V. Dokholyan, A. Deiters, *Chem. Commun.* **2015**, *51*, 5702-5705.
- [48] C. W. Wright, Z. F. Guo, F. S. Liang, *ChemBioChem* **2015**, *16*, 254-261.

- [49] K. M. Schelkle, T. Griesbaum, D. Ollech, S. Becht, T. Buckup, M. Hamburger, R. Wombacher, *Angew. Chem. Int. Ed.* **2015**, *54*, 2825-2829.
- [50] E. R. Ballister, C. Aonbangkhen, A. M. Mayo, M. A. Lampson, D. M. Chenoweth, *Nat. Commun.* **2014**, *5*, 5475.
- [51] E. R. Ballister, S. Ayloo, D. M. Chenoweth, M. A. Lampson, E. L. F. Holzbaur, *Curr. Biol.* **2015**, *25*, R407-R408.
- [52] M. Zimmermann, R. Cal, E. Janett, V. Hoffmann, C. G. Bochet, E. Constable, F. Beaufils, M. P. Wymann, *Angew. Chem. Int. Ed.* **2014**, *53*, 4717-4720.
- [53] P. Liu, A. Calderon, G. Konstantinidis, J. Hou, S. Voss, X. Chen, F. Li, S. Banerjee, J. E. Hoffmann, C. Theiss, L. Dehmelt, Y. W. Wu, *Angew. Chem. Int. Ed.* **2014**, *53*, 10049-10055.
- [54] Y. Akiyama, Y. Nakasone, Y. Nakatani, O. Hisatomi, M. Terazima, *J. Phys. Chem. B* **2016**, *120*, 7360-7370.
- [55] B. D. Zoltowski, L. B. Motta-Mena, K. H. Gardner, *Biochemistry* **2013**, *52*, 6653-6661.
- [56] A. I. Nash, R. McNulty, M. E. Shillito, T. E. Swartz, R. A. Bogomolni, H. Luecke, K. H. Gardner, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9449-9454.
- [57] D. Strickland, Y. Lin, E. Wagner, C. M. Hope, J. Zayner, C. Antoniou, T. R. Sosnick, E. L. Weiss, M. Glotzer, *Nat. Methods* **2012**, *9*, 379-384.
- [58] G. Guntas, R. A. Hallett, S. P. Zimmerman, T. Williams, H. Yumerefendi, J. E. Bear, B. Kuhlman, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 112-117.
- [59] M. J. Kennedy, R. M. Hughes, L. A. Peteya, J. W. Schwartz, M. D. Ehlers, C. L. Tucker, *Nat. Methods* **2010**, *7*, 973-975.
- [60] R. P. Crefcoeur, R. H. Yin, R. Ulm, T. D. Halazonetis, *Nat. Commun.* **2013**, *4*, 1779.
- [61] A. Levskaya, O. D. Weiner, W. A. Lim, C. A. Voigt, *Nature* **2009**, *461*, 997-1001.
- [62] G. G. Yu, H. Onodera, Y. Aono, F. Kawano, Y. Ueda, A. Furuya, H. Suzuki, M. Sato, *Sci. Rep.* **2016**, *6*, 35777.
- [63] W. E. Georgianna, A. Deiters, *ChemBioChem* **2010**, *11*, 301-303.
- [64] G. P. Pathak, D. Strickland, J. D. Vrana, C. L. Tucker, *ACS Synth. Biol.* **2014**, *3*, 832-838.
- [65] R. R. Anderson, J. A. Parrish, *J. Invest. Dermatol.* **1981**, *77*, 13-19.
- [66] E. J. Gomez, K. Gerhardt, J. Judd, J. J. Tabor, J. Suh, *ACS Nano* **2016**, *10*, 225-237.

- [67] K. Fujita, Y. Tanaka, S. Abe, T. Ueno, *Angew. Chem. Int. Ed.* **2016**, *55*, 1056-1060.
- [68] R. Motterlini, J. E. Clark, R. Foresti, P. Sarathchandra, B. E. Mann, C. J. Green, *Circ. Res.* **2002**, *90*, E17-E24.
- [69] P. Mas, P. F. Devlin, S. Panda, S. A. Kay, *Nature* **2000**, *408*, 207-211.
- [70] L. J. Bugaj, A. T. Choksi, C. K. Mesuda, R. S. Kane, D. V. Schaffer, *Nat. Methods* **2013**, *10*, 249-252.
- [71] R. M. Putri, J. W. Fredy, J. J. L. M. Cornelissen, M. S. T. Koay, N. Katsonis, *ChemPhysChem* **2016**, *17*, 1815-1818.
- [72] M. Murugesan, G. Abbineni, S. L. Nimmo, B. R. Cao, C. B. Mao, *Sci. Rep.* **2013**, *3*, 1820.
- [73] D. Hoersch, S. H. Roh, W. Chiu, T. Kortemme, *Nat. Nanotechnol.* **2013**, *8*, 928-932.
- [74] G. A. Woolley, *Nat. Nanotechnol.* **2013**, *8*, 892-893.
- [75] D. Hoersch, T. Kortemme, *Structure* **2016**, *24*, 576-584.
- [76] A. Marx, A. Jagla, E. Mandelkow, *Eur. Biophys. J.* **1990**, *19*, 1-9.
- [77] G. C. Tron, T. Pirali, G. Sorba, F. Pagliai, S. Busacca, A. A. Genazzani, *J. Med. Chem.* **2006**, *49*, 3033-3044.
- [78] M. Borowiak, W. Nahaboo, M. Reynders, K. Nekolla, P. Jalinot, J. Hasserodt, M. Rehberg, M. Delattre, S. Zahler, A. Vollmar, D. Trauner, O. Thorn-Seshold, *Cell* **2015**, *162*, 403-411.
- [79] H. C. Sun, L. Miao, J. X. Li, S. Fu, G. An, C. Y. Si, Z. Y. Dong, Q. Luo, S. J. Yu, J. Y. Xu, J. Q. Liu, *ACS Nano* **2015**, *9*, 5461-5469.
- [80] H. C. Sun, L. L. Zhao, T. T. Wang, G. An, S. Fu, X. M. Li, X. L. Deng, J. Q. Liu, *Chem. Commun.* **2016**, *52*, 6001-6004.
- [81] S. Muramatsu, K. Kinbara, H. Taguchi, N. Ishii, T. Aida, *J. Am. Chem. Soc.* **2006**, *128*, 3764-3769.
- [82] S. Biswas, K. Kinbara, N. Oya, N. Ishii, H. Taguchi, T. Aida, *J. Am. Chem. Soc.* **2009**, *131*, 7556-7557.
- [83] T. Sendai, S. Biswas, T. Aida, *J. Am. Chem. Soc.* **2013**, *135*, 11509-11512.
- [84] S. Biswas, K. Kinbara, T. Niwa, H. Taguchi, N. Ishii, S. Watanabe, K. Miyata, K. Kataoka, T. Aida, *Nat. Chem.* **2013**, *5*, 613-620.

- [85] M. A. Kostianen, O. Kasyutich, J. J. L. M. Cornelissen, R. J. M. Nolte, *Nat. Chem.* **2010**, 2, 394-399.
- [86] W. A. Velema, W. Szymanski, B. L. Feringa, *J. Am. Chem. Soc.* **2014**, 136, 2178-2191.
- [87] A. A. Beharry, O. Sadowski, G. A. Woolley, *J. Am. Chem. Soc.* **2011**, 133, 19684-19687.
- [88] Y. Yang, R. P. Hughes, I. Aprahamian, *J. Am. Chem. Soc.* **2014**, 136, 13190-13193.
- [89] S. Samanta, A. Babalhavaeji, M. X. Dong, G. A. Woolley, *Angew. Chem. Int. Ed.* **2013**, 52, 14127-14130.
- [90] M. X. Dong, A. Babalhavaeji, S. Samanta, A. A. Beharry, G. A. Woolley, *Acc. Chem. Res.* **2015**, 48, 2662-2670.
- [91] J. Olejniczak, C. J. Carling, A. Almutairi, *J. Control. Release* **2015**, 219, 18-30.
- [92] C. Grunwald, K. Schulze, A. Reichel, V. U. Weiss, D. Blaas, J. Piehler, K. H. Wiesmuller, R. Tampe, *Proc. Natl. Acad. Sci. USA* **2010**, 107, 6146-6151.
- [93] N. Laboria, R. Wieneke, R. Tampe, *Angew. Chem. Int. Ed.* **2013**, 52, 848-853.
- [94] N. L. Weisen, C. A. Hommersom, J. Voskuhl, S. Sankaran, A. M. Depauw, N. Katsonis, P. Jonkheijm, J. J. L. M. Cornelissen, *Chem. Commun.* **2017**, 53, 1896-1899.

Chapter 3

Programming allostery in the human serum albumin with light

The human serum albumin (HSA) is a promiscuous protein that is able to bind more than 120 types of ligand, and thus plays a pivotal role for transporting molecules in blood plasma. Allosteric regulation is prominent to the functionality of HSA to control the binding of multiple ligands through mechanisms that have not been entirely elucidated yet. We argue that developing strategies to interfere with allosteric mechanisms reversibly, and in a controlled manner, would thus constitute a relevant approach towards an improved understanding of vital cellular processes, and eventually this might allow controlling their regulation. Here, we perform the site-specific covalent incorporation of a synthetic molecular photo-switch into one of the binding pockets of HSA specifically (subdomain IA). In this engineered protein, light-induced molecular switching occurs at one binding pocket (subdomain IA) and interferes with ligand binding in the neighboring pocket (subdomain IB) and we show that this allosteric mechanism allows the photo-induced switching at IA to influence the overall flexibility of IB pocket. Our approach provides a tool to unravel additional insights into allostery in proteins, and introduces a means to amplify the effects of molecular switching in bio-synthetic systems.

3.1 Introduction

In proteins with multiple binding pockets, binding of a ligand in one binding site often influences the binding of other ligands to one or more binding sites, a process known as allostery.^[1-5] Allosteric regulation of protein activity is central to life and sometimes even referred to as ‘the second secret of life’ - the first one being the genetic code.^[2] However, the mechanisms by which proteins undergo allosteric regulation remain largely unknown, and these complex systems have already revealed a number of paradoxes. For example, tightly packed, fully folded proteins can display a remarkable structural plasticity dictated by their binding pockets. Also, it has been shown that a structural change in the protein backbone is not always necessary for the binding sites to communicate with each other.^[4]

Overall, it appears that allostery is facilitated in dynamic proteins, and in the state of our understanding, allostery is not simply a shape change-induced phenomenon as understood in early years,^[6] but rather a thermodynamic process in an ensemble.^[1-4] Pioneering work has also highlighted the limitations of deducing mechanism from the static structure alone,^[1,3] which calls for tools to manipulate, investigate and eventually engineer the dynamics of allosteric systems. A recent study reports on using temperature and pH control to increase the local conformational entropy of fused proteins that communicate through allostery.^[7] A complementary approach towards allosteric engineering consists in implementing allosteric behavior in wholly synthetic biomimetic systems to control ligand binding and catalysis.^[8,9] However, previous approaches lack of selectivity in space and time and for protein systems operating at physiological conditions, environmental changes such as in pH and temperature could lead to unfolding and loss of activity.

In contrast, light offers high spatiotemporal control while being relatively non-destructive toward proteins and applicable in physiological conditions.^[10-14] In order to make proteins light-switchable, the challenge is to make sure that even in aqueous systems and in the protein environment, the stability and photo-switching properties of molecular switches are preserved. Light-modulated proteins have been developed to control biological activities, such as directing the secondary structures of peptides/proteins,^[15,16] enzymatic catalysis,^[17,18] protein-ligand binding,^[19,20] modulation of protein channels^[21,22] and receptors in neural networks.^[23] Recently, the emergence of photo-switchable allosteric modulators^[24-26] as well

as light-activated allosteric channel and DNA-binding protein^[27,28] marks the era of combining allosteric regulation with tailored optical control.

The human serum albumin (HSA) is characterized by its extraordinary capability to bind over 120 types of ligands, despite its relatively simple monomeric structure.^[29,30] Owing to its promiscuity in ligand binding, HSA is the main carrier of a variety of compounds in blood plasma. Structurally, HSA is a small protein (66 kDa) that is dominated by α -helices and loops.^[29-31] HSA has multiple binding sites (three primary ligand binding sites and seven sites for fatty acid binding) that contribute to its capability of binding a broad range of ligands.^[32] Allostery in such a promiscuous protein remains a challenging phenomenon to decipher, not only due to the inherent structural plasticity of the protein, but also due to the lack of consensus in the type of cooperativity between binding pockets: it can be either a positive cooperation, a negative cooperation, or no effect at all depending on the choice of ligands.^[29] Recent investigations by ultrafast time-resolved spectroscopy have suggested that the allostery of HSA is likely to originate in a ballistic and anisotropic energy flow through connecting helical structures.^[32]

Here, we design and investigate a photo-responsive hybrid system that contributes to unveiling the mechanistic aspects of cooperative actions between distinct pockets/sites of HSA. We achieve an optical control over a specific binding site of HSA (i.e., a site known as subdomain IB), by coupling a single spiropyran photo-switch to a neighboring site (i.e., a cysteine residue at subdomain IA). In this way, we are able to design a photo-switchable HSA where the photo-induced molecular switching occurs at IA site, while the effect is propagated to IB site. This strategy allows to externally interfere with allostery using light as a trigger.

3.2 Results and discussion

3.2.1 Design and synthesis of a photo-switchable hybrid

We connect a spiropyran photo-switch to the subdomain IA site covalently, by modifying the single surface-accessible free cysteine (Cys34) that is available in the native structure (Figure 3.1a). Modification of this single cysteine located in subdomain IA is anticipated to allow a selective photo-switching instead of disrupting the entire protein structure upon

irradiation. Moreover, previous experimental studies have suggested that Cys34 plays a key role in the allosteric regulation of HSA, although the associated mechanism remains largely unknown.^[30] The binding site of HSA examined in this study is located in subdomain IB (residue number 108 – 197), a separate entity that is adjacent to the engineered IA site (residue number 4 – 107).^[30] The modification site of Cys34 is 1.3 – 2.2 nm away from the examined IB binding site (distance relative to binding residues of IB: Lys137, Arg186, Arg117).

Figure 3.1 Design and synthesis of a photo-switchable hybrid protein. a) The hybrid protein is composed by the human serum albumin (HSA) that is covalently connected to a spiropyran photo-switch (SP) via Cys34 (red). IA and IB subdomains of HSA are represented in purple and green, respectively. b) Under irradiation with UV light, the spiropyran photo-switch (SP) converts into its open zwitterionic isomer, the merocyanine (MC). c) Size-exclusion elution profile confirms the formation of a hybrid HSA-SP eluting at $V = 16$ mL. The absorption at $\lambda = 350$ nm (blue line) indicates the presence of spiropyran. Prior to size-exclusion chromatography, the sample is dialyzed overnight against Tris-HCl buffer (pH 7.5). d) UV-visible spectra of HSA-SP hybrid purified by size-exclusion chromatography (blue line) or spin-filtration (red line), revealing a 1:1 ratio of protein over spiropyran switch.

The spiropyran is chosen as the photo-switch, because it undergoes significant changes in its conformation but also because it forms a zwitterion under irradiation with light, which modifies electrostatic interactions locally (Figure 3.1b).^[12] Upon irradiation with ultraviolet (UV) light, the closed spiropyran (SP) is converted to its open, zwitterionic form, the

merocyanine (MC). The conversion of SP to MC is reversible upon irradiation with visible light (Figure 3.1b). The hybrid protein HSA-SP is prepared at room temperature, by coupling the free cysteine of HSA to a nitro-spiropyran bearing a maleimide function, via Michael addition in a Tris-HCl buffer pH 7.5 that also contains 20% DMSO as an organic solvent (to solubilize the photo-switch). The nitro group ensures the stability of the photo-switch in a water-based system by stabilizing the MC (i.e., stabilizing the $-O^-$). It is important to work at pH 7.5 as SP can spontaneously be converted into protonated MC in an acidic environment.^[12] Basic conditions also need to be avoided due to risk of alkaline hydrolysis of the maleimide.^[33]

The presence of SP is indicated by its characteristic absorption with $\lambda_{\max} = 350$ nm at pH 7.5. Furthermore, co-elution of both SP (indicated by the absorption at $\lambda = 350$ nm) and HSA (indicated by the absorption at $\lambda = 280$ nm) around $V = 16$ mL by size-exclusion chromatography indicates that the photo-switch and the protein are coupled covalently (Figure 3.1c and d). We demonstrate that the coupling is quantitatively 1:1 based on absorption values at $\lambda = 280$ nm and $\lambda = 350$ nm (Figure 3.1d). Prolonging the reaction time, from one hour to up to 24 h does not have any effect on the coupling ratio.

Irradiating this albumin hybrid with UV light ($\lambda = 365$ nm, 5 min) triggers the appearance of an absorption band at $\lambda_{\max} = 550$ nm, which indicates the transformation of colorless SP form into the colored MC isomer, and demonstrates that the photochromic properties of the switch are preserved after its covalent attachment to the protein (Figure 3.2a). NMR spectroscopy allows determining the degree of photo-conversion from SP to MC at photostationary state, which is around 40% under our experimental conditions (Figure 3.2b), suggesting that a mixture of HSA-SP and HSA-MC is present at photostationary state. Prolonged irradiation, however, is not recommended as the heat caused by irradiation can disrupt the protein structure and even lead to partial aggregation.^[34] Further, upon irradiation with visible light ($\lambda \geq 420$ nm, 5 min), the band at $\lambda_{\max} = 550$ nm gradually decreases, indicating that some of the MC form switches back to the colorless SP form. UV-visible spectra indicate that this transformation is not fully reversible (Figure 3.2a), which might be related to the fact that the spiropyran also displays a weak absorption at $\lambda \geq 420$ nm, and thus a photostationary state is reached instead of full conversion. It is also known that the MC form is favored and stabilized in a polar environment compared to non-polar environment.^[12]

Figure 3.2 Photo-switchable HSA-SP. a) UV-visible spectra of the hybrid system before any irradiation (black line), after irradiation with UV light (blue line) and after irradiation with Vis light (red line). b) NMR spectra of SP before (red line) and after UV light irradiation (cyan line). c) CD spectra of the hybrid system before (black line), after irradiation with UV light (red line) and after irradiation with Vis light (blue line). d) DLS-based size distribution of the hybrid system before (blue) and after UV light irradiation (red).

We also analyze the effect of irradiation on the global structure of HSA-SP, in terms of conformation and hydrodynamic size. Circular dichroism (CD) spectra in phosphate buffer (pH 7.5) confirm that the HSA primarily consists of alpha helical and coiled segments (Figure 3.2c). Importantly, the CD spectra (Figure 3.2c) and the hydrodynamic size analysis (Figure 3.2d) do not reveal any significant structural perturbations after irradiation in the conditions of the experiment, confirming that the overall structure of the HSA hybrid remains intact. Nevertheless, the CD spectra indicate a slight decrease (around 2.9%) in the alpha helical content upon irradiation (Figure 3.2c, a decrease in the negative CD signal at around 217–228 nm). This decrease is likely associated to conformational changes occurring in the vicinity of the Cys34 as a result of SP to MC conversion.

3.2.2 Optical control over ligand binding to subdomain IB

We examine the behavior of the albumin hybrid, before and after irradiation with UV light, in terms of binding to methyl orange (MO). Methyl orange is a commonly used dye, that is known to bind primarily to subdomain IB,^[35] with an association constant of $2.3 \times 10^5 \text{ M}^{-1}$.^[36] It is also known that the binding of dyes to native HSA is associated with characteristic changes in their UV-visible spectra,^[36,37] which, for example, have been used to determine the concentration of albumin in blood serum.^[38] Here, we monitor spectral changes by UV-visible spectroscopy, in order to follow the binding/release of the ligand (Figure 3.3). For all experiments, an excess of MO ligand (~ 4 -equivalent) is added to HSA-SP ($3.7 \mu\text{M}$) and, before each spectrum is measured, the unbound MO is removed by spin-filtration. We exclude the possibility that this procedure modifies the equilibrium significantly in view of the large values of association constants that characterize the binding of ligands to HSA (typically larger than 10^4 M^{-1}). Also, the optical signature of the material is not modified significantly by spin-filtration (Figure 3.1d). Finally, from differences in UV-vis spectra before and after ligand binding we estimate that concentration of bound MO is $4.7 \mu\text{M}$ (protein:ligand ratio is 1:1.3).

At a fixed concentration of ligand, the absorption of free MO in the visible region ($\lambda_{\text{max}} = 460 \text{ nm}$) is lower than that of bound MO to HSA-SP (Figure 3.3a). Significantly, in the absence of spiropyran, we observe that irradiation of the ligand-HSA complex does not yield any significant spectral changes (Figure 3.3b). This blank experiment indicates that MO is stable and not responsive to UV light in our experimental conditions, thus any spectral modification induced by irradiation is associated to the photo-induced release/binding events primarily.

The binding of the MO ligand to the HSA-SP hybrid does not interfere with the photochromic properties of the photo-switch, as verified by UV-visible spectroscopy (Figure 3.3c). Further examination of the spectral contributions of MO shows a decrease of absorption at $\lambda = 460 \text{ nm}$ after 5 min irradiation with UV light (Figure 3.3d, orange line). The difference can be attributed to the release of MO from subdomain IB. The decrease in the absorption band indicates that about $(16 \pm 2)\%$ of the bound MO ligand is released upon irradiation. Subsequently, about $(8 \pm 1)\%$ of the released MO binds back after subsequent

irradiation with visible light for 5 min; which means that the reversibility is moderate. Each of the ligand release and re-binding experiments were repeated three times *in-situ*.

Figure 3.3 Photo-responsive binding of the methyl orange dye, to subdomain IB. a) MO spectra display a positive shift in the absorption in the visible range when bound to HSA-SP (black line) compared to the free dye (orange line) in Tris-HCl buffer pH 7.5. b) Spectra of HSA-bound MO (in the absence of SP) before (black line) and after UV light irradiation (red line). c) Photo-switching of HSA-SP in the presence of MO characterized with UV-visible spectroscopy, before any irradiation (red line), after UV light irradiation (pink line) and after Vis light irradiation (blue line). The ligand binding does not hinder the reversible conversion of SP to MC upon irradiation. d) Optical response of HSA-SP bound MO (orange line) monitored at $\lambda = 460$ nm.

The photo-induced release of the dye from subdomain IB – whereas the spiropyran switch is connected to subdomain IA, is also corroborated by investigating the diffusion behavior before and after irradiation using Pulsed-Field Gradient Spin-Echo (PFGSE) NMR experiments. The diffusion constant of methyl orange corresponding to the HSA-MO complex raises from $D_{\text{before}} = (4.11 \pm 0.02) \times 10^{-10} \text{ m}^2/\text{s}$ before irradiation, to $D_{\text{after}} = (4.28 \pm 0.02) \times 10^{-10} \text{ m}^2/\text{s}$ after irradiation. This light-induced increase in the diffusion constant confirms qualitatively that the release of the ligand occurs, because free dyes are expected to

diffuse faster, compared to the average of protein-bound dyes, thus resulting in a higher diffusion constant. For comparison, the diffusion constant of the dye, in the absence of albumin and measured in the same conditions is $D = (4.67 \pm 0.02) \times 10^{-10} \text{ m}^2/\text{s}$.

Figure 3.4 Photo-responsive bromocresol green binding to subdomain IB. a) BG spectra display a negative shift in the absorption in the visible range when bound to HSA-SP (black line) compared to free dye (blue line) in Tris-HCl buffer (pH 7.5). b) Spectra of HSA-bound BG (in the absence of SP) before (black line) and after UV light irradiation (red line). c) Photo-switching of HSA-SP in the presence of BG characterized with UV-visible spectroscopy, before any irradiation (red line), after UV light irradiation (pink line) and after irradiation with visible light (blue line). The ligand binding does not hinder the reversible conversion of SP to MC upon irradiation. d) Optical response of HSA-SP bound BG (blue line) monitored at $\lambda = 615 \text{ nm}$.

Next, and aiming to demonstrate the generality of our approach, we performed photo-induced allosteric experiments with bromocresol green (BG), another dye that binds to subdomain IB primarily (Figure 3.4).^[38] The binding of bromocresol green to HSA results in changes in the absorption spectrum that are comparable to those observed for methyl orange (Figure 3.4a). In the native HSA protein and thus in the absence of any spiropyran, no photo-induced spectral changes are ever observed for the HSA-BG complex (Figure 3.4b), which

indicates that BG is stable against irradiation. In the hybrid system, the binding of BG does not interfere with the photochromic properties of the spiropyran (Figure 3.4c). Upon irradiation, an increase in absorption is observed at $\lambda = 615$ nm (Figure 3.4d, blue line), which indicates release of HSA-SP bound BG that occurs upon conversion of the SP-form to the MC-form. Based on similar examination, about $(4.5 \pm 0.5)\%$ of the bound BG is released upon photo conversion of the spiropyran. The lower percentage of BG that is released is likely related to its binding affinity to HSA, which is 7-fold higher than the binding affinity of MO to HSA.^[36,39] Upon subsequent irradiation with visible light ($\lambda \geq 420$ nm), about $(89 \pm 11)\%$ of the released ligand re-binds the protein during MC to SP conversion, which means that BG displays a higher degree of reversibility than MO. The difference in the degree of reversibility possibly originates in BG displaying a naturally higher binding constant.

To further investigate the release of MO from IB pocket, we proceed to determine the binding affinity of MO to the engineered HSA by following the fluorescence of tryptophan that is quenched upon binding of MO to IB pocket (Figure 3.5a, excitation at $\lambda = 279$ nm and emission at $\lambda = 340$ nm). Notably, the only tryptophan residue in the entire HSA structure (Trp214) is located at the connecting helix between subdomain IB and subdomain IIA (Figure 3.5a, inset).^[30] The observed quenching is attributed to energy transfer phenomenon (FRET) between Trp ($\lambda_{\text{emission}} = 300 - 400$ nm) and methyl orange ($\lambda_{\text{excitation}} = 350 - 550$ nm) as the distance between Trp214 and IB binding site (3 – 4 nm) allows such mechanism to occur.^[40] Based on fitting with Stern-Volmer equation (Figure 3.5b), the protein-ligand associations in the different systems reflected by their quenching constants are summarized in Figure 3.5c. The effective quenching constants reveal a 2.7-fold decrease in MO binding upon UV light irradiation.

Although the decrease in the ligand binding affinity is up to 3-fold at best, a significant impact could still be expected as albumin is involved in complex cellular networks. Also, this result compares positively to the magnitude in photo-controlling binding affinity reported in literature, which was a 4-fold decrease of binding affinity in the catabolite activator protein.^[10,19]

Figure 3.5 MO binding to HSA-SP examined based on Trp fluorescence quenching. a) Fluorescence emission of HSA-SP titrated with MO, without any irradiation (left) and upon irradiation with UV light (right) (excitation at $\lambda = 279$ nm and emission at $\lambda = 297 - 450$ nm). Inset: the fluorescent Trp-214 (orange) in HSA. b) Experimental data-points and Stern-Volmer fitting for Trp fluorescence quenching by MO, without and with UV light irradiation (black and red lines, respectively). HSA-SP concentration is $6 \mu\text{M}$. c) Effective quenching constants of MO and the hybrid system, without and with UV light irradiation revealing a 2.7-fold decrease upon irradiation.

3.2.3 Molecular dynamic simulations

In order to shed light on ligand release mechanism, we perform molecular dynamics (MD) simulations of protein binding two MO molecules in the IB pocket using GROMACS 4.5 package^[41] and Amber99SB*-ILDN force field.^[42,43] We adapt the protonation state of amino acid residues, the salt concentration, the minimum distance between protein replicas, pressure and the temperature to correspond to the experimental setup. According to molecular docking simulation performed on an X-ray structure of HSA (PDB entry 4K2C) using Autodock 4.2,^[44] both forms of the photo-switch attached to Cys34 residue in subdomain IA can form a hydrogen bond with Arg144 and Lys137 of the IB pocket (Figure 3.6). Our MD simulations show that the nitro group of both SP and MC prefers the nearest

Arg144 than Lys137 for hydrogen bonds and therefore, we focus on investigating the effect of this non-covalent interaction to the behavior of the IB pocket in terms of affinity for the ligands.

Figure 3.6 Noncovalent interaction between the photo-switch (depicted only for MC) and a) Arg144 and b) Lys137 sidechains of the IB pocket.

From molecular docking simulations, we found that two MO ligands can be inserted in several different ways in the IB pocket. We take an example of two MO ligands which are stacked anti-parallelly (Figure 3.7a), to study the dynamics of protein-ligand system. This configuration of two MO molecules is likely the most favored as it minimizes the strong repulsion between two sulfonate groups by placing them apart from each other. Moreover, the two ligands form a π -stacking interaction with each other. There are in total three hydrogen bonds to the IB pocket established upon binding of two MO, where one interacts with Arg117 and Arg186 and the other interacts with Lys137. From our extensive simulations, we observe that the latter interaction is very important as the ligand release corresponds to a permanent loss of MO–Lys137 hydrogen bond. Compared to the simulations of unmodified/native HSA with two MO stacked similarly, we observe that the stability of MO–Lys137 hydrogen bond is significantly reduced upon attachment of photo-switch, which could be an indication of MO release.

Figure 3.7 Distance fluctuation analysis of domain I upon photo-switching. a) Configuration of two MO stacked anti-parallelly in IB pocket. b) Color coded representation of six and four-helices in the IA and IB subdomains. c) Fluctuation in the flexibility area of domain I helices upon photo-switching, blue and orange-colored area represent regions with higher flexibility in the presence of SP and MC, respectively.

To gain insight on the distinctive effect of SP and MC introduced to the IB pocket, we compute the difference in $C\alpha$ – $C\alpha$ distance fluctuation of amino acid residues of domain I based on the two MD trajectories where a hydrogen bond between the photo-switch and Arg144 is formed. This calculation is intended to pinpoint the change in the flexibility area of domain I upon photo-switching (Figure 3.7b and c). In this analysis, residues that lie on the coil structure are omitted in the calculation as they naturally fluctuate in an uncorrelated way. The orange-colored area in Figure 3.7c indicates the region of HSA that becomes more

flexible in the presence of MC, while blue color indicates a higher flexibility in the HSA-SP state. We observe that the IB pocket fluctuates more in the HSA-MC state as reflected by the predominance of orange-colored region in both the diagonal B1–B4 vs B1–B4 and the off-diagonal A1–A6 vs B1–B4 regions. The former region is related to an increased breathing of IB pocket, while the latter indicates a higher cross-fluctuation between IA and IB subdomains in the HSA-MC state. These findings altogether suggest that the IB pocket is more flexible in the HSA-MC system compared to HSA-SP system.

3.3 Conclusions

We have demonstrated a photo-responsive hybrid system of the allosteric HSA that is covalently modified with a single SP/MC switch at a specific site. Irradiation of the hybrid system results in ligand release, while the overall structure of the protein is preserved. Although the decrease in the ligand binding affinity is up to 3-fold at best, a significant impact could still be expected as albumin is involved in complex cellular networks. The photo-control of ligand binding is achieved in an allosteric manner as the switch is attached to a neighboring site (subdomain IA) of the examined ligand binding site (subdomain IB). Molecular dynamics simulations indicate that the covalent attachment of the switch in IA reduces the stability of ligand binding to key residues of IB pocket. Moreover, IB pocket evidently breathes and fluctuates more in the HSA-MC system compared to HSA-SP system. The remote control of a binding pocket by switching a neighboring site opens new opportunities for externally interfering with allosteric regulation of HSA. We envision that similar engineering strategies could be extended to systems with higher-order structures and further on to supramolecular protein-based hybrids and artificial biomimetic systems.

3.4 Acknowledgements

All molecular dynamic studies were performed by Habiburrahman Zulfikri and Prof. Claudia Filippi. We thank Dr. Jean-Wilfried Fredy and Pichayut Tananchayakul for the synthesis of the photo-switch and collaboration in ligand binding studies and Dr. Alberto Juan for assistance with NMR analysis.

3.5 Materials and methods

3.5.1 General

The human serum albumin (HSA) was purchased from Sigma-Aldrich. All experiments are performed in a Tris-HCl buffer pH 7.5 (containing 150 mM NH₄Cl). Chemicals were purchased from Sigma-Aldrich and used without further purification, unless stated otherwise.

3.5.2 Synthesis of the photo-switch

The light-responsive maleimide-spiropyran switch was synthesized according to described procedure (Figure 3.8).^[45]

Figure 3.8 Synthesis of maleimide-bearing nitrospiropyran

1-(2-hydroxyethyl)-2,3,3-trimethyl-3H-indol-1-ium iodide (3) was synthesized according to the literature method. 2,3,3-trimethyl-3H-indole **1** (4.96 g, 31.2 mmol, 1 eq) was dissolved in dry acetonitrile (10 mL) and 2-iodoethanol (2.55 mL, 32.7 mmol, 1.1 eq) was added and heated to reflux for 24 h under N₂. The reaction mixture was cooled at room temperature, the solid precipitate was filtered, washed with CH₂Cl₂ (30mL*3) to afford purple solid **3** (8.44 g, 82%). MS(ESI): 204.1 [M-H]⁺;

^1H NMR (400 MHz, CD_3CN) δ 7.79 – 7.75 (m, 1H), 7.74 – 7.70 (m, 1H), 7.66 – 7.58 (m, 2H), 4.59 – 4.48 (t, J = 5.0, 5.2 Hz, 2H), 4.00 (dd, J = 10.5, 5.5 Hz, 2H), 3.69 (t, J = 5.8 Hz, 1H), 2.79 (s, 3H), 1.57 (s, 6H);

^{13}C NMR (101 MHz, CD_3CN) δ 198.95, 142.89, 142.23, 130.81, 130.08, 124.38, 118.30, 116.48, 58.84, 55.66, 51.53, 22.83, 15.55.

9,9,9a-trimethyl-2,3,9,9a-tetrahydrooxazolo[3,2-a]indole (4) was synthesized according to the literature method. A solution of **3** (8.35 g, 25.2 mmol, 1 eq), KOH (3.21 g, 57.2 mmol, 2.3 eq) in deionized water (40 mL) was stirred for 30 min at room temperature. After extraction with diethylether (30mL*3), the organic layer was collected, dried with MgSO_4 and evaporated to afford **4** as a yellow oil.

^1H NMR (400 MHz, CD_3CN) δ 7.13 – 7.06 (m, 2H), 6.87 (m, 1H), 6.77 (d, J = 7.8 Hz, 1H), 3.75 (m, 2H), 3.52 (m, 1H), 3.40 – 3.32 (m, 1H), 1.35 (s, 3H), 1.32 (s, 3H), 1.12 (s, 3H);

^{13}C NMR (101 MHz, CD_3CN) δ 152.14, 141.00, 128.38, 123.26, 122.24, 112.79, 109.64, 63.53, 50.40, 47.54, 28.19, 20.94, 17.49.

2-(3',3'-dimethyl-6-nitrospiro[chromene-2,2'-indolin]-1'-yl)ethanol (5) was synthesized according to the literature method. A solution of 5-nitrososalicylaldehyde (2.60 g, 15.6 mmol, 1.5 eq) in ethanol (20 mL) and the obtained crude **4** (2.11 g, 10.4 mmol, 1 eq mmol) was refluxed for 3 h under N_2 . After cooling to room temperature, the mixture was filtered, washed with EtOH and dried to afford **5** as a solid purple (1.07 g, 29%).

^1H NMR (400 MHz, CD_3CN) δ 8.07 (d, J = 2.8 Hz, 1H), 8.00 (dd, J = 9.0, 2.8 Hz, 1H), 7.18 – 7.08 (m, 2H), 7.03 (d, J = 10.4 Hz, 1H), 6.86 – 6.81 (m, J = 7.5, 0.9 Hz, 1H), 6.72 (d, J = 9.0 Hz, 1H), 6.66 (d, J = 7.8 Hz, 1H), 5.99 (d, J = 10.4 Hz, 1H), 3.71 – 3.48 (m, 2H), 3.37 – 3.17 (m, 2H), 2.82 (s, 1H), 1.25 (s, 3H), 1.15 (s, 3H);

^{13}C NMR (101 MHz, CD_3CN) δ 160.44, 148.35, 142.10, 136.97, 128.81, 128.67, 126.57, 123.71, 123.26, 122.71, 120.37, 120.04, 116.25, 108.03, 107.77, 60.88, 53.49, 46.90, 26.15, 20.08.

2-(3',3'-dimethyl-6-nitrospiro[chromene-2,2'-indolin]-1'-yl)ethyl 2-bromoacetate (6) was synthesized according to the literature method. A solution of **4** (0.25 g, 0.7 mmol, 1

eq) was dissolved in dry toluene (6 mL). Pyridine (90 μ L, 1.10 mmol, 1.5 eq) and bromoacetyl bromide (100 μ L, 1.10 mmol, 1.5 eq) were added. The reaction was stirred at room temperature, under N_2 . After stirring for 3 h, water (10 mL) was added and then was extracted with dry diethylether (30mL*3). The organic layer was collected, dried over $MgSO_4$ and evaporated to afford **6** as a yellow brown (0.31 g, 91%).

1H NMR (400 MHz, CD_3CN) δ 8.08 (d, J = 2.8 Hz, 1H), 8.00 (dd, J = 9.0, 2.8 Hz, 1H), 7.20 – 7.10 (m, 2H), 7.06 (d, J = 10.4 Hz, 1H), 6.87 (m, 1H), 6.71 (t, J = 8.3 Hz, 2H), 6.01 (d, J = 10.4 Hz, 1H), 4.33 (m, 1H), 4.28 – 4.19 (m, 1H), 3.94 – 3.81 (m, 2H), 3.52 (m, 1H), 3.40 (dt, J = 15.5, 5.1 Hz, 1H), 1.25 (s, 3H), 1.13 (s, 3H);

^{13}C NMR (101 MHz, CD_3CN) δ 168.21, 160.29, 147.74, 142.21, 136.97, 129.19, 128.76, 126.65, 123.77, 122.90, 122.87, 120.78, 119.97, 116.32, 107.81, 107.72, 64.89, 53.55, 43.07, 27.33, 26.13, 19.94.

2-(3',3'-dimethyl-6-nitrospiro[chromene-2,2'-indolin]-1'-yl)ethyl 2-iodoacetate (1) was synthesized according to the literature method. A solution of **6** (0.31 g, 0.66 mmol, 1 eq) was dissolved in acetone (6 mL) and added NaI (0.12 g, 0.79 mmol, 1.2 eq). After stirring at room temperature for 2 h under Ar, the reaction mixture was filtered. Ethylacetate was added to the mother liquor and the organic phase was washed with H_2O and dried over $MgSO_4$, then evaporated and dried to afford **1** (0.23 g, 68%). MS(ESI): 521.1 $[M-H]^+$;

1H NMR (400 MHz, CD_3CN) δ 8.09 (d, J = 2.8 Hz, 1H), 8.00 (dd, J = 9.0, 2.8 Hz, 1H), 7.21 – 7.11 (m, 2H), 7.06 (d, J = 10.4 Hz, 1H), 6.87 (td, J = 7.5, 0.8 Hz, 1H), 6.72 (t, J = 8.6 Hz, 2H), 6.03 (d, J = 10.4 Hz, 1H), 4.33 – 4.24 (m, 1H), 4.23 – 4.16 (m, 1H), 3.70 – 3.67 (m, 2H), 3.55 – 3.46 (m, 1H), 3.43 – 3.34 (m, 1H), 1.25 (s, 3H), 1.14 (s, 3H);

^{13}C NMR (101 MHz, CD_3CN) δ 169.87, 160.30, 147.76, 142.20, 136.97, 129.19, 128.77, 126.65, 123.77, 122.97, 122.87, 120.76, 119.98, 116.33, 107.83, 107.74, 64.75, 53.57, 43.01, 26.19, 19.95, -4.18.

2-(3',3'-dimethyl-6-nitrospiro[chromene-2,2'-indolin]-1'-yl)ethyl 2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)acetate (2) was synthesized according to the literature method. A solution of **1** (22 mg, 0.047 mmol, 1 eq) in dry acetone (2 mL) was added maleimide (182 mg, 0.188 mmol, 4 eq) and K_2CO_3 (0.010 g, 0.071 mmol, 1.5 eq). After stirring for 4 h at

room temperature under N₂, ethylacetate was added to the reaction mixture and it was washed with 0.1 M HCl and water. The organic phase was then evaporated and dried to afford **8** (0.02 g, 87%). MS(ESI): 489.9 [M-H]⁺;

¹H NMR (400 MHz, CD₃CN) δ 8.08 (d, *J* = 2.7 Hz, 1H), 8.00 (dd, *J* = 9.0, 2.8 Hz, 1H), 7.19 – 7.11 (m, 2H), 7.06 (d, *J* = 10.3 Hz, 1H), 6.89 – 6.84 (m, 1H), 6.81 (s, 1H), 6.73 (d, *J* = 9.0 Hz, 1H), 6.69 (s, 1H), 6.65 (d, *J* = 7.8 Hz, 1H), 5.97 (d, *J* = 10.4 Hz, 1H), 4.34 (m, 1H), 4.20 (dd, *J* = 11.0, 5.6 Hz, 1H), 4.16 (s, 2H), 3.53 – 3.44 (m, 1H), 3.41 – 3.32 (m, 1H), 1.25 (s, 3H), 1.13 (s, 3H);

¹³C NMR (101 MHz, CD₃CN) δ 171.04, 168.64, 160.20, 147.74, 142.13, 136.91, 136.01, 135.52, 129.14, 128.72, 126.58, 123.70, 122.77, 122.69, 120.70, 119.87, 116.25, 107.68, 107.62, 64.54, 53.41, 43.06, 39.32, 26.09, 19.86.

3.5.3 Synthesis and characterization of the hybrid system

A 10-fold excess of nitro-spiropyran bearing a maleimide functional group was added to HSA (in 20 mM Tris-HCl buffer pH 7.5, with 20% DMSO) at room temperature and incubated for an hour. The unbound spiropyran was removed by repeated spin-filtration (Corning Spin-X UF 10k MWCO membrane filters), each for 20 min at 6,000 rpm (Eppendorf centrifuge 5415R) or with the size-exclusion chromatography (preparative column Superose 6 10/100 GL, GE Healthcare FPLC Äkta purifier 900 with a 24 mL bed volume). Prior to injecting the samples into the column, all samples were dialyzed against Tris-HCl buffer pH 7.5 to remove the DMSO using 12 – 14 kDa dialysis membranes (Spectra/Por). The purified hybrid system was further characterized with UV-visible spectrometer (Perkin Elmer Lambda 850).

Calculations for protein:photo-switch ratio based on UV-Vis spectra (Figure 3.1d):

$$A_{350} = [HSA] \times 1,167 \text{ M}^{-1}\text{cm}^{-1} \times 1 \text{ cm} + [\text{spiropyran}] \times 7,994 \text{ M}^{-1}\text{cm}^{-1} \times 1 \text{ cm}$$

$$A_{280} = [HSA] \times 30,456 \text{ M}^{-1}\text{cm}^{-1} \times 1 \text{ cm} + [\text{spiropyran}] \times 13,042 \text{ M}^{-1}\text{cm}^{-1} \times 1 \text{ cm}$$

3.5.4 Photo-responsive ligand binding analyses

A 4-fold excess of ligand (either methyl orange or bromocresol green) was added to the hybrid system (3.7 μM in 20 mM Tris-HCl buffer) at room temperature and incubated for at least 3 h. The unbound ligand was removed with repeated spin-filtration. The samples were

irradiated with UV light ($\lambda = 365$ nm, bluepoint LED Honle Technology, 40 mW/cm^2) for 5 min and subsequently with visible-light ($\lambda \geq 420$ nm, Edmund MI-150 High-intensity Illuminator) for 5 min. The absorption spectra were recorded with a Perkin Elmer Lambda 850 UV-visible spectrometer. For the fluorescence quenching studies, a 3 mL solution of the hybrid system in 20 mM Tris-HCl buffer was placed inside a cuvette with a magnetic stirring system and titrated with the ligand in a microscale addition. After each addition, the solution was incubated and stirred for 3 min. Afterwards, the emission spectra at $\lambda = 297 - 450$ nm (excitation at $\lambda = 279$ nm) were recorded using a Perkin Elmer fluorescence spectrometer. All experiments were reproduced three times.

Calculations for ligand binding/release (calibration curves shown in Figure 3.9):

$$[dye]_{total} = \frac{A_{before\ irradiation}}{\epsilon_{bound} \times 1\text{ cm}}$$

$$[dye]_{bound} \times \epsilon_{bound} \times 1\text{ cm} + [dye]_{free} \times \epsilon_{free} \times 1\text{ cm} = A_{after\ irradiation}$$

$$[dye]_{total} = [dye]_{bound} + [dye]_{free}$$

$$\%release = \frac{[dye]_{free}}{[dye]_{total}} \times 100\%$$

Figure 3.9 Calibration curves of a) free MO, b) HSA-SP-bound MO, c) free BG and d) HSA-SP-bound BG. All experiments are performed in Tris-HCl buffer pH 7.5.

3.5.5 PFGSE-NMR experiments

All NMR experiments were measured using a Bruker Avance II NMR 600 MHz spectrometer, equipped with a triple-nucleus TXI probe head with a z-gradient coil, and a Great 3/10 gradient amplifier. All experiments were performed at 300 K using standard pulse sequences from the Bruker library. Pulsed-Field Gradient Spin-Echo (PFGSE) NMR experiments were carried out by using the pulse sequence of longitudinal-eddy-current delay (LED) with bipolar-gradient pulses, with a diffusion delay (Δ) of 0.2 s and gradient length (δ) of 1200 μ s and 5 ms of eddy-current delay. The gradient strength was varied from 5 to 95%. A total of 128 transients were collected for each of the 20 increment steps with a recycle delay of 2 s. Pseudo-2D DOSY plots were processed with the standard Bruker software; the diffusion coefficients were determined by the T1/T2 “vargrad” SimFit fitting routine. Calibration of the field gradient strength was achieved by measuring the value of translational diffusion coefficient (D) for the residual ^1H signal in D_2O (99.99%, ^2H atom), $D = 1.91 \times 10^{-9} \text{ m}^2/\text{s}$.

3.5.6 Determination of effective quenching constants

The formation of a protein-ligand complex is associated with static quenching of protein fluorescence by the bound ligand. The quenching constant is obtained from the Stern-Volmer equation:

$$\frac{F_0}{\Delta F} = \frac{1}{f_a \cdot K_a \cdot [Q]} + \frac{1}{f_a}$$

where ΔF is the difference in fluorescence emission with and without the quencher Q in the system; Q represents the free quencher; f_a represents the fraction of accessible fluorescence and K_a is the effective quenching constant.^[46,47]

Isothermal microcalorimetry is typically used to determine binding constants experimentally. However, since the MC-form can be converted back to SP-form during the automated measurement, this technique is not applicable here.

3.5.7 Analyses of protein global structure

The conformation of the protein (0.1 μM) before and after irradiation in 10 mM phosphate buffer pH 7.5 was analyzed with a Jasco J-1500 circular dichroism spectrometer. The

hydrodynamic size of the protein before and after irradiation in 20 mM Tris-HCl buffer pH 7.5 was determined with dynamic light scattering (Nanotracc Wave, Microtracc).

3.5.8 Data and schematic representation

All data plotting and mathematical fitting were performed with OriginPro 9.0 software. For CD spectra, the raw data were smoothed using the Savitzky-Golay filter in OriginPro 9.0 software. Protein structures are rendered using PyMOL 1.3 software and chemical structures are drawn using ChemBioDraw Ultra 12.0 software.

3.6 References

- [1] H. N. Motlagh, J. O. Wrabl, J. Li, V. J. Hilser, *Nature* **2014**, *508*, 331-339.
- [2] A. W. Fenton, *Trends Biochem. Sci.* **2008**, *33*, 420-425.
- [3] D. Kern, E. R. P. Zuiderweg, *Curr. Opin. Struc. Biol.* **2003**, *13*, 748-757.
- [4] C.-J. Tsai, A. del Sol, R. Nussinov, *J. Mol. Biol.* **2008**, *378*, 1-11.
- [5] A. C. M. Ferreón, J. C. Ferreón, P. E. Wright, A. A. Deniz, *Nature* **2013**, *498*, 390-394.
- [6] J.-P. Changeux, *Nat. Rev. Mol. Cell. Biol.* **2013**, *14*, 819-829.
- [7] J. H. Choi, A. H. Laurent, V. J. Hilser, M. Ostermeier, *Nat. Commun.* **2015**, *6*, 6968.
- [8] R. A. Brown, V. Diemer, S. J. Webb, J. Clayden, *Nat. Chem.* **2013**, *5*, 853-860.
- [9] G. Y. Tonga, Y. Jeong, B. Duncan, T. Mizuhara, R. Mout, R. Das, S. T. Kim, Y.-C. Yeh, B. Yan, S. Hou, V. M. Rotello, *Nat. Chem.* **2015**, *7*, 597-603.
- [10] W. Szymański, J. M. Beierle, H. A. V. Kistemaker, W. A. Velema, B. L. Feringa, *Chem. Rev.* **2013**, *113*, 6114-6178.
- [11] A. A. Beharry, G. A. Woolley, *Chem. Soc. Rev.* **2011**, *40*, 4422-4437.
- [12] R. Klajn, *Chem. Soc. Rev.* **2014**, *43*, 148-184.
- [13] D. P. Nguyen, M. Mahesh, S. J. Elsässer, S. M. Hancock, C. Uttamapinant, J. W. Chin, *J. Am. Chem. Soc.* **2014**, *136*, 2240-2243.
- [14] J. Broichhagen, D. Trauner, *Curr. Opin. Chem. Biol.* **2014**, *21*, 121-127.
- [15] J. R. Kumita, O. S. Smart, G. A. Woolley, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3803-3808.
- [16] G. A. Woolley, *Acc. Chem. Res.* **2005**, *38*, 486-493.

- [17] B. Schierling, A.-J. Noël, W. Wende, L. T. Hien, E. Volkov, E. Kubareva, T. Oretskaya, M. Kokkinidis, A. Römpf, B. Spengler, A. Pingoud, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1361-1366.
- [18] M. Aizawa, K. Namba, S. Suzuki, *Arch. Biochem. Biophys.* **1977**, *182*, 305-310.
- [19] M. Bose, D. Groff, J. Xie, E. Brustad, P. G. Schultz, *J. Am. Chem. Soc.* **2006**, *128*, 388-389.
- [20] G. A. Woolley, A. S. I. Jaikaran, M. Berezovski, J. P. Calarco, S. N. Krylov, O. S. Smart, J. R. Kumita, *Biochemistry* **2006**, *45*, 6075-6084.
- [21] A. Koçer, M. Walko, W. Meijberg, B. L. Feringa, *Science* **2005**, *309*, 755-758.
- [22] M. R. Banghart, M. Volgraf, D. Trauner, *Biochemistry* **2006**, *45*, 15129-15141.
- [23] I. Tochitsky, M. R. Banghart, A. Mourrot, J. Z. Yao, B. Gaub, R. H. Kramer, D. Trauner, *Nat. Chem.* **2012**, *4*, 105-111.
- [24] S. Pittolo, X. Gómez-Santacana, K. Eckelt, X. Rovira, J. Dalton, C. Goudet, J.-P. Pin, A. Llobet, J. Giraldo, A. Llebaria, P. Gorostiza, *Nat. Chem. Biol.* **2014**, *10*, 813-815.
- [25] M. Volgraf, P. Gorostiza, S. Szobota, M. R. Helix, E. Y. Isacoff, D. Trauner, *J. Am. Chem. Soc.* **2007**, *129*, 260-261.
- [26] M. Schönberger, D. Trauner, *Angew. Chem. Int. Ed.* **2014**, *53*, 3264-3267.
- [27] M. Volgraf, P. Gorostiza, R. Numano, R. H. Kramer, E. Y. Isacoff, D. Trauner, *Nat. Chem. Biol.* **2006**, *2*, 47-52.
- [28] D. Strickland, K. Moffat, T. R. Sosnick, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10709-10714.
- [29] P. Ascenzi, M. Fasano, *Biophys. Chem.* **2010**, *148*, 16-22.
- [30] G. Fanali, A. di Masi, V. Trezza, M. Marino, M. Fasano, P. Ascenzi, *Mol. Aspects Med.* **2012**, *33*, 209-290.
- [31] A. S. Pitek, S. A. Jameson, F. A. Veliz, S. Shukla, N. F. Steinmetz, *Biomaterials* **2016**, *89*, 89-97.
- [32] G. F. Li, D. Magana, R. B. Dyer, *Nat. Commun.* **2014**, *5*, 3100.
- [33] M. N. Khan, *J. Pharm. Sci.* **1984**, *73*, 1767-1771.
- [34] M. J. Davies, *Biochem. Biophys. Res. Commun.* **2003**, *305*, 761-770.
- [35] F. Zsila, *Mol. Pharm.* **2013**, *10*, 1668-1682.

- [36] R. Ambrosetti, R. Bianchini, S. Fisichella, M. Fichera, M. Zandomenoghi, *Chem. Eur. J.* **1996**, 2, 149-156.
- [37] F. L. Rodkey, *Clin. Chem.* **1965**, 11, 478-487.
- [38] J. W. Keyser, *Clin. Chim. Acta* **1961**, 6, 445-447.
- [39] V. D. Trivedi, I. Saxena, M. U. Siddiqui, M. A. Qasim, *IUBMB Life* **1997**, 43, 1-8.
- [40] P. C. Ray, Z. Fan, R. A. Crouch, S. S. Sinhaa, A. Pramanika, *Chem. Soc. Rev.* **2014**, 43, 6370-6404.
- [41] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* **2005**, 26, 1701-1718.
- [42] R. B. Best, G. J. Hummer, *Phys. Chem. B.* **2009**, 113, 9004-9015.
- [43] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, *Proteins* **2010**, 78, 1950-1958.
- [44] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, 16, 2785-2791.
- [45] M. -Q. Zhu, T. Chen, G. -F. Zhang, C. Li, W. -L. Gong, Z. -Q. Chena, M. P. Aldreda, *Chem. Commun.* **2014**, 50, 2664-2666.
- [46] Y. -J. Hu, Y. Liu, X. -H. Xiao, *Biomacromolecules* **2009**, 10, 517-521.
- [47] M. van de Weert, *J. Fluoresc.* **2010**, 20, 625-629.

Chapter 4

Bacterial encapsulins as *in vitro* and *in vivo* protein-based nanoplatfoms

Bacterial encapsulins are highly organized structures that are composed of protein building blocks, making up icosahedral cage-like structures with the size of 20 – 25 nm. The outer and inner surface of encapsulins can be modified, either chemically or genetically, and the internal cavity can be used to template, store and arrange molecular cargo within a defined space. Recent years have witnessed the increasing interest on the application of bacterial semi-organelle encapsulins for bio-nanotechnology. Here, we characterize the native structure of B. linens encapsulins with both native and foreign cargo using cryo-electron microscopy (cryo-EM). Furthermore, by harnessing the confined enzyme (i.e., a peroxidase), we demonstrate the functionality of the encapsulin for an in vitro surface-immobilized catalysis in a cascade pathway with an additional enzyme, glucose oxidase. We also demonstrate the in vivo functionality of the encapsulin for cellular uptake using mammalian macrophages. Unraveling both the structure and functionality of the encapsulins paves the way towards transforming biological nanocompartments into functional systems.

Part of this chapter is published in R.M. Putri, J.J.L.M. Cornelissen, M.S.T. Koay, *ChemPhysChem* **2015**, 16, 911-918. Part of this chapter is included in a submission to *ACS Nano* **2017**.

4.1 Introduction

Protein-based nanocages have gained great interest in nanotechnology and have been developed into effective delivery agents (nanocarriers) and nanoreactors.^[1-3] In recent years, a new class of nanocages from bacteria called the encapsulins (Figure 4.1) has emerged as promising nanoplateforms, characterized by their robust nature.^[4-9] Nevertheless, compared to other protein-based nanocages, such as viral capsids,^[1,10] ferritins^[11,12] and heat-shock proteins,^[13] the development of encapsulins for nanoplateforms is only in its infancy.^[6] Previous studies highlighted that the encapsulins confine a functional protein-based cargo, such as an enzyme or a ferritin-like protein, thus being recognized as a bacterial semi-organelle. The encapsulin from *Brevibacterium linens*, for instance, naturally houses a dye-decolorizing peroxidase (DyP) within its cavity, which is involved in oxidative stress. DyP is assembled as a trimer of dimers, i.e., a 240 kDa hexamer. The specific encapsulation mechanism of DyP is mediated by its C-terminal end, which interacts specifically with a defined region of the encapsulin inner surface. We recently demonstrated that fusion of the DyP C-terminal end to a heterologous protein, such as the teal fluorescent protein (TFP), allows its packaging in these nanocontainers.^[14]

Despite the recent sharp increase in the number of studies on encapsulins,^[4,5,7-9,14-20] structural characterization of encapsulins at the molecular level is yet to be explored. Encapsulins from *Myxococcus xanthus* (PDB entry 4PT2)^[17] and *Pyrococcus furiosus* (PDB entry 2E0Z)^[21] assemble into T = 3 (180-subunit, ~32 nm) nanocompartments, whereas encapsulin from *Thermotoga maritima* (PDB entry 3DKT)^[6] assembles into a T = 1 (60-subunit, 24 nm) particle (Figure 4.1). Both icosahedral particles pack ferritin-like proteins that protect cells from oxidative stress.^[6,17] Notably, encapsulin fold is similar to the capsid protein fold of Hong Kong 97 (HK97)-like virions^[22] (including dsDNA bacteriophages and herpesviruses), which constitute the most successful self-replicating system on Earth. Thus, encapsulins and the capsid proteins of HK97-like viruses might share a common evolutionary origin. The structural characteristics at the molecular level of encapsulins that confine enzymes is still poorly understood. Deciphering the structure would be the first step towards understanding the enzyme confinement in bacteria and the functioning of encapsulins as semi-organelles for biochemical reactions. To characterize the protein structure, as an

alternative to the well-known protein crystallization, cryo-electron microscopy (cryo-EM) recently emerges as a powerful technique.

Figure 4.1 Structure of encapsulin from the bacterium *T. maritima* (PDB entry 3DKT).^[6] The protein cage is composed of 12 pentamers, whereby the monomer (blue) and one pentamer (red) are shown. The protein cargo (yellow) occupies the internal cavity as a result of interactions with a conserved peptide sequence (purple).

In this chapter, the structural characterization of *B. linens* encapsulin with native DyP cargo (DyP-E) and foreign TFP cargo (TFP-E) is presented using cryo-EM reconstruction. Furthermore, the functionality of *B. linens* encapsulins as nanoplateforms for *in vitro* and *in vivo* purposes is demonstrated. We harness the confined DyP in the native encapsulin to perform a tandem catalysis *in vitro* on a surface in a cascade manner. For the *in vivo* study, we use the modified encapsulin with a TFP cargo to investigate the particle uptake by murine macrophages.

4.2 Results and discussion

4.2.1 Characterization of encapsulin stability and structure

The encapsulin from *B. linens* is recombinantly expressed in *E. coli* and mechanical lysis (i.e., ultrasonication) is performed to disrupt the cell and retrieve the encapsulin particles. The particles are purified with 38% sucrose cushion and 10 – 50% sucrose gradient to remove other cellular proteins. The final step of purification involves a size-exclusion chromatography (SEC) that indicates the presence of the encapsulin peak at $V = 11 - 12$ mL

(Figure 4.2A, red line) monitored at $\lambda = 280$ nm in the native environment (Tris-HCl buffer pH 7.5).

Figure 4.2 Characterization of encapsulin at different pH values. (A) Size-exclusion profiles of encapsulin at acidic, native and basic pH monitored at $\lambda = 280$ nm, revealing a single peak of encapsulin ($V = 11 - 12$ mL or $V = 9$ mL) for each profile. (B) Size-exclusion profiles of encapsulin at native and highly basic pH, showing an extra peak (number 3) at $V = 15$ mL apart from the typical encapsulin peak (number 1 and 2). (C) DLS-based size distributions of encapsulin at acidic, native and basic pH. (D) TEM images of encapsulin at acidic, native and basic pH, revealing intact spherical particles under all conditions. (E) Denaturing gel electrophoresis revealing the encapsulin band (~ 28 kDa) present for all samples. Lanes 1, 2, 3 correspond to peaks 1, 2, 3 in Figure 4.2B, respectively, lane 4 corresponds to the protein marker.

To investigate the robust nature of *B. linens* encapsulin,^[20] we characterize the stability of native DyP-E against pH changes and higher ionic strength. The stability of encapsulin particles in 7 different pH values is examined (i.e., acidic condition at pH 3 – 5, native condition at pH 7.5, basic condition at pH 9 – 11). Acetate buffer is used to prepare pH 3, pH 4 and pH 5 buffers, Tris-HCl buffer is used for pH 7.5 and pH 9 buffers, and phosphate buffer is used for pH 10 and pH 11 buffers. For each pH value, the encapsulin samples are incubated overnight and subsequently characterized using SEC, dynamic light scattering (DLS) and transmission electron microscopy (TEM), respectively. Upon mixing and overnight incubation of encapsulin at pH 3 and pH 4 buffers, particle aggregation is suspected to occur

as the solutions become cloudy. Centrifugation of the solutions results in white pellets and therefore these samples are not further analyzed.

Apart from pH 3 and pH 4, the encapsulin particles seem to be stable at all tested pH values, which is in agreement with a previous report.^[23] Characterization with SEC (Figure 4.2A and B), DLS (Figure 4.2C) and TEM (Figure 4.2D) respectively shows that uniform and spherical structures could be observed in all samples. Notably, the elution profiles from SEC are different at basic pH (pH 9 and 10): instead of eluting at $V = 11 - 12$ mL (characteristic for native encapsulin particles), the encapsulins incubated at higher pH elute at a lower volume ($V = 9$ mL), suggesting an increase in particle size, possibly due to aggregation. TEM analysis (Figure 4.2D) shows that typical spherical particles are still present at basic pH although a tendency to aggregate is indeed observed. Aggregation of particles is clearly seen when the incubation at elevated pH is prolonged to 2 days (Figure 4.2C), which could be partially reversed by adjusting the pH back to around neutral. Furthermore, we also note that a second peak at $V = 15$ mL appears in the case of encapsulin at pH 11 (Figure 4.2B), indicating the formation of smaller particles as it elutes at a higher volume than intact particles ($V = 11 - 12$ mL). This phenomenon is not observed at lower pH values. The species eluting at $V = 15$ mL is confirmed to be encapsulin-based by SDS-PAGE (Figure 4.2E, red arrow) since the protein band appears at the same position as the band of native encapsulin at pH 7.5. Hence, it is concluded that native *B. linens* encapsulin partially disassembles at pH 11. Similar particle disintegration at elevated pH has been reported for *T. maritima* encapsulin.^[15]

To investigate the effect of increasing ionic strength, DyP-E at native conditions (pH 7.5) and elevated pH (pH 9) are incubated with either 1 M NaCl, 1 M MgCl₂ or 1 M CaCl₂ in the solution. At both pH 7.5 and pH 9, the encapsulin particles elute at a similar elution volume compared to the samples with no salt added (Figure 4.3). This finding indicates that the increase in ionic strength (up to the tested range) does not significantly compromise the stability of encapsulin particles. This is in stark contrast, for instance, with the Cowpea Chlorotic Mottle plant virus (CCMV) that already disassembles when the ionic strength is increased to ~1 M at an elevated pH.^[24]

Figure 4.3 Characterization of encapsulin at different ionic strength. (a) Size-exclusion profiles of native encapsulin at native pH (pH 7.5) in the presence of different salts monitored at $\lambda = 280$ nm, revealing a single peak of encapsulin ($V = 10 - 12$ mL) for each profile. (b) Size-exclusion profiles of native encapsulin at elevated pH (pH 9) in the presence of different salts monitored at $\lambda = 280$ nm, revealing a single peak of encapsulin ($V = 9$ mL) for each profile.

Figure 4.4 Characterization of encapsulin in the presence of organic solvent. TEM images of native encapsulin (size around 20 nm) in the presence of (a) DMSO (left: 20% right: 40%) and (b) ethanol (left: 20% right: 40%). All samples are in Tris-HCl buffer pH 7.5.

As for several (nano)technological applications non-aqueous conditions are required, the effect of addition of organic solvents is examined. Here, we choose two most common organic solvents added to protein solutions (usually for chemical modifications),^[25,26] which are ethanol and DMSO. Based on characterization with TEM, we could still observe the presence of intact, spherical particles upon the addition of up to 40% DMSO to the encapsulin solution at pH 7.5 (Figure 4.4). However, the number of intact particles is noticeably low and most of them are already disassembled/degraded into smaller fragments, even already in the presence of 20% DMSO. Similarly, addition of 40% ethanol also results in predominantly degraded fragments and only a very small amount of intact particles could be observed (Figure 4.4). On the contrary, addition of up to 20% ethanol does not seem to affect the structures as they are highly comparable to the native encapsulin structures (Figure 4.4).

We use 3D cryo-EM to analyze the native structures of encapsulin from *B. linens* loaded with its natural cargo, DyP (DyP-E), or with TFP (TFP-E). Non-loaded encapsulin (nl-E) is included as a control (Figure 4.5A-C). Whereas DyP- and TFP-loaded encapsulin particles are homogeneous, with an internal density corresponding to the cargo, nl-E shows two types of particles, either empty or containing some material in the internal cavity at a 1:5 ratio (Figure 4.5A, white and black arrows, respectively). The presence of material inside nl-E is attributed to random/statistical packaging of cellular components that might occur during *in vivo* assembly; such phenomenon is also indicated by previously reported broad distribution of masses of nl-E particles analyzed by native mass spectrometry.^[14,20] A 3D reconstruction (3DR) is obtained for each of the three sets of particles (Figure 4.5D-F). Based on a 0.5 Fourier shell correlation (FSC) threshold, the resolution is 11.4 Å for nl-E, and 13.5 and 15 Å for TFP-E and DyP-E, respectively. These values suggest that structural integrity is better preserved in nl-E than in cargo-loaded encapsulins, in accordance with previous biophysical analyses.^[20] Our biochemical analysis further indicates that nl-E particles do not undergo partial disassembly even at elevated pH (i.e., a tendency to aggregate is observed at pH 11 instead of forming smaller particles, Figure 4.6). The outer diameter, determined from spherically averaged radial density plots of the 3DR, is 228 Å for nl-E and TFP-E, but DyP-E measured 232 Å. Despite this ~2% size increase, the inner diameter is identical for the three maps. Encapsulin maps are icosahedral shells with an average thickness of 28 – 30 Å, and the particle is based on a T=1 lattice.

Figure 4.5 Cryo-EM images and 3DR of non-loaded, DyP- and TFP-loaded encapsulin from *B. linens*. (A–C) Cryo-electron micrographs of purified non-loaded encapsulin (nl-E) (A), DyP-loaded encapsulin (DyP-E) (B), and TFP-loaded encapsulin (TFP-E) (C). In A, white and black arrows indicate empty nl-E and nl-E with packed material, respectively. Bar, 500 Å. (D–F) 3DR with icosahedral symmetry of nl-E (D), DyP-E (E) and TFP-E (F) viewed along a twofold axis of symmetry contoured at 2 σ above the mean density. Arrows indicate pores on the outer surface. (G–I) 3DR with no symmetry of nl-E (G), DyP-E (H) and TFP-E (I). Top row, the radially color-coded outer surfaces viewed along a twofold axis of symmetry contoured at 2 σ above the mean density; bottom row, 3DR with the front half of the protein shell removed. Cargo is yellow (nl-E), red (DyP-E) or green (TFP-E).

Figure 4.6 Characterization of empty/non-loaded encapsulin. (a) Denaturing gel electrophoresis revealing the encapsulin band (~ 28 kDa, green arrow) and the cargo DyP band (~ 50 kDa, red arrow). Lane 1 corresponds to protein marker and lane 2 and 3 correspond to empty/non-loaded (nl-E) and native (DyP) particles, respectively. (b) TEM image of nl-E at pH 7.5. (c) Size-exclusion profiles of nl-E at pH 7.5 and pH 11, revealing only a single peak of encapsulin ($V = 10 - 11$ mL).

At the moderate resolutions achieved, some features are distinguished; the capsid is formed by 12 slightly outward-protruding pentamers, each consisting of five curved, elongated subunits. Particles have $\sim 5 - 8$ Å diameter holes that extend through the capsid. These pores are located around the threefold axes (Figure 4.5D-F, arrows), and might serve as channels through which solutes can be interchanged with the cytoplasm. To infer cargo organization, 3DR of DyP-E, TFP-E and nl-E particles are calculated without imposing icosahedral symmetry to avoid smearing cargo features (Figure 4.5G-I). Although the three maps show smooth topography (average resolution ~ 24 Å), the cargo density exhibits distinct features. Whereas DyP-related density is connected to the inner surface of DyP-E at one three-fold axis, TFP-related density, which occupies a large part of the inner cavity, is clearly connected at six locations, in addition to other possible connections at six other locations. The internal density of nl-E capsids is much smaller, ~ 30 and 15% less compared to those of DyP-E and TFP-E.

Encapsulin maps for *B. linens* are similar to the map for the *T. maritima* T=1 encapsulin particle ($230 - 240$ Å in diameter). To evaluate this structural resemblance, and considering a similar encapsulin fold, we perform docking analysis of the *T. maritima* encapsulin crystallographic model (PDB entry 3DKT) into the cryo-EM density maps of non-loaded and loaded T = 1 encapsulins. A total of 60 monomers of *T. maritima* encapsulin accounts well for the protein shell of *B. linens* nl-E (Figure 4.7A). This model includes the C-terminal

extension of the packaged protein, a ferritin-like protein (Flp) from *T. maritima* (Figure 4.7A, yellow). In addition, most pores in the *T. maritima* encapsulin shell co-localize with the pores in *B. linens* nl-E (Figure 4.7B, dashed circles).

Figure 4.7 Pseudo-atomic models of nl-E, DyP-E and TFP-E. (A) nl-E inner surface viewed along a threefold axis, with docked *T. maritima* encapsulin atomic coordinates. Encapsulin monomers at the threefold axis are depicted in red, green and blue, and their corresponding Flp C-terminal ends in yellow. (B) nl-E outer surface viewed along a threefold axis. There are numerous pores at and around the threefold axis (dashed circles). The cargo density has been removed computationally. Black symbols indicate icosahedral symmetry axes. (C) DyP-E map with C3 symmetry from inside, with docked *T. maritima* encapsulin (blue) and DyP density (red). (D) DyP cryo-EM density extracted from the map shown in C; arrows indicate connections to the inner encapsulin surface (left). DyP density calculated from negative staining electron microscopy^[6] (right). (E) TFP-E map from inside, with docked *T. maritima* encapsulin (blue) and TFP density (green). The TFP atomic structure (PDB entry 2HQB) is shown at the same scale (top right).

The role of the DyP C-terminal extension is, as it is for Flp, to target the cargo protein to the binding sites on the encapsulin interior. Considering that DyP is organized as a trimer of dimers, we calculate a 3DR of DyP-E with C3 symmetry (Figure 4.7C). The map resolution is slightly improved compared to the map calculated with no symmetry (25 Å vs 28 Å). Furthermore, DyP is solved as a density with three connections that match with the location of the Flp C-terminal extensions (Figure 4.7D, left). The *B. linens* DyP, expressed in the

absence of encapsulin, is used to calculate a map from negatively stained particles (Figure 4.7D, right, EMD entry 1530).^[6] This recombinantly expressed DyP is similar to our native encapsulated DyP model. The cryo-EM hexameric DyP map highlights the structural differences between the two trimers; the trimer connected to the encapsulin capsid interior probably has ordered C-terminal extensions, whereas the other three C-terminal extensions are probably disordered. The cryo-EM packed hexameric DyP is observed as an elongated density (~ 114 Å long x ~ 92 Å wide; $\sim 3.3 \times 10^5$ Å³ volume), compatible with the dimensions and volume of a recombinant hexameric DyP ($\sim 104 \times 104$ Å; 2.85×10^5 Å³). The cryo-EM TFP-related density is $\sim 5.1 \times 10^5$ Å³ (Figure 4.7E); a TFP monomer (PDB entry 2HQK; a TFP derived from *Clavularia* sp.)^[27] is 4.7×10^4 Å³, which suggests that 10 – 12 TFP copies/particle are encapsulated, compatible with previous mass spectrometry analysis.^[14] Despite their similar molecular weights, the packing efficiency of TFP is greater than that of DyP, suggesting that the quaternary structure of the cargo is a determining factor for packaging.

4.2.2 *In vitro* study of encapsulin as nanoreactors on surface

To demonstrate the functionality of encapsulins for *in vitro* applications, we investigate the performance of native DyP-E as a bio-nanoreactor. The confined DyP inside the encapsulin cavity allows the particles to catalytically oxidize dye molecules, for example 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) or ABTS dye, while reducing hydrogen peroxide into water.^[19,28] We immobilize the encapsulin particles on a glass surface in order to ease the particle recovery as well as to demonstrate the catalysis on a surface, which is commonly preferred for *in vitro* application such as protein-based sensors.^[29,30] The glass surface is first modified with pentafluorophenyl silicate (Figure 4.8A) that introduces surface reactivity and allows chemical reaction with lysine residues present on the encapsulin surface.^[31] We have reported that the lysine residues of *B. linens* encapsulin are indeed available for chemical modification.^[18]

To confirm that the particles are covalently immobilized onto the surface using this approach, we perform a contact angle measurement on the encapsulin-modified glass surface, which results in the contact angle of 55 – 60° as opposed to 75 – 85° for PFPS-modified glass surface (before adding the encapsulin). The decrease in the contact angle after encapsulin

addition indicates a successful immobilization as the presence of encapsulin changes the surface properties into a more polar state (in comparison to the surface with PFPS only). Furthermore, as shown in Figure 4.8B, we are able to visualize the encapsulin particles on the surface using scanning electron microscopy (SEM). Based on multiple images recorded with SEM, we estimate the percentage of coverage area to be $35 \pm 5\%$ with up to 1.16×10^{11} encapsulin particles immobilized on the surface (area: $1 \times 1.5 \text{ cm}^2$).

Figure 4.8 Surface-immobilized encapsulin as a bio-nanoreactor. (A) Immobilization strategy of encapsulin particles on a glass surface using PFPS molecules as linkers. (B) SEM image of surface-immobilized encapsulin showing the presence of nanometer-sized particles (~24 nm) on the surface. (C) Schematic representation of a tandem system consisting of GOx and DyP-E. The initial substrate glucose is oxidized by GOx while producing H_2O_2 , which in turn is converted by Dyp into water while producing green-colored ABTS radical. (D) Catalytic assay of the tandem system monitored at $\lambda = 410 \text{ nm}$ based on the absorption of the ABTS radical cation.

After surface attachment, we investigate the catalytic activity by carrying out a cascade reaction on the surface. The encapsulin enzymatic activity is tested based on the oxidation of ABTS dye while reducing H_2O_2 produced by an additional enzyme called glucose oxidase (GOx) that is added into the solution (Figure 4.8C). We choose the well-studied tandem GOx-peroxidase system as it has been proven useful as a biosensor for glucose detection,^[28,32,33] thus providing our system with a possible functionality. The oxidation of ABTS into its green-colored radical form is monitored at $\lambda = 410 \text{ nm}$ for up to 6 min (Figure 4.8D) at

different glucose concentrations. In our experiments, GOx (1 μM) is added as the final compound into the reaction mixture as it kick-starts the entire cascade reaction. Noticeably, the Dyp-catalyzed formation of ABTS-based radical does not start immediately after GOx addition ($t = 0$, Figure 4.8D), which is attributed to GOx catalyzed H_2O_2 production and diffusion to the surface-immobilized, encapsulated DyP. As expected, this lag phase is shortened by increasing the concentration of glucose as more H_2O_2 is produced with increasing glucose concentration. In a similar way, the lag phase decreases when the surface-immobilized encapsulin is added into a pre-mixed glucose-GOx solution to allow GOx-catalyzed reactions with H_2O_2 already present (Figure 4.9), which does not change the reaction rate.

Figure 4.9 Different order of addition resulting in different lag phases. At a fixed concentration of all substrates and enzymes, the lag phase is longer if the substrate glucose is added as the final component to the system (black line, reaction starts at $t = 5$ min). On the contrary, if the surface-immobilized encapsulin is added to the mixture of GOx-glucose, the lag phase is significantly shortened (red line).

Based on the maximum slope ($\max dA_{410}/dt$), we calculate the reaction rates (i.e., the rate of the ABTS radical cation being formed) to be 2.67 $\mu\text{M}/\text{min}$ for 0.5 mM glucose, 1.57 $\mu\text{M}/\text{min}$ for 3.2 mM glucose and 1.83 $\mu\text{M}/\text{min}$ for 5 mM glucose. Based on these results, we assume that 0.5 mM glucose is already sufficient to achieve the maximum reaction rate on the surface (i.e., $V_{\max} = 2.67 \mu\text{M}/\text{min}$) since increasing glucose concentrations does not lead to a higher reaction rate and likely results in substrate inhibition instead.^[34] Considering the number of the encapsulins immobilized (1.16×10^{11} particles), we calculate the apparent

turnover number of ($V_{\max}$ /total protein concentration) to be $k_{\text{cat}} = 115.4 \text{ s}^{-1}$ (within the reaction volume of 500 μL , total protein concentration = $3.86 \times 10^{-10} \text{ M}$). The obtained turnover number is comparable in magnitude to reported turnover number for (non-encapsulated) DyP in literature (i.e., $k_{\text{cat}} = 224 \text{ s}^{-1}$),^[35] indicating that the surface immobilization does not lead to a significant activity loss. We have demonstrated here the catalytic activity of surface-immobilized encapsulin *in vitro*, paving the way towards the use of *B. linens* encapsulins as bio-nanoreactors and/or biosensors.

4.2.3 *In vivo* study of encapsulin as an agent for cellular infection

To evaluate the use of *B. linens* encapsulins for *in vivo* applications, we investigate the uptake of the encapsulin with a fluorescent protein cargo (TFP-E) by mammalian J774 macrophages. A more detailed biochemical analysis of TFP-E particles has been reported in a previous work by our group.^[14] Notably, the modified particles remain intact and stable over two years of storage at 4°C (Figure 4.10), further confirming their robust nature even with a non-native cargo.

Figure 4.10 Characterization of modified encapsulin with a fluorescent protein TFP cargo (TFP-E). Size distribution of TFP-E particles over two years of storage at 4°C, revealing the presence of intact particles with the expected size of around 20 nm.

After the addition of TFP-E to the macrophage cells and subsequent incubation for 4 h, images are taken with a fluorescent microscope to visualize the TFP cargo (excitation at $\lambda = 460 - 490 \text{ nm}$, emission at $\lambda = 525 \text{ nm}$). As the negative control, cells that are not treated

with any encapsulin particles remain present after the removal of culture medium but do not show any fluorescence signal (Figure 4.11A and B). For both lower and higher concentration of TFP-E (i.e., 0.013 μM and 0.06 μM), the treated macrophages show a high fluorescent signal as shown in Figure 4.11C-H. The green-colored signal that resembles the circular shape of the cells indicates that the TFP cargo is taken up by the macrophages, confirming an uptake of TFP-E particles. There is no visible difference between the higher and lower concentration of TFP-E and the resulting uptake at the examined concentrations. Based on the fluorescence signal from the cells, TFP-E appears to enter the macrophages effectively yet stays clear of the nucleus.

Figure 4.11 Fluorescence microscopy images of J774 macrophages treated with TFP-E. (A-B) The negative control consisting of macrophages that are not treated with TFP-E does not show any fluorescence signal. (C-E) Macrophages treated with a lower concentration of TFP-E (0.013 μM) showing fluorescence from TFP cargo protein taken up by the cells. (F-H) Macrophages treated with a higher concentration of TFP-E (0.06 μM) showing fluorescence from TFP cargo protein taken up by the cells.

This study shows that the encapsulins can be used to confine a foreign protein cargo and preserve their structural integrity for a prolonged time *in vitro*, and the cargo can be taken up by mammalian cells. Further studies are needed to identify whether the spherical particles stay intact inside the cells as well as the uptake mechanism.

4.3 Conclusion

We have provided structural insight into the encapsulins from *B. linens* using cryo-EM, confirming that the native icosahedral, cage-like assemblies confine a single peroxidase DyP within its cavity. Using the same analysis, we also characterize the structure of modified encapsulin particles with non-native cargo (i.e., a fluorescent protein). Compared to the empty encapsulin (nl-E) the protein cages that contain a cargo are slightly larger, which is in line with our previous analysis.^[20] The data further confirms the directive role the C-terminal fragments on the cargo proteins play in their inclusion,^[6,14] allowing the design and preparation of encapsulin containers with a variety of guest systems. To investigate the functionality of encapsulins as a promising agent for nanoplateforms, we demonstrate a tandem *in vitro* catalysis proving that the encapsulin is able to play a role as bio-nanoreactors on surface as well as the *in vivo* cellular uptake of modified encapsulins by mammalian macrophages. Our results provide both functional and structural basis of *B. linens* encapsulins, paving the way towards the development of the encapsulins as functional assemblies for nanotechnology, for instance as therapeutic agents and biosensors.

4.4 Acknowledgements

We thank Prof. Nenad Ban, Dr. Markus Sutter and Dr. W. Frederik Rurup for the donation of plasmids. We acknowledge Aikaterini A. Rousou and Robin Klem for the collaboration on the experiments regarding the stability of encapsulin and *in vivo* studies. Aijie Liu and Dr. Christoph Traulsen are acknowledged for assistance with encapsulin immobilization. We thank Prof. José R. Castón and team (Centro Nacional Biotecnología, Madrid, Spain) for cryo-EM measurement and image reconstruction, Dr. Enrico G. Keim for assistance with TEM imaging and Dr. Mark A. Smithers for assistance with SEM imaging.

4.5 Materials and methods

4.5.1. General

B. linens encapsulin was recombinantly expressed in *E. coli* and purified based on reported procedures.^[14,18] Encapsulin samples were stored in “encapsulin storage buffer” containing 20 mM Tris-HCl, 150 mM NH₄Cl, 1 mM β -mercaptoethanol at pH 7.5. For pH variation studies, 100 mM acetate buffer (CH₃COONa-CH₃COOH) was used to prepare pH 3, pH 4 and pH 5 buffers, 20 mM Tris-HCl buffer was used for pH 7.5 and pH 9 buffers, and 10 mM phosphate buffer (Na₂HPO₄-NaOH) was used for pH 10 and pH 11 buffers. All chemicals were purchased from Sigma-Aldrich unless stated otherwise. All incubations and reactions were conducted at room temperature unless stated otherwise.

For data processing and representation, all plotting and mathematical calculations were performed with OriginPro 9.0 software. Protein structures were rendered using PyMOL 1.3 software and chemical structures were drawn using ChemBioDraw Ultra 12.0 software.

4.5.2. Cryo-EM and image processing

Non-loaded, DyP- and TFP-loaded *B. linens* encapsulin (5 μ L) were applied to one side of Quantifoil R 2/2 holey grids, blotted and plunged into liquid ethane in a Leica EM CPC cryofixation unit. Samples were analyzed in a Tecnai G2 electron microscope equipped with a field emission gun operating at 200kV, and images were recorded under low-dose conditions with a FEI Eagle CCD at a detector magnification of 69,444X (2.16 Å/pixel sampling rate). Image processing operations were performed using Xmipp^[36] and Relion^[37] packages integrated in the Scipion platform.^[38] Graphic representations were produced by UCSF Chimera.^[39] The Xmipp automatic picking routine was used to select 6,785, 19,591 and 29,680 individual particle images of non-loaded, DyP- and TFP-loaded encapsulin, respectively. A 1.2 – 4.1 μ m defocus range was determined for each image with CTFind4.^[40] Particle images were extracted, normalized and downsampled to a factor of 2, with a final sampling ratio of 4.32 Å/pixel. Using the Relion routine, a two-dimensional (2D) classification was performed to discard low-quality particles, and 5,357, 15,941 and 19,779 isometric particles were selected for non-loaded, DyP- and TFP-loaded encapsulin, respectively. A 3D classification was run using the structure of *Thermotoga maritima*

encapsulin (PDB entry 3DKT), low-pass filtered to 40 Å, as an initial model. When icosahedral symmetry was imposed, a single class with 4,246, 12,833 and 15,976 particles was obtained for non-loaded, DyP- and TFP-loaded encapsulin, respectively. When assuming no symmetry for non-loaded, DyP- and TFP-loaded encapsulin, three classes were obtained, but no significant differences were observed between them at the resolutions achieved and particles were refined together. DyP-loaded encapsulin particles were analyzed considering C3 symmetry and a single class with 9,930 particles was selected. These data sets were used to obtain the final 3DR using the Relion autorefinement routine. Resolutions of 3D with icosahedral symmetry were estimated from two independent half-datasets using the 0.5 (or 0.3) criterion of the Fourier shell correlation (FSC), and the values for non-loaded, DyP- and TFP-loaded encapsulin were 11.4 (10.7), 15.1 (13.5) and 13.5 (12) Å, respectively. Similarly, resolution values for asymmetric 3DR were 27.3 (23.8), 28.3 (25) and 24.2 (22.7) Å. For the DyP-loaded encapsulin with C3 symmetry, resolution was 24.7 (22.8) Å. The Chimera fitting tool was used to dock the atomic crystallographic model *T. maritima* encapsulin into our cryo-EM encapsulin maps.

4.5.3. Characterization of encapsulin stability

Size-exclusion profiles of encapsulin upon pH and ionic strength variation were obtained by injecting and running 500 µL of each samples (~15 µM) into a Superose 6 preparative column 10/100 GL (GE Healthcare FPLC Äkta purifier 900 with a 24 mL bed volume). The hydrodynamic size distribution of the particles was determined using a particle analyzer Nanotrac Wave (Microtrac). For particle imaging with TEM, 5 µL of a sample was applied onto Formvar-carbon coated grids and the liquid was drained after 30 s. Afterwards, 5 µL of a staining solution consisting of uranyl acetate (1% w/v) was added onto the grids and the liquid was drained after 1 min. For protein characterization with denaturing gel electrophoresis (SDS-PAGE), experiments using 12% polyacrylamide gel were conducted based on procedures in literature^[41] and Bio-Safe Coomassie (Bio-Rad) was used for visualization of protein bands.

4.5.4. Immobilization of encapsulin and catalytic assay

Substrates (glass or silicon) were rinsed with water and activated by immersion/cycling piranha solution ($\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2 = 3:1$) and rinsed with water and ethanol, then dried with a stream of nitrogen. !Caution: Piranha solution reacts violently with organic material and has to be treated with utmost care! Perfluorophenyl-11-(triethoxysilyl)undecanoate (PFPS) was deposited by substrate immersion into a PFPS solution (DCM, 10 mM) for 24 h at room temperature under argon atmosphere. The glass wafer was rinsed with DCM to remove unreacted reagent and dried in a stream of N_2 . Encapsulin particles were deposited from buffered solution (0.2 M phosphate buffer, pH 7.2). The particles were drop-coated on the flat glass substrate overnight in a closed vessel to avoid solvent evaporation. Afterwards, the glass was rinsed with buffer solution to remove non-immobilized particles.

For catalytic assay, 200 μL of glucose (3 different final concentrations: 0.5 mM, 3.2 mM, 5 mM) and 200 μL of ABTS (2.5 mM) in Tris-HCl buffer pH 7.5 was placed inside a cuvette together with the glass surface with the immobilized encapsulin particles. Afterwards, 100 μL of GOx (1 μM) in PBS buffer pH 7.4 was added into the cuvette to start the reaction and the absorption value at $\lambda = 410 \text{ nm}$ was immediately recorded for 6 min using a Perkin Elmer Lambda 850 UV-visible spectrometer. As control experiments, a PFPS-modified glass surface without encapsulin particles was used for similar kinetic studies.

The reaction rate based on radical ABTS production was calculated as follows:

$$\text{reaction rate} = \frac{d[\text{radical ABTS}]}{dt} = \frac{dA}{dt} \times \frac{1}{\epsilon \times b}$$

where A is the absorbance of radical ABTS at $\lambda = 410 \text{ nm}$, ϵ is the extinction coefficient of radical ABTS at $\lambda = 410 \text{ nm}$ ($36000 \text{ M}^{-1} \text{ cm}^{-1}$) and b is the cuvette path length (1 cm). dA/dt corresponds to the maximum slope of the kinetic plots in Figure 4.8D and Figure 4.9. The apparent turnover number k_{cat} was calculated based on V_{max} /total protein concentration. Although the particles were immobilized on the glass surface, the catalytic assay was performed in solution (i.e., the modified surface was fully immersed inside a cuvette). The total protein concentration was calculated based on the number of particles on the surface and the volume of reaction.

4.5.5. Cell experiments

B. linens encapsulin containing mTFP (monomeric teal fluorescent protein) were recombinantly produced in *E. coli* and purified using the same protocol established for native *B. linens* encapsulin. The concentration of TFP-E was 0.70 μM . Murine macrophage cells (J774) were cultivated in a DMEM. Following the cultivation, 100 μL of ~2000 cells were plated per well on a 96-well plate and either 0, 5 or 10 μL of TFP-E was added into the well (0, 0.013, 0.06 μM TFP-E). The treated cells were incubated at 37°C, 5% CO_2 for 4 h to allow the cells to take up the particles. Following the incubation, the medium was removed and the cells were rinsed with phosphate buffer saline (10 mM PBS, pH 7.4) to further remove non-absorbed species. For the visualization based on fluorescence, the cells were imaged using a fluorescence microscope (Olympus TH4-200 with a X-Cite series 120pc Q Laser from Lumen Dynamics, excitation at $\lambda = 460 - 490 \text{ nm}$, emission at $\lambda = 525 \text{ nm}$).

4.6 References

- [1] Uchida, M.; Klem, M. T.; Allen, M.; Suci, P.; Flenniken, M.; Gillitzer, E.; Varpness, Z.; Liepold, L. O.; Young, M.; Douglas, T. *Adv. Mater.* **2007**, *19*, 1025-1042.
- [2] Peek, L. J.; Middaugh, C. R.; Berkland, C. *Adv. Drug Deliv. Rev.* **2008**, *60*, 915-928.
- [3] Lee, E. J.; Lee, N. K.; Kim, I. S. *Adv. Drug Deliv. Rev.* **2016**, *106*, 157-171.
- [4] Giessen, T. W. *Curr. Opin. Chem. Biol.* **2016**, *34*, 1-10.
- [5] Choi, B.; Moon, H.; Hong, S. J.; Shin, C.; Do, Y.; Ryu, S.; Kang, S. *ACS Nano* **2016**, *10*, 7339-7350.
- [6] Sutter, M.; Boehringer, D.; Gutmann, S.; Guenther, S.; Prangishvili, D.; Loessner, M. J.; Stetter, K. O.; Weber-Ban, E.; Ban, N. *Nat. Struct. Mol. Biol.* **2008**, *15*, 939-947.
- [7] Moon, H.; Lee, J.; Min, J.; Kang, S. *Biomacromolecules* **2014**, *15*, 3794-3801.
- [8] Tamura, A.; Fukutani, Y.; Takami, T.; Fujii, M.; Nakaguchi, Y.; Murakami, Y.; Noguchi, K.; Yohda, M.; Odaka, M. *Biotechnol. Bioeng.* **2015**, *112*, 13-20.
- [9] Corchero, J. L.; Cedano, J. *Microb. Cell. Fact.* **2011**, *10*, 92.
- [10] Kushnir, N.; Streatfield, S. J.; Yusibov, V. *Vaccine* **2012**, *31*, 58-83.
- [11] Zhen, Z.; Tang, W.; Todd, T.; Xie, J. *Expert Opin. Drug Deliv.* **2014**, *11*, 1913-1922.
- [12] Zeth, K.; Hoiczky, E.; Okuda, M. *Trends Biochem. Sci.* **2016**, *41*, 190-203.

- [13] Choi, S. H.; Kwon, I. C.; Hwang, K. Y.; Kim, I. S.; Ahn, H. J. *Biomacromolecules* **2011**, *12*, 3099-3106.
- [14] Rurup, W. F.; Snijder, J.; Koay, M. S. T.; Heck, A. J. R.; Cornelissen, J. J. L. M. *J. Am. Chem. Soc.* **2014**, *136*, 3828-3832.
- [15] Cassidy-Amstutz, C.; Oltrogge, L.; Going, C. C.; Lee, A.; Teng, P.; Quintanilla, D.; East-Seletsky, A.; Williams, E. R.; Savage, D. F. *Biochemistry* **2016**, *55*, 3461-3468.
- [16] Contreras, H.; Joens, M. S.; McMath, L. M.; Le, V. P.; Tullius, M. V.; Kimmey, J. M.; Bionghi, N.; Horwitz, M. A.; Fitzpatrick, J. A.; Goulding, C. W. *J. Biol. Chem.* **2014**, *289*, 18279-18289.
- [17] McHugh, C. A.; Fontana, J.; Nemecek, D.; Cheng, N. Q.; Aksyuk, A. A.; Heymann, J. B.; Winkler, D. C.; Lam, A. S.; Wall, J. S.; Steven, A. C.; Hoiczky, E. *EMBO J.* **2014**, *33*, 1896-1911.
- [18] Putri, R. M.; Fredy, J. W.; Cornelissen, J. J. L. M.; Koay, M. S. T.; Katsonis, N. *ChemPhysChem* **2016**, *17*, 1815-1818.
- [19] Rahmanpour, R.; Bugg, T. D. H. *FEBS J.* **2013**, *280*, 2097-2104.
- [20] 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.; Heck, A. J. R. *Biomacromolecules* **2016**, *17*, 2522-2529.
- [21] Akita, F.; Chong, K. T.; Tanaka, H.; Yamashita, E.; Miyazaki, N.; Nakaishi, Y.; Suzuki, M.; Namba, K.; Ono, Y.; Tsukihara, T.; Nakagawa, A. *J. Mol. Biol.* **2007**, *368*, 1469-1483.
- [22] Abrescia, N. G.; Bamford, D. H.; Grimes, J. M.; Stuart, D. I. *Annu. Rev. Biochem.* **2012**, *81*, 795-822.
- [23] Valdes-Stauber, N.; Scherer, S. *Appl. Environ. Microbiol.* **1994**, *60*, 3809-3814.
- [24] Lavelle, L.; Michel, J. P.; Gingery, M. *J. Virol. Methods.* **2007**, *146*, 311-316.
- [25] Teramoto, H.; Nakajima, K.; Takabayashi, C. *Biomacromolecules* **2004**, *5*, 1392-1398.
- [26] Basle, E.; Joubert, N.; Pucheault, M. *Chem. Biol.* **2010**, *17*, 213-227.
- [27] Ai, H. W.; Henderson, J. N.; Remington, S. J.; Campbell, R. E. *Biochem. J.* **2006**, *400*, 531-540.
- [28] Lian, X.; Chen, Y.-P.; Liu, T.-F.; Zhou, H.-C. *Chem. Sci.* **2016**, *7*, 6969-6973.
- [29] Sassolas, A.; Blum, L. J.; Leca-Bouvier, B. D. *Biotechnol. Adv.* **2012**, *30*, 489-511.

- [30] Homaei, A. A.; Sariri, R.; Vianello, F.; Stevanato, R. *J. Chem. Biol.* **2013**, *6*, 185-205.
- [31] Lockett, M. R.; Phillips, M. F.; Jarecki, J. L.; Peelen, D.; Smith, L. M. *Langmuir* **2008**, *24*, 69-75.
- [32] Heller, A.; Feldman, B. *Chem. Rev.* **2008**, *108*, 2482-2505.
- [33] Shlyahovsky, B.; Li, D.; Katz, E.; Willner, I. *Biosens. Bioelectron.* **2007**, *22*, 2570-2576.
- [34] Nicol, M. J.; Duke, F. R. *J. Biol. Chem.* **1966**, *241*, 4292-4293.
- [35] Salvachúa, D.; Prieto, A.; Martínez, A. T.; Martínez, M. J. *Appl. Environ. Microbiol.* **2013**, *79*, 4316-4324.
- [36] Marabini, R.; Masegosa, I. M.; San Martin, M. C.; Marco, S.; Fernandez, J. J.; de la Fraga, L. G.; Vaquerizo, C.; Carazo, J. M. *J. Struct. Biol.* **1996**, *116*, 237-240.
- [37] Scheres, S. H. *J. Struct. Biol.* **2012**, *180*, 519-530.
- [38] de la Rosa-Trevin, J. M.; Quintana, A.; Del Cano, L.; Zaldivar, A.; Foche, I.; Gutierrez, J.; Gomez-Blanco, J.; Burguet-Castell, J.; Cuenca-Alba, J.; Abrishami, V.; Vargas, J.; Oton, J.; Sharov, G.; Vilas, J. L.; Navas, J.; Conesa, P.; Kazemi, M.; Marabini, R.; Sorzano, C. O.; Carazo, J. M. *J. Struct. Biol.* **2016**, *195*, 93-99.
- [39] Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605-1612.
- [40] Rohou, A.; Grigorieff, N. *J. Struct. Biol.* **2015**, *192*, 216-221.
- [41] Shapiro, A. L.; Vinuela, E.; Maizel, J. V. *Biochem. Biophys. Res. Commun.* **1967**, *28*, 815-820.

Chapter 5

Labeling encapsulin with light-switchable fluorophores

The emergence of bacterial nanocages as nano-carriers to penetrate cells calls for the development of tracking strategies that would enable in vivo studies. Here, a strategy to label bacterial nanocages with photo-switchable fluorophores is presented, to facilitate their imaging by super-resolution microscopy. We demonstrate the functionalization of the encapsulin from Brevibacterium linens with a spiropyran photo-switch, which is not fluorescent, by covalent attachment to the amine residues at the outer encapsulin shell. Upon alternating irradiation with ultraviolet and visible light, the spiropyran switches forth and back to its fluorescent merocyanine photo-isomer and thus the fluorescence can be switched on and off, reversibly. The bacterial compartments preserve their structural integrity upon covalent modification and over at least five irradiation cycles.

This chapter is published in R.M. Putri, J.W. Fredy, J.J.L.M. Cornelissen, M.S.T. Koay and N. Katsonis, *ChemPhysChem* **2016**, 17, 1815-1818 (special issue on molecular machines), and inside cover page.

5.1 Introduction

The work presented in Chapter 4 has highlighted how encapsulins emerge as promising building blocks for nanobiotechnology.^[1-8] For *in vivo* applications, the ability of encapsulins as nanoplateforms to penetrate cells while carrying functional molecules has been demonstrated,^[3,8] but further understanding of their *in vivo* behavior calls for the development of non-invasive imaging and tracking strategies, while not compromising the self-assembly. Overcoming the diffraction barrier for microscopy can be achieved in live cells by using super-resolution fluorescence imaging,^[9-15] by photo-switching fluorescence on and off to statistically turn on a small fraction of fluorophores only.^[9, 14-18] Such technique known as Stochastic Optical Reconstruction Microscopy (STORM) involves the mapping of fluorescent points by activating only a subset of fluorophores stochastically at any given time while the rest are switched off (Figure 5.1).^[9] Super-resolution images are then reconstructed from multiple time-resolved fluorescence images at different areas. Moreover, the possibility to switch fluorescence on or off might reveal useful to avoid overlapping signals in the presence of various labeled structures. With this in mind, we set out to label bacterial nanocages with photo-switchable fluorophores.

Figure 5.1 Principle of STORM imaging. An object is labeled with fluorophores (orange circles) that can be switched between an off (dark) state and on (fluorescent) state by light irradiation at different wavelengths, for instance, red laser to switch off and green laser to switch on (depicted by red and green arrows). Initially, the fluorophores are all present in their dark state (black circles) and in each cycle, a green laser is used to switch on only a fraction of the fluorophores. From multiple imaging cycles, optically resolvable sets of active fluorophores are then obtained and used to reconstruct the overall image. Reprinted with permission from [9]. Copyright © 2006, Macmillan Publishers Limited, part of Springer Nature.

Natural fluorophores such as the green fluorescent protein shows an on/off blinking behavior under examination at the single molecule level, and thus they have been used as fluorescence probes in live-cell imaging.^[19, 20] However, despite their biocompatibility, they suffer drawbacks compared to synthetic fluorophores, which are often brighter than their biological counterparts,^[10] allow chemical modifications on surfaces^[21, 22] and are less likely to interfere with protein folding because they are smaller than fluorescent proteins.^[17] Consequently, synthetic and photo-switchable fluorescence probes have emerged as powerful tools for *in vivo* imaging,^[20] including diarylethenes^[23] and spiropyran derivatives.^[24] However, their moderate brightness remains an issue, which is why multiple fluorophores are often grafted at the surface of the same object.^[21, 25] Here, we graft multiple spiropyran switches on the surface of the encapsulin. The spiropyran switches can be converted reversibly into their fluorescent photo-isomer upon irradiation with light, thus allowing a controlled on/off fluorescence.^[26-31] In addition, the fluorescent isomer of spiropyran provides sufficient brightness for biological applications in live-cell imaging.^[10, 24, 32]

5.2 Results and discussion

5.2.1 Covalent attachment of spiropyran to encapsulin

B. linens encapsulin counts 240 lysine residues that are accessible on its surface, which accounts for four lysines per monomer, the amine group of which are used as active sites for chemical modification. The sequence of encapsulin monomer is given below based on the DNA database: UniProtKB - Q45296 (LIN18_BRELN). Lysine residues are highlighted in yellow.

MNNLYRELAPIGPAWAEIEEEEARRTFKRNIAGRRIVDVAGPTGFETSAVTTGHI
RDVQSETSGQLQVKQRIVQEYIELRTPFTVTRQAIDDVARGSGDSWDWQPVKDAA
TTIAMAEDRAILHGLDAAGIGGIVPGSSNAAVAIPDAVEDFADAVAQALSVLRT
VGVDGPYSLLLSSAEYTKVSESTDHGYPHLSRQLGAGEIHWAPALEGALLVS
TRGGDYELHLGQDLSIGYYSHDSETVELYLQETFGFLALTDESSVPLSL

Our approach involves the functionalization of spiropyran with a succinimide moiety (Figure 5.2a, black), which reacts with the amine group of lysines in order to form an amide bond.

Labeling encapsulin with light-switchable fluorophores

The spiropyran (closed, off-state) is then converted to the colored merocyanine form (open, on-state) upon irradiation with UV light. Without UV irradiation, a spontaneous conversion of spiropyran form to merocyanine form can also occur in the dark in aqueous environment and is particularly favored in acidic environment, hence, all samples are kept under ambient light (to stabilize the spiropyran form) and in a neutral solution (in PBS buffer pH 7.4).^[27] This conversion is reversible upon subsequent irradiation with visible light (Figure 5.2b). The merocyanine also undergoes a slow thermal relaxation into spiropyran in the dark, although the open (charged) form is more stabilized in a polar environment.^[27]

Figure 5.2 Covalent attachment of spiropyran-succinimide onto encapsulin. a) Coupling of encapsulin and spiropyran via amine-succinimide reaction. b) Photo-isomerization of non-fluorescent spiropyran (blue) to fluorescent merocyanine (red).

Coupling the encapsulin nanocages to the spiropyran is achieved by adding 1000-times excess of spiropyran to the purified encapsulin solution at pH 7.4, containing 10% DMSO. After overnight incubation, the non-reacted spiropyran and DMSO are removed by dialysis and size-exclusion chromatography (SEC). In this process, non-covalently bound spiropyran molecules are removed by the combination of overnight dialysis and 48-times dilution throughout the SEC column (a buffered saline is used as the eluent to avoid electrostatic interactions).^[33] The chromatogram reveals a protein peak at around $V = 11 - 12$ mL which is characteristic for encapsulin particles (Figure 5.3a).

Figure 5.3 Labeling encapsulin with a spiropyran photo-switch. a) Size-exclusion chromatogram showing the characteristic encapsulin elution at $V = 11 - 12$ mL as well as the characteristic spiropyran absorption at $\lambda = 350$ nm at the corresponding elution volume. b) UV-visible spectrum of encapsulin labeled with spiropyran switches.

Encapsulin particles, owing to their sizes of 20 – 25 nm, always elute at around $V = 11 - 12$ mL, monitored at protein primary absorption at $\lambda = 280$ nm. When the encapsulins are modified with spiropyran switches, another wavelength is monitored during the chromatography (i.e., the characteristic absorption for spiropyran) that is at $\lambda = 350$ nm. The modification with the switches does not change the size of encapsulin particles significantly since the hybrid system still elutes at $V = 11 - 12$ mL during the chromatography (Figure 5.3a). The absorption at $\lambda = 350$ nm at the corresponding elution volume indicates covalent coupling of the spiropyran to the encapsulin as non-covalently bound spiropyran has been removed. The presence of spiropyran is further confirmed by UV-visible spectroscopy (Figure 5.3b). The small absorption band observed at around $\lambda = 520$ nm prior to any irradiation (Figure 5.3b) is likely due to a conversion of the spiropyran form into the merocyanine form, which occurs spontaneously in water.

To quantify the labeling, the concentration of encapsulin is determined based on gel densitometry, while the concentration of spiropyran could be determined based on the absorption at $\lambda = 350$ nm in Figure 5.3b. Since the encapsulin does not absorb at this wavelength, we are able to use the absorbance at $\lambda = 350$ nm of the hybrid system to determine the spiropyran concentration in the hybrid system ($\epsilon = 7,994 \text{ M}^{-1} \text{ cm}^{-1}$), which results in the value of 1.16 mM of spiropyran. On the other hand, the encapsulin monomer is shown in SDS-PAGE gel (Figure 5.4), corresponding to the band at around 28 kDa (lanes 2,

3, 4). Importantly, lane number 3 confirms the present of the encapsulin component in the hybrid of encapsulin-spiropyran. By comparing the intensity of the encapsulin band in lane 3 and the intensity of the encapsulin standard in lane 4 (known concentration = 0.14 mM), we could roughly estimate the concentration of the encapsulin in the hybrid system that is 0.01 mM. Furthermore, we could estimate the number of spiropyran molecules per one encapsulin particle that is $\sim (1.16 \text{ mM}/0.01 \text{ mM}) = 116$ ($\sim 49\%$ coverage).

Figure 5.4 Gel densitometry to determine encapsulin concentration. Denaturing gel electrophoresis confirms the presence of encapsulin in the hybrid sample (lane 3) based on the characteristic encapsulin band at around 28 kDa (the monomer of the encapsulin cage), which is comparable to the native encapsulin (lane 2 and lane 4). The intensity of the encapsulin band in lane 4 is used as a standard (known concentration = 0.14 mM) to estimate the concentration of the encapsulin in the hybrid system (lane 3). The systematic comparison of band intensity (gel densitometry) is performed with Bio-Rad Image Lab software.

5.2.2 Photo-switchable fluorescence of modified encapsulin

The photo-isomerization of labeled encapsulin in PBS (pH 7.4) is examined by irradiating the sample with UV light for 2 min ($\lambda = 365 \text{ nm}$) in solution. The photo-isomerization of spiropyran into merocyanine is indicated by the appearance of an absorption band around $\lambda = 540 \text{ nm}$ (Figure 5.5a and b). This conversion is complete after 2 min of irradiation and prolonged irradiation does not significantly increase the merocyanine concentration. We also observe the appearance of merocyanine fluorescence (Figure 5.5c and d) at $\lambda = 615 \text{ nm}$ upon excitation at $\lambda = 535 \text{ nm}$. The excitation with 535 nm light for fluorescence measurement is unlikely to significantly revert the merocyanine form to spiropyran form because it occurs

only for a short time (around 1 – 2 min of spectrum scanning), while the conversion back to spiropyran form requires 7 min of continuous irradiation of visible light with considerable intensity (flux density: 46 mW/cm²).

Figure 5.5 Photo-triggered on/off fluorescence of spiropyran-labeled encapsulin. a) UV-visible spectra showing the characteristic absorption band of merocyanine at $\lambda = 540$ nm. b) Reversible photo-switching of spiropyran to merocyanine monitored at $\lambda = 540$ nm. c) Emission spectra showing the fluorescence of merocyanine at $\lambda = 615$ nm upon excitation at $\lambda = 535$ nm. d) Reversible conversion of on-state merocyanine to off-state spiropyran monitored at $\lambda = 615$ nm.

We irradiate the encapsulin-merocyanine with visible light ($\lambda \geq 420$ nm) to investigate the reversibility. The disappearance of the absorption band at $\lambda = 540$ nm, after 7 min of continuous irradiation, indicates that the merocyanine form is completely reverted to the spiropyran form (Figure 5.5a). Similarly, we observe a decrease of the fluorescence of merocyanine at $\lambda = 615$ nm after irradiation with visible light (Figure 5.5c). We further repeat the ultraviolet light and visible light irradiation cycle up to five times. After the fifth cycle, the merocyanine displays absorbance and fluorescence that amount to half their initial values (Figure 5.5b and 5.5d). The alternating decrease and increase of the fluorescence during the cycles confirms the photo-switching of the fluorophore (Figure 5.5d). The merocyanine form attached to the protein also undergoes thermal relaxation into the spiropyran form in the dark,

which can also be observed by the disappearance of its absorption band at $\lambda = 540 - 550$ nm (Figure 5.6). However, this process is much slower in comparison to the conversion back with continuous visible light irradiation.

Figure 5.6 Thermal relaxation of merocyanine into spiropyran in aqueous system (the switch is attached to a model protein, human serum albumin). The relaxation is indicated by the disappearance of absorption band at $\lambda = 550$ nm, which is much slower (within hours) in comparison to the light-triggered conversion (within 7 min).

5.2.3 Structural integrity of encapsulin particles

To confirm that the encapsulin particles are stable upon covalent modification and photo-switching, we characterize the hydrodynamic size of the particles before and after irradiation by dynamic light scattering (DLS). DLS analysis reveals a diameter of about 20 – 25 nm, which indicates that the functionalized encapsulins remain intact after irradiation with both UV and visible light (Figure 5.7). The intact particles after irradiation are also visualized with TEM (Figure 5.8a). Nevertheless, structures with sizes exceeding 100 nm are also observed after the fifth cycle, which are likely aggregates forming upon the protein degradation caused by prolonged irradiation (Figure 5.8b).^[34]

Figure 5.7 Structural integrity of labeled encapsulin measured by DLS. a) Size distribution of encapsulin particles throughout the alternating irradiation with UV and visible light peaking at diameter = 20 – 25 nm. b) Comparison of size distribution before any irradiation (black), after UV irradiation of the first cycle (red) and after visible irradiation of the fifth cycle (blue) showing that the structural integrity of the particles is preserved throughout the cycles.

Figure 5.8 Effect of irradiation on modified encapsulin. a) TEM image of encapsulin particles (diameter = 20 – 25 nm) after irradiation with UV light (for 2 min), confirming that the majority of the particles are still intact upon irradiation. b) Size distribution of encapsulin particles (diameter ~ 20 nm) after 5 cycles of irradiation showing that the intact encapsulin particles are present but larger structures also form (diameter >100 nm).

5.3 Conclusions

We have demonstrated a strategy to label encapsulin nanocages isolated from *Brevibacterium linens* with a spiropyran photo-switch. The fluorescence of the hybrid nanocompartment resulting from this operation can be photo-switched on-command for multiple cycles. The encapsulin particles retain their structural integrity, both upon covalent modification with the dye and upon light irradiation. These results provide opportunities towards tracking protein-based nanocages in live cells.

5.4 Acknowledgements

Dr. Jean-Wilfried Fredy is gratefully acknowledged for the design and synthesis of the photo-switch. We thank Dr. Enrico G. Keim for assistance with TEM imaging and Federico Lancia for help with characterizing the photo-switch.

5.5 Materials and methods

5.5.1 General

B. linens encapsulin is recombinantly expressed and purified according to the protocol in literature.^[1,2] All experiments are performed at pH 7.4 in a phosphate buffered saline (0.01 M PBS, Sigma-Aldrich). Chemicals are purchased from Sigma-Aldrich unless stated otherwise.

5.5.2 Synthesis of spiropyran-succinimide

The synthesis of product **1** to **4** (Figure 5.9) is carried out according to literature procedures.^[35,36]

1-(3-Carbomethoxypropyl)-3,3-dimethyl-2-methyleneindoline (1). To a solution of 2,3,3-trimethyl-3H-indole (500 mg, 3.14 mmol, 1 eq) in 5 mL of chloroform, methyl 4-bromobutyrate (600 mg, 3.30 mmol, 1.05 eq) is added. The reaction mixture is stirred under N₂ for 24 h at 70°C. The chloroform is evaporated and diethylether is added to obtain a

precipitate. The solid is extensively washed with diethylether. 490 mg is obtained as a purple solid.

^1H NMR (400 MHz, CD_3CN) δ 7.92 – 7.86 (m, 1H), 7.74 – 7.69 (m, 1H), 7.66 – 7.61 (m, 2H), 4.49 – 4.41 (m, 2H), 3.63 (s, 3H), 2.78 (s, 3H), 2.61 (dd, J = 8.6, 5.0 Hz, 2H), 2.23 – 2.13 (m, 3H), 1.55 (s, 6H).

^{13}C NMR (101 MHz, CD_3CN) δ 197.8, 173.9, 143.0, 130.9, 130.2, 124.4, 116.3, 55.6, 52.4, 48.3, 30.9, 23.4, 22.6, 14.9.

ESI MS (m/z) 260.02, calculated for $[\text{C}_{16}\text{H}_{22}\text{NO}_2]^+ = 260.16$.

1'-(3-Carbomethoxypropyl)-3',3'-dimethyl-6-nitrospiro[2H-1]-benzopyran-2,2'-indoline (2).

To a solution of 5-nitrosalicylaldehyde (68 mg, 0.41 mmol, 1 eq) in 3 mL of ethanol, the compound **1** (140 mg, 0.41 mmol, 1 eq) dissolved in 2 mL of ethanol is added slowly. The reaction mixture is stirred at 80°C for 5 h. The solvent is evaporated and the residue is purified by chromatography on silica gel to afford the product as a purple solid (58 mg, 56%)

^1H NMR (400 MHz, CD_3CN) δ 8.07 (d, J = 2.8 Hz, 1H), 8.00 (d, J = 2.8 Hz, 1H), 7.98 (d, J = 2.8 Hz, 1H), 7.15 (td, J = 7.7, 1.3 Hz, 1H), 7.11 (dd, J = 7.3, 0.8 Hz, 1H), 7.02 (m, 1H), 6.87 – 6.81 (m, 1H), 6.71 (m, 1H), 6.66 (d, J = 7.8 Hz, 1H), 5.95 (d, J = 10.4 Hz, 1H), 3.57 (s, 3H), 3.26 – 3.10 (m, 2H), 2.32 (tt, J = 8.1, 4.1 Hz, 2H), 1.92 – 1.75 (m, 2H), 1.26 (s, 3H), 1.15 (s, 3H).

^{13}C NMR (101 MHz, CD_3CN) δ 174.44, 160.39, 148.13, 142.08, 137.11, 129.11, 128.72, 126.60, 123.73, 122.81, 122.74, 120.41, 119.97, 116.27, 107.93, 107.81, 53.36, 52.00, 43.58, 31.81, 26.27, 24.77, 20.00.

ESI MS (m/z) 409.28, calculated for $[\text{C}_{23}\text{H}_{24}\text{N}_2\text{O}_5\text{H}]^+ = 409.17$.

1'-(3-Carboxypropyl)-3',3'-dimethyl-6-nitrospiro[2H-1]benzopyran-2,2'-indoline (3).

To a solution of **2** (150 mg) in THF (3 mL), 1 mL of NaOH 10% is added. The reaction mixture is stirred at room temperature for 18 h. Afterwards, it is acidified with a solution of HCl (1 M). The aqueous phase is extracted with ethylacetate and concentrated to afford the product as a yellow solid. (124 mg, 86%).

^1H NMR (400 MHz, CD_3CN) δ 10.06 (d, $J = 0.4$ Hz, 1H), 8.70 (dd, $J = 11.1, 2.9$ Hz, 1H), 8.42 (dd, $J = 9.2, 2.8$ Hz, 1H), 8.10 (d, $J = 2.8$ Hz, 2H), 8.02 (dd, $J = 9.0, 2.8$ Hz, 2H), 7.21 – 7.12 (m, 5H), 7.05 (dd, $J = 10.2, 4.3$ Hz, 2H), 6.86 (td, $J = 7.5, 0.9$ Hz, 2H), 6.74 (d, $J = 9.0$ Hz, 2H), 6.70 (d, $J = 7.8$ Hz, 2H), 3.31 – 3.13 (m, 4H), 2.38 – 2.28 (m, 5H), 1.85 (dddd, $J = 15.4, 13.8, 10.7, 6.2$ Hz, 6H), 1.27 (s, 8H), 1.18 (d, $J = 6.1$ Hz, 3H).

^{13}C NMR (101 MHz, CD_3CN) δ 197.6, 174.7, 166.8, 160.4, 148.2, 142.1, 137.1, 132.5, 130.9, 129.1, 128.7, 126.6, 123.7, 122.8, 122.7, 120.4, 120, 119.5, 116.3, 107.9, 107.8, 53.4, 43.64, 31.5, 26.3, 24.7, 20.0.

ESI MS (m/z) 395.09, calculated for $[\text{C}_{22}\text{H}_{22}\text{N}_2\text{O}_3\text{H}]^+ = 395.16$.

Spiropyran-succinimide (4). Compound **3** (50 mg, 0.13 mmol, 1 eq) and N-hydroxysuccinimide (15 mg, 0.13 mmol, 1 eq) are dissolved in 5 mL in dried acetonitrile. DCC (26 mg, 0.13 mmol, 1 eq) is added and the reaction mixture is stirred at room temperature for 18 h under nitrogen. The solvent is evaporated and the residue is purified by chromatography over silica gel with hexane/ethylacetate to afford the product as a purple solid (20 mg, 32%).

^1H NMR (400 MHz, CD_3CN) δ 8.07 (d, $J = 2.7$ Hz, 1H), 7.99 (dd, $J = 9.0, 2.7$ Hz, 1H), 7.17 – 7.07 (m, 2H), 7.03 (dd, $J = 10.4, 6.2$ Hz, 1H), 6.88 – 6.81 (m, 1H), 6.74 – 6.63 (m, 2H), 5.98 (dd, $J = 10.4, 7.0$ Hz, 1H), 3.32 – 3.11 (m, 2H), 2.76 (s, 2H), 2.66 (m, 2H), 2.59 – 2.56 (m, 2H), 2.36 – 2.26 (m, 2H), 1.25 (s, 3H), 1.15 (s, 3H).

^{13}C NMR (101 MHz, CD_3CN) δ 169.7, 168.2, 159.1, 146.8, 140.8, 136.0, 128.08, 127.8, 125.3, 122.4, 121.5, 121.4, 119.3, 118.7, 114.7, 106.5, 52.0, 42.3, 41.8, 27.9, 25.0, 23.4, 18.7

ESI MS (m/z) 491.94, calculated for $[\text{C}_{26}\text{H}_{25}\text{N}_3\text{O}_7\text{H}]^+ = 492.17$.

Figure 5.9 Synthesis of succinimide-bearing nitrospiropyran

5.5.3 Covalent attachment of spiropyran and encapsulin

A 1000-fold excess of nitro-spiropyran bearing a succinimide is added to an encapsulin solution to a total volume of 500 μL (in 0.01 M PBS pH 7.4 containing 10% DMSO) at room temperature and incubated overnight. The non-reacted spiropyran is removed by size-exclusion chromatography (SEC, preparative column Superose 6 10/100 GL, GE Healthcare FPLC Äkta purifier 900 with a 24 mL bed volume). Prior to injecting the samples into the column, all samples are dialyzed against PBS to remove the DMSO using 12 – 14 kDa dialysis membranes (Spectra/Por).

5.5.4 Photo-switchable fluorescence analysis

500 μL of the solution containing the encapsulin-spiropyran hybrid is irradiated with UV light ($\lambda = 365 \text{ nm}$, bluepoint LED Hönle Technology, 40 mW/cm^2) for 2 min and subsequently with visible-light ($\lambda \geq 420 \text{ nm}$, Edmund MI-150 High-intensity Illuminator) for 7 min to complete a cycle of irradiation. After each irradiation, the absorption spectrum is

recorded with a Perkin Elmer Lambda 850 UV-visible spectrometer and the emission spectrum at $\lambda = 555 - 670$ nm (excitation at $\lambda = 535$ nm) is recorded using a Perkin Elmer fluorescence spectrometer.

5.5.5 Characterization of encapsulin structure

The hydrodynamic size of the encapsulin solution (~500 μ L) before and after irradiation in 0.01 mM PBS (pH 7.4) is determined with dynamic light scattering (Nanotracer Wave, Microtrac). In addition to DLS analysis, confirmation of the structural integrity is also provided by TEM (300 kV acceleration voltage, Phillips CM 30) and SEC analysis.

5.6 References

- [1] M. Sutter, D. Boehringer, S. Gutmann, S. Guenther, D. Prangishvili, M. J. Loessner, K. O. Stetter, E. Weber-Ban, N. Ban, *Nat. Struct. Mol. Biol.* **2008**, *15*, 939-947.
- [2] W. F. Rurup, J. Snijder, M. S. T. Koay, A. J. R. Heck, J. J. L. M. Cornelissen. *J. Am. Chem. Soc.* **2014**, *136*, 3828-3832.
- [3] H. Moon, J. Lee, J. Min, S. Kang, *Biomacromolecules* **2014**, *15*, 3794-3801.
- [4] A. Tamura, Y. Fukutani, T. Takami, M. Fujii, Y. Nakaguchi, Y. Murakami, K. Noguchi, M. Yohda, M. Odaka, *Biotechnol. Bioeng.* **2015**, *112*, 13-20.
- [5] R. M. Putri, J. J. L. M. Cornelissen, M. S. T. Koay, *ChemPhysChem* **2015**, *16*, 911-918.
- [6] L. Lavelle, J. P. Michel, M. Gingery, *J. Virol. Methods.* **2007**, *146*, 311-316.
- [7] C. A. Hommersom, B. Matt, A. van der Ham, J. J. L. M. Cornelissen, N. Katsonis, *Org.Biomol. Chem.* **2014**, *12*, 4065-4069.
- [8] B. Choi, H. Moon, S. J. Hong, C. Shin, Y. Do, S. Ryu, S. Kang, *ACS Nano* **2016**, *10*, 7339-7350.
- [9] M. J. Rust, M. Bates, X. W. Zhuang, *Nat. Methods* **2006**, *3*, 793-795.
- [10] C. J. MacNevin, D. Gremyachinskiy, C. W. Hsu, L. Li, M. Rougie, T. T. Davis, K. M. Hahn, *Bioconjugate Chem.* **2013**, *24*, 215-223.
- [11] S. van de Linde, M. Heilemann, M. Sauer, *Annu. Rev. Phys.Chem.* **2012**, *63*, 519-540.
- [12] S. N. Uno, M. Kamiya, T. Yoshihara, K. Sugawara, K. Okabe, M. C. Tarhan, H. Fujita, T. Funatsu, Y. Okada, S. Tobita, Y. Urano, *Nat. Chem.* **2014**, *6*, 681-689.

- [13] S. W. Hell, J. Wichmann, *Opt. Lett.* **1994**, *19*, 780-782.
- [14] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindisser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, H. F. Hess, *Science* **2006**, *313*, 1642-1645.
- [15] S. T. Hess, T. P. K. Girirajan, M. D. Mason, *Biophys. J.* **2006**, *91*, 4258-4272.
- [16] H. Bock, C. Geisler, C. A. Wurm, C. Von Middendorff, S. Jakobs, A. Schonle, A. Egner, S. W. Hell, C. Eggeling, *Appl. Phys. B.* **2007**, *88*, 161-165.
- [17] K. Hirabayashi, K. Hanaoka, M. Shimonishi, T. Terai, T. Komatsu, T. Ueno, T. Nagano, *Chem. Eur. J.* **2011**, *17*, 14763-14771.
- [18] S. W. Hell, *Science* **2007**, *316*, 1153-1158.
- [19] R. M. Dickson, A. B. Cubitt, R. Y. Tsien, W. E. Moerner, *Nature* **1997**, *388*, 355-358.
- [20] M. Fernandez-Suarez, A. Y. Ting, *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 929-943.
- [21] T. Udayabhaskararao, P. K. Kundu, J. Ahrens, R. Klajn, *ChemPhysChem* **2016**, *17*, 1805-1809.
- [22] N. Katsonis, M. Lubomska, M. M. Pollard, B. L. Feringa, P. Rudolf, *Prog. Surf. Sci.* **2007**, *82*, 407-434.
- [23] Y. Zou, T. Yi, S. Z. Xiao, F. Y. Li, C. Y. Li, X. Gao, J. C. Wu, M. X. Yu, C. H. Huang, *J. Am. Chem. Soc.* **2008**, *130*, 15750-15751.
- [24] W. Wan, M. Q. Zhu, Z. Y. Tian, A. D. Q. Li, *J. Am. Chem. Soc.* **2015**, *137*, 4312-4315.
- [25] A. Fihey, A. Perrier, W. R. Browne, D. Jacquemin, *Chem. Soc. Rev.* **2015**, *44*, 3719-3759.
- [26] Y. Kalisky, T. E. Orlowski, D. J. Williams, *J. Phys. Chem.* **1983**, *87*, 5333-5338.
- [27] R. Klajn, *Chem. Soc. Rev.* **2014**, *43*, 148-184.
- [28] M. Q. Zhu, L. Zhu, J. J. Han, W. Wu, J. K. Hurst, A. D. Q. Li, *J. Am. Chem. Soc.* **2006**, *128*, 4303-4309.
- [29] W. Szymanski, J. M. Beierle, H. A. V. Kistemaker, W. A. Velema, B. L. Feringa, *Chem. Rev.* **2013**, *113*, 6114-6178.
- [30] Z. Y. Tian, A. D. Q. Li, *Acc. Chem. Res.* **2013**, *46*, 269-279.
- [31] Z. Y. Tian, W. W. Wu, W. Wan, A. D. Q. Li, *J. Am. Chem. Soc.* **2009**, *131*, 4245-4252.
- [32] E. Deniz, M. Tomasulo, R. A. DeFazio, B. D. Watson, F. M. Raymo, *Phys. Chem. Chem. Phys.* **2010**, *12*, 11630-11634.

- [33] P. Hong, S. Koza, E. S. P. Bouvier, *J. Liq. Chromatogr. Relat. Technol.* **2012**, *35*, 2923-2950.
- [34] M. J. Davies, *Biochem. Biophys. Res. Commun.* **2003**, *305*, 761-770.
- [35] P. Remon, M. Hammarson, S. M. Li, A. Kahnt, U. Pischel, J. Andreasson, *Chem. Eur J.* **2011**, *17*, 6492-6500.
- [36] P. J. Wu, J. L. Chen, C. P. Chen, Y. H. Chan, *Chem. Commun.* **2013**, *49*, 898-900.

Chapter 6

Light-fueled assembly of encapsulin and chaperone into hybrid superstructures

Life gains its sustainability from a continuous input of energy. Re-engineering the principles underlying biological functionality into artificial systems would allow developing new design paradigms for adaptive matter. Previously reported assemblies of semi-synthetic protein systems primarily focus on integrative self-assembly under equilibrium conditions, while extending the design to non-equilibrium dynamic order has proven challenging due to the complexity. Here, we demonstrate that higher degrees of organization emerge in self-assembled protein systems where proton is released gradually under irradiation with light, provided that proteins undergo partial unfolding events. The giant superstructures are formed with defined edges confining protein-based substructures. Their fueled formation is mediated by the presence of a spiropyran-based photoacid, which releases its proton upon irradiation, leading to partial unfolding and exposure of hydrophobic segments of proteins and thus favoring self-association. Simultaneously, the photoacid switches from a hydrophilic state to a hydrophobic state and thus becomes an active building block in the formation of the giant assemblies. As the photo-induced switching and pH drop are rapidly reversible in the absence of light, the superstructures spontaneously disperse when the irradiation is stopped.

6.1 Introduction

In some cases, extended nanostructures and ordered states of protein assemblies in Nature are sustained by a constant flow of energy or fuel. This is exemplified by the assembly of cytoskeletal protein microtubules, where the consumption of a chemical fuel generates a dynamic order that allows carrying complex tasks.^[1,2] Drawing inspiration from the dynamic interplay between fuel and matter in biological processes thus calls for strategies to move beyond architectures under equilibrium conditions to a dynamic order of interacting components.^[3] The distinctive feature of a dynamic order is that it allows a highly reconfigurable system. However, previous attempts on directing higher-order assembly of proteins *in vitro* have relied primarily on the incorporation of bridging interactions (e.g., disulfide bonds,^[4] dendrons,^[5] or metal coordination^[6]) that are encoded in the structure of the proteins by design. Consequently, assembly of such systems occurs spontaneously as the cue for the assembly is embedded in the structure, resulting in architectures under equilibrium conditions. The disassembly would require addition of chemicals that can disrupt the bridging moieties such as a reducing agent or a chelating agent.

Moving beyond spontaneously-assembled architectures, synthetic systems have demonstrated a dynamic assembly that is sustained by a continuous input of a fuel, using metal nanoparticles,^[7] macrocycle^[8] and chemical gelators as the building blocks.^[9] To our knowledge, such system that is based on protein building blocks has not been reported, while protein architectures offer higher complexity, monodispersity and biocompatibility. In multimeric self-assembled proteins, such as cage-like proteins of virus and bacteria, environmental conditions such as pH, ionic strength and polarity of medium govern their assembly behavior.^[10] Dynamic self-association of such proteins can be achieved by reversibly adjusting the pH to change the charges, conformations and properties of proteins to induce self-association in a reversible manner.

In order to generate an energy-dependent dynamic order of proteins, the pH condition that favors self-association should be achieved in a reversible manner and sustained upon an input of energy. To enable this, we use a spiropyran-based photoacid (i.e., the open, protonated merocyanine form, MCH⁺) that reversibly releases its proton upon irradiation with visible light.^[11] Using this molecule, previous reports demonstrated the decrease of pH

up to 2 units within only 20 s of irradiation with visible light upon the conversion of MCH⁺ form to spiropyran (SP) form.^[11,12] The release of the proton from the photoacid is rapidly reversible as the MCH⁺ form is more stable than the SP form (half-life $t_{1/2} \sim 70$ s),^[11] therefore altogether allowing a rapidly reversible control over the assembly. The spontaneous and immediate reversibility of pH control has enabled reversible photo-control of gold nanoparticles assembly,^[13] ring-opening polymerization,^[14] formation and swelling of supramolecular hydrogels,^[15,16] photo-activation of antibiotic^[17] and photo-activation of a biological ion channel.^[18]

6.2 Results and discussion

6.2.1 Light-fueled assembly of hybrid superstructures

Here, we use as building blocks the bacterial nanocompartment encapsulin from *T. maritima* (abbreviated as TM-Enc) that is co-purified with GroEL chaperone upon recombinant expression in *E. coli* (Figure 6.1a).^[19] Both TM-Enc and GroEL are self-assembled multimeric proteins with highly defined structures that favor self-association under extreme pH values.^[20] The native encapsulin particles adapt an icosahedral cage-like structure, composed of 60 copies of identical protein subunits that self-assemble into a cage with the size of around 20 – 25 nm (Figure 6.1b).^[21] We recently reported that encapsulin particles remain intact under irradiation with light.^[22]

GroEL chaperone from *E. coli*, a tetradecameric cylindrical protein (~15 nm in diameter, Figure 6.1b), is responsible for assisting the folding process of polypeptides into their native protein conformation.^[23-25] The so-called apical domain located at the opening of cylindrical GroEL consists of hydrophobic surfaces that are responsible for polypeptide binding in the GroEL functioning.^[23,24] Furthermore, GroEL molecules have a tendency to stack into fibrillar assembly upon agitation or in the presence of surfactant such as sodium dodecyl sulfate (SDS).^[20,26] Recently, a design of photo-responsive stacking of GroEL into fibrillar structures is reported based on chemical incorporation of a spiropyran photoswitch that is anchored in the apical domain of GroEL.^[27-29] The closed SP form is converted to the open MC form upon irradiation with UV light.^[28,29] In the designed system, the stacking of GroEL relies on the formation of MC–Mg²⁺–MC bridges while the SP form does not accommodate

such interaction with Mg^{2+} .^[28,29] This structure can be switched back upon irradiation with visible light and in the presence of a radical scavenger.^[28]

Figure 6.1 TM-Encapsulin and GroEL proteins. a) Mass spectrum of co-purified proteins indicating the presence of two species, which are the GroEL ($m/z \sim 12500$ kDa) and TM-Enc ($m/z \sim 17500$ kDa). b) Structure of TM-Enc (red, PDB: 3DKT) and GroEL (yellow, PDB: 1SS8) and the visualization of the particles in their native state (in buffer Tris-HCl pH 7.5) using TEM.

We anticipate that the decrease of pH upon irradiation of photoacid MCH^+ in water would induce association of protein molecules into larger species (Figure 6.2a). The pH of the water-based system containing the proteins ($4.3 \mu M$, TM-Enc:GroEL $\sim 5\%:95\%$) and the photoacid MCH^+ (0.17 mM, 2% ethanol) decreases from 5.90 to 4.05 upon irradiation with visible light ($\lambda \geq 420$ nm, 30 s). Based on the absorbance at $\lambda_{max} = 421$ nm (Figure 6.2b), we achieve $\sim 45\text{--}50\%$ conversion of MCH^+ to SP upon irradiation under our experimental conditions, which does not increase further upon prolonged irradiation up to 5 min. Nevertheless, the pH drop is considered sufficient for our design as it already decreases around 2 units in 30 s, significantly changing the system from a mildly acidic/neutral to an acidic condition.

Figure 6.2 Formation of protein-based superstructures in the presence of photoacid. a) Schematic of reversible assembly of TM-Enc (red arrow) and GroEL (yellow arrow) into the giant superstructures in the presence of photoacid MCH^+ . The particles are visualized in TEM images. Upon irradiation with visible light for 30 s, the MCH^+ form releases its proton and switches to the SP form. The process is reversible in the absence of light source (i.e., incubation for 30 min in the dark). b) Chemical structures of MCH^+ and SP and the corresponding UV-visible spectra demonstrating the reversible photo-conversion of photoacid monitored at $\lambda_{\text{max}} = 421 \text{ nm}$.

Figure 6.3 Photo-switching of photoacid in the presence of proteins. a) UV-visible spectra of proteins only (Prot only) and proteins-photoacid (ProtMCH) before and after irradiation with visible light, showing a decrease of absorbance at $\lambda_{\text{max}} = 421 \text{ nm}$ upon irradiation. b) Reversion of the absorbance of photoacid to its initial state upon incubation in the dark monitored at $\lambda_{\text{max}} = 421 \text{ nm}$ for 300 s.

Importantly, the presence of the proteins in solution does not significantly affect the photo-switching of MCH^+ to SP (Figure 6.3a). Furthermore, the back process from SP to MCH^+ occurs rapidly in the dark (Figure 6.3b), leading to a complete reversibility in 30 min (Figure 6.2b).

Figure 6.4 Emergence of higher organization in a hybrid system, under irradiation with light. Formation of larger assemblies upon irradiation, based on size distribution profiles from a) dynamic light scattering and b) TEM analyses. The superstructures with defined edges (black arrow) confining smaller substructures (blue arrow), shown by c) SEM and d) TEM images. e) TEM images of dispersed superstructures upon subsequent incubation in the dark for 1 h, showing the TM-Enc (red arrow) and GroEL fragments (yellow arrow).

In the presence of photoacid, irradiation with visible light ($\lambda \geq 420$ nm) leads to the formation of larger and finite structures in solution, revealed by dynamic light scattering

(Figure 6.4a). Examination with negative-stained electron microscopy confirms the presence of large, organized, spherical superstructures with a broad distribution of sizes, ranging from 100 nm – 1.2 μm (Figures 6.2a, 6.4b-d). Based on TEM analysis, most of the circular structures are of 100 – 300 nm in diameter (Figure 6.4b). Examination with SEM reveals the presence of defined edges of the superstructures (Figure 6.4c). Such defined edges are found in unilamellar vesicles consisting of lipid bilayers.^[30] In our case, based on TEM analysis, the edges seem to consist of densely associated protein materials (Figure 6.4d). Furthermore, TEM also reveals the presence of confined smaller protein-based substructures inside the assemblies (in average ≤ 12 nm, *see* Figure 6.4d), which are less dense than the edges. The superstructures spontaneously disperse when the irradiation is stopped (Figure 6.4e). These results confirm the light-dependent, reversible formation of the giant superstructures as illustrated in Figure 6.5.

Figure 6.5 Schematic of light-fueled assembly of hybrid superstructures.

Notably, in the absence of photoacid, the protein system remains intact and does not show any response to light irradiation ($\lambda \geq 420$ nm, 30 s), confirmed by analyses with dynamic light scattering and fluorescence spectroscopy (Figure 6.6). In the absence of photoacid, when the pH is decreased from 7 to 4 using an acid salt (i.e., potassium hydrogen phthalate), the protein system in water visibly forms large aggregate (Figure 6.7). pH 4 is relatively close to the isoelectric points of both TM-Enc and GroEL (i.e., 4.89 and 4.76, respectively)^[31], at which the proteins would carry zero net charge overall. Measurement of zeta potential reveals

that the surface charge of the proteins changes from -30.4 mV to -1.99 mV upon decreasing the pH from 7 to 4. The decrease of zeta potential indicates a significant loss of surface charge at pH 4 that results in the loss of electrostatic repulsion between particles, leading to the formation of the observed aggregates (Figure 6.7).

Figure 6.6 Stability of proteins against irradiation. a) DLS-based size distributions of proteins before and after irradiation with visible light. b) Emission profiles of proteins before and after irradiation with visible light (excitation at $\lambda = 290$ nm).

Figure 6.7 Aggregation of particles in acidic environment in the absence of photoacid. a) DLS-based size distribution profile and b) TEM images of proteins at pH 4 (0.245 M KHP solution pH 4.01) in water environment, revealing the absence of any defined superstructure in the absence of photoacid.

As a control experiment, we also examine the behavior of the photoacid in the same fashion while decreasing the concentration of the proteins to be 10-fold less. In this experiment, the formation of any defined superstructure is not observed, indicating that such assemblies cannot be obtained in the lack of proteins (Figure 6.8a). Furthermore, the formation of the superstructures does not occur when the building blocks GroEL and TM-Enc are both replaced with an encapsulin from *B. linens* (BL-Enc) that is treated in the same

fashion (Figure 6.8b). BL-Enc is structurally comparable with TM-Enc in terms of symmetry and number of subunits (i.e., 60 subunits), however they are constituted of entirely different amino acid sequences and appear to differ in terms of stability in the presence of photoacid (Figure 6.8b).

Figure 6.8 Aggregation of particles in the lack or absence of GroEL/TM-Enc system. TEM images of a) highly diluted GroEL/TM-Enc system and b) BL-Enc system after irradiation with visible light in the presence of photoacid, revealing the absence of any defined superstructure.

6.2.2 Structural changes of proteins upon assembly

TEM images indicate that the proteins already undergo structural changes upon the addition of photoacid, even before any irradiation starts: the spherical, cage-like structure of TM-Enc is largely retained but seems to be more irregular in shape (Figure 6.2a) in comparison to the native state in buffer Tris-HCl pH 7.5 (Figure 6.1b). On the other hand, the circular structure of GroEL in the native state (Figure 6.1b) is disrupted into smaller fragments (Figure 6.2a). Using circular dichroism (CD) spectroscopy, we attempt to investigate whether the proteins undergo further conformational changes upon the drop in

pH that follows irradiation. The CD spectrum of the proteins-photoacid system at $\lambda = 240$ – 290 nm is modified significantly upon irradiation (Figure 6.9). This modification is reversible upon incubation in the dark (Figure 6.9). However, as the photoacid on its own also raises CD signal in the same range of wavelength, it is difficult to infer the conformation of the proteins based on CD spectra.

Figure 6.9 CD spectra of protein (Prot), photoacid (MCH) and photoacid-proteins (ProtMCH) before irradiation, after irradiation (for 30 s) and after subsequent incubation in the dark (for 30 min).

Figure 6.10 TEM images of GroEL/TM-Enc system after heating treatment at 80°C for 1.5 h to denature the GroEL fragments. Images are recorded after irradiation with visible light in the presence of photoacid, revealing the absence of any defined superstructure.

Nevertheless, as the protein component is predominantly composed of GroEL (~95%), we could anticipate that the behavior of the protein component is governed by the response of GroEL to the changes in the system. With this in mind, we proceed to denature the GroEL by heating at 80°C for 1.5 h (GroEL denatures at $65 - 70^{\circ}\text{C}$),^[32,33] whereas TM-Enc is previously reported to tolerate such treatment.^[19] In this case, addition of photoacid and

subsequent irradiation with visible light does not result in the formation of the superstructures (Figure 6.10). As the thermally treated GroEL fails to result in any superstructure even in the presence of TM-Enc in the system, we postulate that the GroEL fragments might be the key component for the assembly.

Figure 6.11 Emission profiles of photoacid-proteins after subtraction by emission profiles of photoacid only monitored at $\lambda_{\text{max}} = 340$ nm (excitation at $\lambda = 290$ nm) before irradiation, after irradiation (30 s) and after subsequent incubation in the dark (30 min).

The confined protein substructures in Figure 6.4d are smaller than the structure of cage-like TM-Enc, hinting at possible conformational changes of TM-Enc upon assembly of the superstructures. Since GroEL does not have any tryptophan residue (Trp) while TM-Enc consists of 5 Trp per monomer, we use fluorescence spectroscopy to examine whether the conformation of TM-Enc remains unchanged upon the assembly. Trp fluorescence intensity directly represents whether Trp residues are buried in the core of the protein, which indicates a folded conformation, or exposed to water environment in the case of (partially) unfolded proteins. Upon excitation at $\lambda = 290$ nm, we observe a 23% decrease in the emission at $\lambda = 330 - 350$ nm (Figure 6.11) as the response to irradiation with visible light ($\lambda \geq 420$ nm), which indicates a partial unfolding of TM-Enc that accompanies the assembly into the superstructures. Subsequent incubation in the dark for 30 min increases the emission at $\lambda = 330 - 350$ nm back to its initial state (Figure 6.11), demonstrating the reversibility of TM-Enc partial unfolding when the superstructures disperse in the dark. The result suggests that TM-Enc does undergo structural or conformational changes accompanying the assembly of the superstructures. This could be attributed to the fact that our experimental condition upon

irradiation with visible light (i.e., pH 4) is relatively close to the isoelectric pH of TM-Enc (i.e., pH 4.89), which could result in the reduced stability of the structure.

6.2.3 Emergence of light-dependent order by inclusion of photoacid in the superstructures

Although both partial unfolding and loss of charge of protein molecules are likely to induce self-association upon light irradiation, those phenomena fail to explain why the assembled superstructures adopt a curvature that leads to large circular bodies. It is evident from the control experiment in the absence of the photoacid (Figure 6.7) that the photoacid is the origin of the difference in the obtained structures. The chemical characteristics of the photoacid MCH^+ and its photoisomer SP (Figure 6.2b) prompt us to postulate that the irradiation with visible light triggers two events to occur simultaneously: first, the expected release of the proton from MCH^+ that results in the pH decrease from 5.9 to 4.0; and second, the significant increase in the hydrophobicity of the system as the closed SP form is much more hydrophobic compared to the open, protonated MCH^+ form. The SP form could be the part of the hydrophobic tail while the sulfonate group at the end of the alkyl chain could be the hydrophilic head (Figure 6.2b). Moreover, the photoacid is likely to have interacted with the protein molecules to form ordered superstructures because on its own it cannot form any defined superstructure, as evident from control experiments summarized in Figure 6.8.

In the previously described results, the proteins are first added with the photoacid MCH^+ (initial pH = 5.9) and then the mixture is irradiated with visible light that results in the formation of organized superstructures (Figure 6.4). To corroborate the possible interaction between the photoacid and the protein molecules, we switch the order of light irradiation *vs* addition of the components: the photoacid MCH^+ is first irradiated with visible light leading to the formation of SP and decreasing the pH to 4; then, the mixture of proteins is added into the formed SP. In this control experiment, we observe the formation of aggregates (Figure 6.12) that are comparable to the aggregates shown in Figure 6.7 (i.e., control experiment using an acid salt instead of the photoacid). Based on this experiment, while the proteins do aggregate as their natural response to acidic condition (i.e., pH 4), the pre-mixing of proteins and photoacid seems to be necessary to obtain the ordered structures. This result suggests

that the pre-mixing step might have allowed sufficient interaction between protein molecules and MCH^+/SP molecules to form such superstructures.

Figure 6.12 TEM images of GroEL/TM-Enc system added to a pre-irradiated photoacid solution, revealing the absence of any superstructure.

To further confirm the interaction between proteins and photoacid experimentally, we investigate the diffusion behavior of photoacid only vs photoacid-proteins using microscale thermophoresis (MST). MST measurement detects the mobility of dissolved particles as a response to microscopic temperature gradients inside a capillary (raised around 5 K in the beam center). Upon temperature gradients, the mobility of the particles inside the capillary depends on the size and hydration shell of the particles.

The photoacid is excited at $\lambda = 470$ nm and the thermophoresis is monitored based on emission at $\lambda = 520$ nm (Figure 6.13). It is important to note that the proteins do not give any fluorescence signal upon excitation at $\lambda = 470$ nm (Figure 6.13a), ensuring that the MST measurement detects the photoacid only. Furthermore, the fluorescence of the photoacid is not affected by the presence of proteins, confirmed by similar emission profiles in the absence and in the presence of proteins (Figure 6.13a). The emission at $\lambda = 485 - 600$ nm decreases upon irradiation as the MCH^+ is partially converted to SP (Figure 6.13a), which suggests that the MCH^+ is the only species that fluoresces. As shown in Figure 6.13b, MCH^+ exhibits negative thermophoresis (i.e., particles move from lower to higher temperatures), which is commonly observed for small species.^[34] Importantly, the fluorescence signals of MCH^+ in the presence of proteins are significantly lower than in the absence of proteins (Figure 6.13b), indicating an increase in size that slows down the mobility of the photoacid in the presence

of proteins. This finding suggests that the fluorescent MCH⁺ molecules readily interact with the proteins upon mixing.

Figure 6.13 Emission and thermophoresis profiles. a) Emission profiles of proteins only (Prot), photoacid only (MCH) and proteins-photoacid (ProtMCH) before and after irradiation (30 s), excitation at $\lambda = 470$ nm. b) Thermophoresis profiles of photoacid only (MCH) and proteins-photoacid (ProtMCH) before irradiation (left) and after irradiation for 30 s (right), excitation at $\lambda = 470$ nm and emission at $\lambda = 520$ nm.

Our results indicate that the superstructures obtained upon irradiation are formed as the result of interaction between proteins and photoacid. As the MCH⁺ partially switches into the more hydrophobic SP upon irradiation, the hybrid of proteins-photoacid presumably behaves like giant amphiphiles,^[35] thus assembling into circular superstructures (Figure 6.5). The interaction of the photoacid and the proteins might have been enabled by partial unfolding of both TM-Enc and GroEL that exposes their hydrophobic residues (Figure 6.5). Moreover, our results suggest the importance of GroEL fragments to form the superstructures. These fragments might correspond to the apical domain of GroEL that has been reported to enable interaction with amphiphilic molecules such as SDS.^[26] We propose that the exposure of GroEL apical domain upon addition of photoacid allows interaction with the photoacid in our system. The apical domain is also reported to be the key component in the formation of fibrillar stacks of GroEL both upon addition of SDS and upon agitation.^[20,26] While no circular higher-order assemblies have been reported for GroEL-based system, the apical domain is also likely to play a role in the formation of the superstructures due to its ability to interact with hydrophobic molecules. Further structural characterizations will allow pinpointing the interaction hotspots.

The obtained assemblies that confine smaller substructures are arguably comparable to the naturally occurring casein micelles.^[36] Structural similarities can be drawn for the building blocks as casein also possesses hydrophobic residues and flexible conformations like GroEL chaperone. Previous studies have even reported the ability of casein to function as a molecular chaperone.^[36] Both proteins also have the tendency to form fibrillar structures, although naturally casein prefers to form amorphous circular bodies (around 300 nm in diameter) confining high density nanoclusters of calcium phosphate along with channels and water-filled cavities.^[36-38]

6.3 Conclusions

We have demonstrated a dynamic order upon fueling hybrid bio-synthetic system with light. The system consists of *T. maritima* encapsulin and GroEL chaperone that are partially unfolded to accommodate the assembly process. The dynamic assembly into organized superstructures is enabled by a spiropyran-based photoacid that undergoes reversible deprotonation and structural change upon irradiation. Moreover, the photoacid molecules are likely to have interacted with the proteins to form such structures. Hence, the photo-switching of the photoacid plays a role beyond pH control by increasing the hydrophobicity of the system. In view of GroEL functioning as a chaperone and TM-Enc as a natural nanoconfinement system, the presented assemblies hold promise as a transient vessel to entrap, accumulate and store functional materials, which could be released and dispersed on-command. The on-command release feature is enabled in our system owing to the spontaneous disassembly of superstructures in the absence of energy source.

6.4 Acknowledgements

We thank Federico Lancia and Dr. Tibor Kudernac for fruitful brainstorming and discussions. We acknowledge Dr. Enrico G. Keim for assistance with TEM imaging and Dr. Joost Snijder and Prof. Albert Heck (Utrecht University) for assistance with mass spectrometry analysis and Agustin Enciso Martinez and Dr. Cees Otto for assistance with SEM analysis.

6.5 Materials and methods

6.5.1 General

TM-Enc and GroEL were recombinantly co-expressed in *E. coli* and co-purified according to the procedure in literature.^[21,22] As a control system, BL-Enc was separately expressed in *E. coli* and purified according to similar procedure. The proteins were stored in Tris-HCl buffer pH 7.5 (containing 150 mM NH₄Cl) at 4 °C and dialyzed against MilliQ water (12 – 14 kDa dialysis membrane, Spectra/Por) at room temperature prior to all experiments. The spiropyran-based photoacid MCH⁺ was purchased from Syncom (Groningen, The Netherlands) and then dissolved and stored in ethanol at room temperature. All experiments were performed in MilliQ (containing 150 mM NaCl and 2% ethanol). All chemicals were purchased from Sigma-Aldrich unless stated otherwise. All experiments were performed at room temperature unless stated otherwise.

6.5.2 Reversible formation of hybrid superstructures

500 µL of a sample solution containing proteins (4.3 µM) and photoacid (0.17 mM) in MilliQ (containing 150 mM NaCl and 2% ethanol) was prepared as the initial condition (prior to irradiation). Afterwards, the mixture was irradiated with visible light ($\lambda \geq 420$ nm, Edmund MI-150 High-intensity Illuminator, 2 mW/cm²) for 30 s. The absorption spectra were recorded with a Perkin Elmer Lambda 850 UV-visible spectrometer and the photo-switching was monitored at $\lambda_{\text{max}} = 421$ nm. For the reversion, the mixture after irradiation was incubated in the dark for 30 – 60 min and the UV-visible spectrum was recorded following the incubation. For control experiments, solutions containing photoacid only and proteins only were separately prepared and treated in the same fashion. Due to the absence of photoacid, the pH of the solution containing proteins only was decreased using a commercial acid salt, potassium hydrogen phthalate (0.245 M KHP solution pH 4.01, Certipur®). The pH of all system was measured using a SartoriusTM pH meter (Thermo Fisher Scientific). Zeta potential of proteins was measured using a Zetasizer Nano (Malvern).

6.5.3 Characterization of hybrid superstructures

The assembled superstructures were characterized using dynamic light scattering (DLS), negative-stained transmission electron microscopy (TEM) and scanning electron microscopy (SEM). For DLS analyses, 500 μL of a proteins-photoacid sample was prepared and measured using a Nanotracs Wave, Microtracs particle analyzer. For control experiments, a solution containing only the proteins was prepared and treated in the same fashion. For electron microscopy analyses, samples (5 μL) were applied onto Formvar-carbon coated grids. After 30 s, the excess of liquid was drained. Uranyl acetate (5 μL , 1% w/v) was added for negative staining and the excess of liquid was drained after 1 min and air-dried for 1 h at room temperature. For TEM and SEM respectively, the samples were examined using FEG-TEM (Phillips CM 30) and JSM-6610LV-SEM (JEOL). For SEM analysis, the grid was placed on top of a stainless steel substrate (μRIM) prior to measurement.

6.5.4 Characterization of structural changes of proteins

The structure of proteins was analyzed based on examination of negative-stained TEM images, circular dichroism (CD) spectra and fluorescence spectra. The CD spectra were recorded for 500 μL of sample using a Jasco J-1500 circular dichroism spectrometer at $\lambda = 190 - 350$ nm. The emission spectra of 500 μL of sample at $\lambda = 300 - 440$ nm (excitation at $\lambda = 290$ nm) were recorded using a Perkin Elmer fluorescence spectrometer. For control experiments, solutions containing photoacid only and proteins only were separately prepared and treated in the same fashion.

6.5.5 Analysis of photoacid inclusion in the superstructures

The interaction between photoacid and proteins was inferred based on fluorescence spectroscopy and microscale thermophoresis (MST) analyses. The emission spectra of 500 μL of sample at $\lambda = 485 - 600$ nm (excitation at $\lambda = 470$ nm) were recorded using a Perkin Elmer fluorescence spectrometer. For control experiments, solutions containing photoacid only and proteins only were separately prepared and treated in the same fashion. For MST analyses, around 3 – 5 μL of each sample was loaded into a hydrophilic glass capillary and the fluorescence-based thermophoresis profiles for all samples were recorded using a

Monolith NT.115 Blue/Green Nanotemper (excitation at $\lambda = 470$ nm, emission at $\lambda = 520$ nm). The microscopic gradient of temperature inside the capillary was generated by a built-in IR-laser module ($\lambda = 1475$ nm at 120 mW laser power). For all measurements, the IR laser was switched on at $t = 5$ and switched off at $t = 30$ s, corresponding to Figure 6.13b. All measurements were performed at room temperature of approximately 20°C.

6.6 References

- [1] C. Conde, A. Caceres, *Nat. Rev. Neurosci.* **2009**, *10*, 319-332.
- [2] H. Hess, J. L. Ross, *Chem. Soc. Rev.* **2017**.
- [3] B. A. Grzybowski, W. T. Huck, *Nat. Nanotechnol.* **2016**, *11*, 585-592.
- [4] M. Uchida, B. LaFrance, C. C. Broomell, P. E. Prevelige, Jr., T. Douglas, *Small* **2015**, *11*, 1562-1570.
- [5] M. A. Kostiaainen, O. Kasyutich, J. J. L. M. Cornelissen, R. J. M. Nolte, *Nat. Chem.* **2010**, *2*, 394-399.
- [6] C. C. Broomell, H. Birkedal, C. L. P. Oliveira, J. S. Pedersen, J. A. Gertenbach, M. Young, T. Douglas, *Soft Matter* **2010**, *6*, 3167-3171.
- [7] R. Klajn, K. J. Bishop, B. A. Grzybowski, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10305-10309.
- [8] G. Ragazzon, M. Baroncini, S. Silvi, M. Venturi, A. Credi, *Nat. Nanotechnol.* **2015**, *10*, 70-75.
- [9] J. Boekhoven, W. E. Hendriksen, G. J. M. Koper, R. Eelkema, J. H. van Esch, *Science* **2015**, *349*, 1075-1079.
- [10] R. M. Putri, J. J. L. M. Cornelissen, M. S. T. Koay, *ChemPhysChem* **2015**, *16*, 911-918.
- [11] Z. Shi, P. Peng, D. Strohecker, Y. Liao, *J. Am. Chem. Soc.* **2011**, *133*, 14699-14703.
- [12] L. Florea, K. Wagner, P. Wagner, G. G. Wallace, F. Benito-Lopez, D. L. Officer, D. Diamond, *Adv. Mater.* **2014**, *26*, 7339-7345.
- [13] P. K. Kundu, D. Samanta, R. Leizrowice, B. Margulis, H. Zhao, M. Borner, T. Udayabhaskararao, D. Manna, R. Klajn, *Nat. Chem.* **2015**, *7*, 646-652.
- [14] C. Fu, J. Xu, C. Boyer, *Chem. Commun.* **2016**, *52*, 7126-7129.
- [15] C. Maity, W. E. Hendriksen, J. H. van Esch, R. Eelkema, *Angew. Chem. Int. Ed.* **2015**, *54*, 998-1001.

- [16] T. Satoh, K. Sumaru, T. Takagi, T. Kanamori, *Soft Matter* **2011**, 7, 8030-8034.
- [17] Y. Luo, C. M. Wang, P. Peng, M. Hossain, T. L. Jiang, W. Fu, Y. Liao, M. Su, *J. Mater. Chem. B* **2013**, 1, 997-1001.
- [18] O. S. Shafaat, J. R. Winkler, H. B. Gray, D. A. Dougherty, *ChemBioChem* **2016**, 17, 1323-1327.
- [19] C. Cassidy-Amstutz, L. Oltrogge, C. C. Going, A. Lee, P. Teng, D. Quintanilla, A. East-Seletsky, E. R. Williams, D. F. Savage, *Biochemistry* **2016**, 55, 3461-3468.
- [20] J. Chen, H. Yagi, P. Sormanni, M. Vendruscolo, K. Makabe, T. Nakamura, Y. Goto, K. Kuwajima, *FEBS Lett.* **2012**, 586, 1120-1127.
- [21] M. Sutter, D. Boehringer, S. Gutmann, S. Gunther, D. Prangishvili, M. J. Loessner, K. O. Stetter, E. Weber-Ban, N. Ban, *Nat. Struct. Mol. Biol.* **2008**, 15, 939-947.
- [22] R. M. Putri, J. W. Fredy, J. J. L. M. Cornelissen, M. S. T. Koay, N. Katsonis, *ChemPhysChem* **2016**, 17, 1815-1818.
- [23] A. L. Horwich, G. W. Farr, W. A. Fenton, *Chem. Rev.* **2006**, 106, 1917-1930.
- [24] M. Mayhew, A. C. da Silva, J. Martin, H. Erdjument-Bromage, P. Tempst, F. U. Hartl, *Nature* **1996**, 379, 420-426.
- [25] M. Hayer-Hartl, A. Bracher, F. U. Hartl, *Trends Biochem. Sci.* **2016**, 41, 62-76.
- [26] J. Chen, H. Yagi, Y. Furutani, T. Nakamura, A. Inaguma, H. Guo, Y. Kong, Y. Goto, *Sci. Rep.* **2014**, 4, 5614.
- [27] S. Biswas, K. Kinbara, T. Niwa, H. Taguchi, N. Ishii, S. Watanabe, K. Miyata, K. Kataoka, T. Aida, *Nat. Chem.* **2013**, 5, 613-620.
- [28] T. Sendai, S. Biswas, T. Aida, *J. Am. Chem. Soc.* **2013**, 135, 11509-11512.
- [29] S. Biswas, K. Kinbara, N. Oya, N. Ishii, H. Taguchi, T. Aida, *J. Am. Chem. Soc.* **2009**, 131, 7556-7557.
- [30] G. Basanez, M. B. Ruiz-Arguello, A. Alonso, F. M. Goni, G. Karlsson, K. Edwards, *Biophys. J.* **1997**, 72, 2630-2637.
- [31] S. K. Molugu, J. Li, R. A. Bernal, *J. Chromatogr. B* **2015**, 1007, 93-99.
- [32] J. Martin, A. L. Horwich, F. U. Hartl, *Science* **1992**, 258, 995-998.
- [33] B. Sot, S. Banuelos, J. M. Valpuesta, A. Muga, *J. Biol. Chem.* **2003**, 278, 32083-32090.
- [34] B. Gotsmann, U. Durig, *Appl. Phys. Lett.* **2005**, 87.
- [35] Q. Luo, C. Hou, Y. Bai, R. Wang, J. Liu, *Chem. Rev.* **2016**, 116, 13571-13632.

- [36] C. Holt, J. A. Carver, H. Ecroyd, D. C. Thorn, *J. Dairy Sci.* **2013**, *96*, 6127-6146.
- [37] R. Trejo, T. Dokland, J. Jurat-Fuentes, F. Harte, *J. Dairy Sci.* **2011**, *94*, 5770-5775.
- [38] D. G. Dalgleish, P. A. Spagnuolo, H. D. Goff, *Int. Dairy J.* **2004**, *14*, 1025-1031.

Summary

Living materials are based on proteins that adapt and change in structure and function continuously when in use; cellular microtubules, ATP synthases and ribosomes are but a few examples. Breathing life into semi-synthetic materials would allow improved understanding over protein regulation and dynamics, and how their integration into complex systems can lead to emergent functions across length scales. The focus of my research is to achieve optical control over functional proteins by connecting them to artificial molecular switches, with the aim to amplify molecular switching events across length scales and gain understanding over cooperative and systematic regulations of proteins. We choose light because it offers spatiotemporal selectivity, is compatible with a wide range of phases and relatively non-destructive towards protein systems. Interfering with mechanisms such as allosteric communication or hierarchical self-assembly not only paves the way towards an improved understanding of cooperative or collective effects in living matter, but it is also associated with the generation of new classes of smart bio-hybrids. The works presented in this thesis involve diverse functional proteins, ranging from an allosteric transport protein, the human serum albumin (Chapter 3), to proteins forming cages (Chapter 4, 5, 6) all the way to chaperone proteins that are known to assist cellular protein folding (Chapter 6).

The work focuses on the use of spiropyran derivatives as photo-switches, because their irradiation induces not only a change in conformation, but also it yields a zwitterionic form that is bound to influence a protein environment. Chapter 3 and 5 demonstrate covalent incorporation of spiropyran switches into a cysteine and lysines that are present in the studied proteins. The covalent coupling involves reacting maleimide or succinimide moieties (attached to the photo-switchable units) with the proteins via thiol-Michael addition or amine-NHS ester reaction, respectively. Chapter 3 describes how such structural modification allows photo-controlling the natural allosteric regulation of the human serum albumin (HSA), and thus to modulate ligand binding by irradiation with light. On the other hand, modification of a bacterial encapsulin, a self-assembled protein cage that confines

Summary

smaller proteins in its cavity (Chapter 4 and 5), yields local modifications only, and does not compromise the structural integrity and the behavior of the entire assembly.

Chapter 4 and 5 further confirm and demonstrate that bacterial encapsulins are robust, catalytically active and biocompatible. They can also be labeled with light-switchable fluorophores on their lysines via reaction with succinimide-bearing spiropyran switches, and thus they are good candidates as smart nanoplateforms for *in vitro* and *in vivo* research purposes.

Chapter 6 demonstrates the light-fueled assembly of proteins by adding photo-switches into a medium containing the encapsulin and GroEL chaperone. This strategy enables addition of a large amount of photo-switches, which, in this context, results in the changes of the medium upon irradiation, i.e., pH and hydrophobicity. Mediated by a spiropyran-based photoacid, the described approach leads to an assembly of transient superstructures that depend on light irradiation to prevent disassembly. This assembly thus presents a dynamic and reversible order of a protein system that is sustained by a continuous input of energy, and this mechanism is reminiscent of the GTP-fueled assembly of cellular microtubules.

Overall, we have developed photo-responsive protein-based hybrids in which dynamic regulations such as allostery and self-assembly are prominent to protein functionality. Via incorporation of photo-switches, optical control has been demonstrated across length scales, from a monomeric structure to highly defined multimeric architectures made of proteins. The presented research provides a means to interfere with dynamic regulations of proteins and supports strategies towards the development of biocompatible and smart materials.

Samenvatting

Levende materialen zijn gebaseerd op eiwitten die zich continu aanpassen en van structuur veranderen wanneer ze actief zijn; cellulaire microtubuli, ATP synthases en ribosomen zijn slechts enkele voorbeelden. Het door mensen gemaakte materialen levend maken kan zorgen voor een verbeterd begrip van eiwitregulatie en -dynamiek en meer kennis over hoe de integratie van complexe systemen kan leiden tot nieuwe functies over lengte schalen. De focus van mijn onderzoek is om optische controle te verkrijgen over functionele eiwitten door ze te verbinden met artificiële moleculaire schakelaars. Het doel is om moleculaire schakelingen te versterken over lengte schalen en meer kennis te vergaren over de coöperatieve en systemische regulering van eiwitten. We hebben gekozen voor licht omdat het spatio-temporele selectiviteit biedt, compatibel is met een groot scala aan fases en het relatief non-destructief is met betrekking tot eiwitten. Het interfereren met mechanismen zoals allosterische communicatie of hiërarchische zelfassemblage zet de deur open voor een verbeterd begrip van coöperatieve of collectieve effecten in levende materialen, maar het is ook geassocieerd met de vorming van nieuwe klassen van slimme bio-hybriden. Het werk dat in deze thesis wordt gepresenteerd betreft diverse functionele eiwitten, reikend van allosterische transport eiwitten, albumine van humaan serum (Hoofdstuk 3), via eiwitten die kooien vormen (Hoofdstuk 4, 5, 6), tot chaperonne-eiwitten die bekend staan te assisteren in cellulaire eiwit vouwing (Hoofdstuk 6).

Dit werk focust op het gebruik van spiropyran derivaten als foto-schakelaars, omdat de bestraling ervan niet alleen een verandering in conformatie induceert, maar ook omdat het een zwitterion oplevert dat de eiwit omgeving zal beïnvloeden. Hoofdstuk 3 en 5 demonstren covalente incorporatie van spiropyran schakelaars met een cysteine en lysines die aanwezig zijn in de bestudeerde eiwitten. De covalente koppeling werkt door maleimide of succinimide groepen (die verbonden zijn met de foto-schakelaars) te laten reageren met eiwitten via thiol-Michael additie of een amine-NHS ester reactie, respectievelijk.

Hoofdstuk 3 beschrijft hoe dergelijke structurele modificaties foto-controle van de natuurlijke allosterische regulering van albumine uit humaan serum (HSA) toestaan en

zodoende ligand binding moduleren via bestraling met licht. Anderzijds, modificatie van bacteriële encapsulines, zelfassemblerende eiwit kooien die een kleiner eiwit inkapselen (Hoofdstuk 4 en 5), resulteert slechts in lokale modificaties en heeft geen invloed op de structurele integriteit en het gedrag van het gehele construct.

Hoofdstuk 4 en 5 bevestigen en demonstreren verder dat bacteriële encapsulines robuust, katalytisch actief en biocompatibel zijn. Ze kunnen ook gelabeld worden met licht-geschakelde fluoroforen die zijn gebonden aan lysines door middel van een reactie met spiropyran schakelaars die een succinimide-groep hebben. Zodoende zijn het geschikte kandidaten om te dienen als slimme nanoplatformen voor *in vitro* en *in vivo* onderzoek.

Hoofdstuk 6 demonstreert de licht-gedreven assemblage van eiwitten via foto-schakelaars in een medium dat encapsulines en de chaperonne GroEL bevat. Deze strategie laat de toevoeging van grote hoeveelheden foto-schakelaar toe, welke, in deze context, resulteert in veranderingen in het medium na bestraling, zoals veranderingen in pH en hydrofobiciteit. Terwijl het gemedieerd wordt door een spiropyran-gebaseerde fotozuur, leidt de beschreven methode tot een assemblage van vesikel-achtige superconstructen die afhankelijk zijn van bestraling met licht om te voorkomen dat ze uiteen vallen. Deze vorm van assemblage presenteert een dynamische en reversibele orde van een eiwit systeem dat wordt onderhouden door een continue invoer van energie, en dit mechanisme doet denken aan de GTP-gevoede assemblage van cellulaire microtubuli.

Concluderend hebben we licht-responsieve eiwit-gebaseerde hybriden gemaakt waarin dynamische regulatie zoals allosterie en autoassemblage prominent zijn over eiwit functionaliteit. Door middel van de incorporatie van foto-schakelaars, optische controle is aangetoond over lengte schalen, van een monomerische structuur tot zeer gedefinieerde multimerische structuren die bestaan uit eiwitten. Het hier gepresenteerde onderzoek geeft een manier om te interfereren met de dynamische regulatie van eiwitten en het ondersteunt strategieën voor de ontwikkeling van biocompatibele en slimme materialen.

Acknowledgements

With the completion of this dissertation, I would like to take the opportunity to thank everyone who has supported me through my PhD study and my stay in The Netherlands. First of all, I am grateful to my supervisors, **Jeroen** and **Nathalie**, for welcoming me into their research groups, providing guidance and facilities to conduct scientific works and develop my skills and giving me the opportunities to join a lot of international conferences. Jeroen, I was a Master student in your group first and afterwards a PhD student, it was a great pleasure to have been a part of the BNT group in these past five years. Looking back, I am grateful that all this time I have been given immense opportunity and research experiences that have undoubtedly shaped my scientific life and helped me become a better researcher and scientist. Nathalie, thank you for introducing me to the fascinating world of molecular switches, it was a new field to me and I am happy to have learned a lot in the past years and to have conducted interesting projects on the hybrids of proteins and photo-switches. I greatly appreciate that you welcomed me into your research group and gave me a lot of opportunities, guidance and support during my study. I would also like to express my sincere gratitude to **Melissa** for your guidance, support and advice. It was a great pleasure to have worked with you and thank you for participating in my graduation committee. Furthermore, I would like to sincerely thank all my committee members: **Prof. Andrew Woolley** from the University of Toronto, **Prof. Pepe Castón** from CSIC Madrid, **Prof. Mireille Claessens**, **Prof. Jurriaan Huskens** and **Prof. Herman ten Kate** from the University of Twente for reading my manuscript and the opportunity of having my PhD defense under your consideration.

I would like to express my gratitude to the staff of MnF/BNT cluster: **Tibor**, **Jurriaan**, **Pascal**, **Jos** and **Wim** for scientific advice, suggestions and discussions during the colloquium and meetings. Tibor, thank you for your suggestions and fruitful discussion regarding the final chapter of my thesis. I would like to thank **Nicole** and **Izabel** for all your kind help regarding the many arrangements of schedules, conferences and especially regarding my scholarship and administration as an international student that might have been complicated at times. **Marcel**, **Regine**, **Richard** and **Bianca**, thank you for all your kind help and support in the labs and the nice chats that we had.

Acknowledgements

Furthermore, I am very grateful to my scholarship provider, **Indonesia Endowment Fund of Education** or Lembaga Pengelola Dana Pendidikan (LPDP) who has supported my PhD study throughout these four years and to **Prof. Zeily Nurachman** from Institut Teknologi Bandung (ITB), Indonesia for your recommendation and scientific training that I received while performing my Bachelor and Master study at ITB under your supervision.

The accomplishment of this dissertation would have not been possible without the fruitful collaborations with the research group of Prof. Pepe Castón from CSIC Madrid, **Habiburrahman** and **Prof. Claudia Filippi** from MESA+ University of Twente, **Agustin** and **Dr. Cees Otto** from MIRA University of Twente. Thank you and it was an honor to have the opportunity to collaborate with your teams. Furthermore, I would like to thank fellow students and postdocs in the BNT/MnF cluster who have worked together and collaborated with me on my projects: **Wilfried, Federico, Robin, Aijie, Christoph, Mark de R., Alberto** and my Master student, **Katerina**. I am very thankful for your help, suggestions and discussion and I wish you all the best with your study and future plans. Robin, thank you for your help with my Samenvatting. I would like to thank **Dr. Rico Keim** for your help with TEM measurements and the nice chats about Indonesia as your family used to live in my hometown.

Aijie and Wilfried, thank you for being my paranymphs and for all your support during my PhD, thank you for not only being scientifically involved in my projects, but more so, for being such good friends to me. Aijie, we started about the same time and I am happy that we could finish our study also about the same time. In these past four years, suffice to say that we have been through a lot together and that we have helped each other out throughout. I sincerely wish you all the best towards the completion of your PhD and for your future plans. Wilfried, thank you for always helping me out and cheering me up, you really are a funny and supportive friend, in addition to a smart scientist. I wish you all the best with your future plans and I am sure you will succeed (I see it..!) I would also like to sincerely thank **Meow (Supa), Xiu, Yao, Jenny (Shuqin), Liulin, He** and **Wang Yi** for your help, suggestions, funny and nice chats, laughter and jokes, and of course nice foods! I cherish the moments we shared together in Enschede (and beyond!) and I sincerely hope that we could meet again

someday, perhaps in a different part of the world ☺ I wish you all the best with your study and your future plans. Yao, special thanks to you for giving me a Chinese name!

Melanie, you were the first person that I met from the group and I am grateful for all the time we worked together and your supervision during my Master study. **Supitch**, thank you for being such a good friend and for your help and advice during my stay in Enschede. **Rianne** and **Stan**, thank you for your many suggestions and feedback during colloquium and meetings, I really appreciate it, and also thank you for the nice chats outside the labs. **Gülistan**, I admire your hard work, thank you for the nice chats that we had and I wish you all the best for your study and research. **Wouter**, we went to the Global Young Scientists Summit in Singapore together and it was a nice experience, thank you for the nice chats that we had and I wish you all the best for your study. I would like to thank all the members and former members of BNT/MnF cluster who have made my stay in Twente really enjoyable and for your help in so many different ways: **Raquel, Emanuela, Sarah, Rik, Alexis, Ben, Tushar, Shirish, Jenny, Bettina, Janneke, Jasper, Milou, Roberto, Alejandro, Laura, Pichayut, Fern, Anne-Déborah, Alexander R., Mark V., Maike, Almudena, Daniele, Jacopo, Sara, Andrea, Erik, Alexander M., Nico, Gaurav, Pramod, Nicole, Pia, Ines, Sander, Dhanya, Salmeen** and fellow students in the BNT/MnF cluster.

Being far away from home can be difficult at times, but the Indonesian community in Enschede has truly provided me a second home. I am very grateful to my Indonesian friends who have supported me throughout, as it would have been difficult had I gone through this journey alone. I have met a lot of people from Indonesia in these past five years in Enschede, I forever cherish the moment we shared together and I am thankful to have made new friends and family half the world away from home. For those who are still in the middle of their study, I sincerely wish you all the very best towards the finish line and remember to never give up.

Finally and especially, I would like to thank my beloved family and friends, without whom I would not have made it this far. I am sincerely grateful for your constant support and never-ending encouragement. **Mami, Papi, Pranata** dan **Tiwi**, alhamdulillah rabbil'alam, akhirnya selesai sudah perjuangan selama lima tahun ini, terimakasih atas dukungan,

Acknowledgements

semangat dan do'a yang mengalir tak henti-hentinya. Semoga ilmu ini kiranya berkah dan kita semua selalu di dalam rahmat dan lindungan Allah SWT.

Enschede, September 2017

Rindia Maharani Putri

About the author

Rindia Maharani Putri was born on May 28th 1991 in Padang, Indonesia. She pursued her bachelor degree in Chemistry at Institut Teknologi Bandung, Indonesia, and graduated in July 2011. During her bachelor project, she studied the isolation and characterization of lipase under the supervision of Prof. Zeily Nurachman. Afterwards, in January and April 2013, she received her master degrees in Chemical Engineering at the University of Twente and in Chemistry at Institut Teknologi Bandung, respectively. Her master research, entitled “Synergistic Biocatalysts inside Cowpea Chlorotic Mottle Virus-based Nanoreactors” was carried out in the Biomolecular Nanotechnology group, University of Twente under the supervision of Prof. Jeroen Cornelissen, Dr. Melissa Koay and Dr. Melanie Brasch.

Since September 2013 she started her PhD program in the Biomolecular Nanotechnology group under the supervision of Prof. Jeroen Cornelissen, Prof. Nathalie Katsonis and Dr. Melissa Koay. The aim of her project is to engineer proteins into dynamic building blocks for light-responsive hybrid (semi-synthetic) systems. The results of this project are described in this thesis.

List of publications

1. **Rindia M. Putri**, Jeroen J.L.M. Cornelissen and Melissa S.T. Koay. Self-assembled Cage-like Protein Structures, *ChemPhysChem* **2015**, *16*, 911-918.
2. **Rindia M. Putri**, Jean Wilfried Fredy, Jeroen J.L.M. Cornelissen, Melissa S.T. Koay and Nathalie Katsonis. Labelling Bacterial Nanocages with Photo-switchable Fluorophores, *ChemPhysChem* **2016**, *17*, 1815-1818.
3. Liulin Yang, Aijie Liu, Shuqin Cao, **Rindia M. Putri**, Pascal Jonkheijm, Jeroen J.L.M. Cornelissen. Self-assembly of Proteins: Towards Supramolecular Materials, *Chem. Eur. J.* **2016**, *22*, 15570–15582.
4. **Rindia M. Putri**, Mark V. de Ruiter, Jeroen J.L.M. Cornelissen. CCMV-like Particles as Enzymatic Nanoreactors, *in press*. "Virus-derived Nanoparticles for Advanced Technologies", in the lab protocol series 'Methods in Molecular Biology', G. Lomonossoff and C. Wege, Eds.; Humana Press: Springer Science + Business Media New York, USA.
5. Melanie Brasch, **Rindia M. Putri**, Mark V. de Ruiter, Daniel Luque, Melissa S.T. Koay, José R. Castón and Jeroen J.L.M. Cornelissen. Assembling Enzymatic Cascade Pathways Inside Virus-based Nanocompartments Using Dual-tasking Nucleic Acid, *J. Am. Chem. Soc.* **2017**, *139*, 1512-1519.

Publications in preparation

6. **Rindia M. Putri**, Daniel Luque, Robin Klem, Aikaterini A. Rousou, Aijie Liu, Christoph H.-H. Traulsen, W. Frederik Rurup, Melissa S.T. Koay, José R. Castón and Jeroen J.L.M. Cornelissen. Structural Characterization of Native and Modified Encapsulins as Nanoplatforms for Catalysis and Cellular Uptake, *submitted to ACS Nano*.
7. **Rindia M. Putri**, Federico Lancia, Agustin Enciso Martinez, Cees Otto, Melissa S.T. Koay, Tibor Kudernac, Jeroen J.L.M. Cornelissen and Nathalie Katsonis.

Opportunistic Behavior of Artificial Switches in Protein Environment, *in preparation*.

8. **Rindia M. Putri**, Habiburrahman, Jean Wilfried Fredy, Pichayut Tananchayakul, Jeroen J.L.M. Cornelissen, Melissa S.T. Koay, Claudia Filippi and Nathalie Katsonis. Light-programmed Allostery in an Engineered Human Serum Albumin and Mechanistic Insight from Molecular Dynamics, *in preparation*.
9. **Rindia M. Putri**, Jeroen J.L.M. Cornelissen and Nathalie Katsonis. Strategies to Modulate Allosteric Regulation and Self-assembly of Proteins using Light, *in preparation*.
10. Aijie Liu, Wei Zhu, Mark V. de Ruiter, Stan J. Maassen, **Rindia M. Putri**, Liulin Yang, C. Jeffrey Brinker and Jeroen J.L.M. Cornelissen. Interfacial Crosslinking of Viral Protein Cage-inspired Free-standing Thin Films with Tunable Functionalities, *in preparation*.