João Paulo Palma Jacinto

Mestre em Engenharia de Micro e Nanotecnologias

Exploring dynamic structures of bionanocages

Dissertação para obtenção do Grau de Doutor em
Biofísica e Bioquímica das Radiações, Especialidade em Biofísica

Orientador: Prof. Doutora Alice S. Pereira,
Professora Associada,
Faculdade de Ciências e Tecnologias – Universidade Nova de Lisboa

Co-orientador: Prof. Doutor Pedro Tavares,
Professor Auxiliar,
Faculdade de Ciências e Tecnologias – Universidade Nova de Lisboa

Júri:

Presidente: Prof. Doutor Paulo Manuel Assis Loureiro Limão-Vieira, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Arguentes: Prof. Doutor Cedrick Dicko, Lund University, Suécia

Doutor Enrique Carbó-Argibay, International Iberian Nanotechnology Laboratory, Portugal

Vogais: Prof. Doutor Ilko Bald, Potsdam University, Alemanha

Prof. Doutor Paulo Manuel Assis Loureiro Limão-Vieira, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Doutora Isabel Cristina da Costa Garcia Timóteo, ITQB Universidade NOVA de Lisboa

Prof. Doutora Alice S. Pereira, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

junho 2021

Universidade Nova de Lisboa Faculdade de Ciências e Tecnologia

João Paulo Palma Jacinto

Mestre em Engenharia de Micro e Nanotecnologias

Exploring dynamic structures of bionanocages

Dissertação para obtenção do Grau de Doutor em Biofísica e Bioquímica das Radiações, Especialidade em Biofísica

junho 2021

Exploring dynamic structures of bionanocages

To my family

Acknowledgements

The work presented in this Thesis was in its vast majority performed in the Molecular Biophysics group (REQUIMTE/UCBIO, FCT-UNL). I would like to thank everyone who contributed either directly or indirectly to this work, and made it possible.

I would like to express my sincere gratitude to my supervisors, Professor Alice S. Pereira and Professor Pedro Tavares, for the opportunity that was to perform this work under their guidance. I can not thank enough all the scientific advice through these years. I am grateful for the time, help, encouragement and for the terrific contribution to my personal and scientific development. This, I am sure would not have happened without their presence in my life. The memories will be deeply cherished, thank you!

To Dr Ilko Bald (Hybrid Nanostructures Research Group, Institut für Chemie - Physikalische Chemie, Potsdam Universität) for the warm and friendly welcome to his research group in order to perform part of this work. For all the help with the origami experiments and data interpretation. It was a pleasure to work alongside him and his group.

I also thank Dr Søren Vrønning Hoffmann and Dr Nykola Jones (Department of Physics and Astronomy - Centre for Storage Ring Facilities, Aarhus University) for all the guidance with data acquisition and interpretation of the SRCD experiments, and for making me feel welcomed during my stay in Aarhus.

I would also like to thank Professor Paulo Limão-Vieira for the scientific discussions with regards to CD and for making all the beam times extremely enjoyable.

I also express thanks to Dr Haydyn Mertens from EMBL – BioSaxs P12 beamline for the help and guidance with beamtime proposals submission and subsequent acquisition procedures and treatment of the SAXS data featured in this thesis.

I thank Dr César Laia (Photochemistry Group REQUIMTE/UCBIO, FCT-UNL) for the access to DLS equipment and also the discussions regarding the technique.

Many thanks to the Radiation Biology and Biophysics Doctoral Training Programme (RaBBiT PD/00193/2012) and to Fundação para a Ciência e Tecnologia, Ministério da Educação e Ciência (FCT-MCES) for granting me the PhD fellowship (SFRH/BD/135056/2017). I also thank the STSM funding by COST Action (CA15126 MOBIEU), that enabled part of this work.

I would like to show my deepest appreciation towards my friends from the Molecular Biophysics Group, namely Ana Almeida, Daniela Penas, João Guerra, Nídia Almeida, Nuno

Coelho and Raquel Pacheco. Work would have been much harder without your support, thank you for all the good times.

At last, I sincerely thank my family and closest ones. This thesis is also a result of their support and continuous presence in my life. Their encouragement and reassurance were vital to my well-being and I am grateful for not having made this journey alone. Thank you!

Abstract

Mini-ferritins also known as Dps (DNA-binding protein from starved cells) belong to the ferritin family. They are composed of 12 identical subunits forming a hollow quasi spherical structure with ~ 9 nm outer diameter and 5 nm inner diameter. They can store up to 500 iron atoms per shell, being a vital piece in iron homeostasis. These proteins are able to oxidize Fe^{2+} ions using either H_2O_2 or O_2, store it as a ferric mineral and release it in the ferrous form to the medium, when needed for various cellular processes. Dps proteins are also known to bind DNA thus protecting it from hazardous factors.

In this Thesis the oligomeric states of Dps from *Marinobacter (M.) hydrocarbonoclasticus* were studied under different ionic strength conditions, in order to take advantage of the system for biotechnological uses. The present work has shown that the oligomerization is a dynamic process that can be controlled by the ionic strength and that the N-terminal tail is necessary for the dodecameric assembly. The DNA binding activity of Dps was also explored using DNA origami technology, which indicated that binding activity seems to occur in a wide range of pH values, being however not clear if there is a preferential binding to single or double stranded DNA.

The biotechnological potential and versatility of this system to be employed as nanoreactor or a building block for the synthesis of novel enhanced supramolecular polymers were also demonstrated. The possibility to synthesise different metal nanoparticles/materials with constrained size inside the protein nanocage, using a simple procedure was proven.

Keywords: Dps; oligomerization; DNA binding; structural dynamics; synchrotron radiation circular dichroism; supramolecular protein polymer.

Resumo

As mini-ferritinas também conhecidas como Dps (DNA-binding protein from starved cells) pertencem à família proteica das ferritinas. São compostas por 12 subunidades idênticas que formam uma estrutura quase esférica com cerca de 9 nm de diâmetro externo e 5 nm de diâmetro interno. Conseguem armazenar até 500 átomos de ferro por proteína, sendo uma peça vital no processo homeostático de ferro. Estas proteínas são capazes de oxidar Fe^{2+} através do uso de H_2O_2 ou O_2, consequentemente são também capazes de mineralizar e libertar o ferro na sua forma ferrosa para o meio exterior, quando necessário para vários processos celulares. A Dps é também conhecida pela sua capacidade de ligar DNA e dessa forma protegê-lo de fatores nocivos.

Neste trabalho os estados oligoméricos da Dps proveniente de *Marinobacter (M.) hydrocarbonoclasticus* foram estudados em diferentes condições de força iónica de modo a permitir a utilização biotecnológica do sistema. É evidenciado que o estado oligómerico da Dps pode ser controlado através da otimização da força iónica do tampão, e também a necessidade da existência da região N-terminal para a existência do decâmero. A atividade de ligação ao DNA foi explorada com recurso à tecnologia de DNA origami, indicando que a ligação ocorre numa gama extensa de valores de pH, não sendo clara uma preferência a cadeias simples ou duplas de DNA.

O potencial biotecnológico e a versatilidade deste sistema para ser aplicado enquanto reator ou bloco de construção para a síntese de novas estruturas poliméricas supramolecurares foi também demonstrada. A possibilidade de sintetizar diferentes nanopartículas/materiais metálicos de tamanho limitado no interior da esfera proteica utilizando um procedimento simples foi provado.

Palavras-chave: Dps; oligomerização; ligação a DNA; dinâmica estrutural; radiação de sincrotrão dicroísmo circular; redes poliméricas supramoleculares.

Contents

List of figures

List of tables

List of Abbreviations and Symbols

χ^2 – Chi-squared test

ΔH – Enthalpy

AFM – Atomic force microscopy

Bfr – Bacterioferritin

bp – Base pair

CD – Circular dichroism

Cryo-EM – Cryogenic electron microscopy

Da – Dalton

DLS – Dynamic light scattering

Dps – DNA-binding protein from starved cells

DSC – Differential scanning microscopy

dsDNA – Double stranded DNA

DTT – Dithiothreitol

EDTA – Ethylenediamine tetraacetic acid

EMSA – Electrophoretic mobility shift assay

FOCs – Ferroxidase centres

FPLC – Fast protein liquid chromatography

Ftn – Classical ferritin

IPTG – Isopropyl β-D-1-thiogalactopyranoside

LAP – Lithium phenyl-2,4,6-trimethylbenzoylphosphinate

LB – Luria-Bertani medium

LMW – Low molecular weight

MALDI-TOF MS – Matrix-assisted laser desorption/ionization time of flight mass spectrometry

MOPS – 3-Morpholinopropane-1-sulfonic acid

MWCO –Molecular weight cut-off

MW – Molecular weight

NPs – Nanoparticles

NRMSD – Normalized root-mean-square deviation

PCR – Polymerase chain reaction

PDI – Polydispersity index

pI – Isoelectric point

Rg – Radius of gyration

R_H – Hydrodynamic radius

ROS – Reactive oxygen species

SAXS – Small angle X-ray scattering

SD – Standard deviation

SDS-PAGE – Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SEC – Size exclusion chromatography

SEC-MALS – Size exclusion chromatography – multi-angle light scattering

SERS – Surface-enhanced Raman spectroscopy

SRCD – Synchrotron radiation circular dichroism

ssDNA – Single stranded DNA

TEM – Transmission electron microscopy

T_m – Melting temperature

UV-Vis – Ultraviolet-visible spectroscopy

VLP – Virus-like particle chaperones

WT – Wild type

XRD – X-ray powder diffraction

1

Introduction

1.1 Ferritin family of proteins

In prokaryotes, the ferritin family is composed of several types of proteins such as ferritins, constituted by 24 subunits (24-mer) (Ftn), heme-containing bacterioferritins (Bfr) also composed of 24-mer and 12-mer mini-ferritins or DNA-binding proteins from starved cells (Dps) [1]. This family is extensive (***Figure 1-1***), and has been shown to perform several physiological functions, such as metal homeostasis and DNA protection [2].

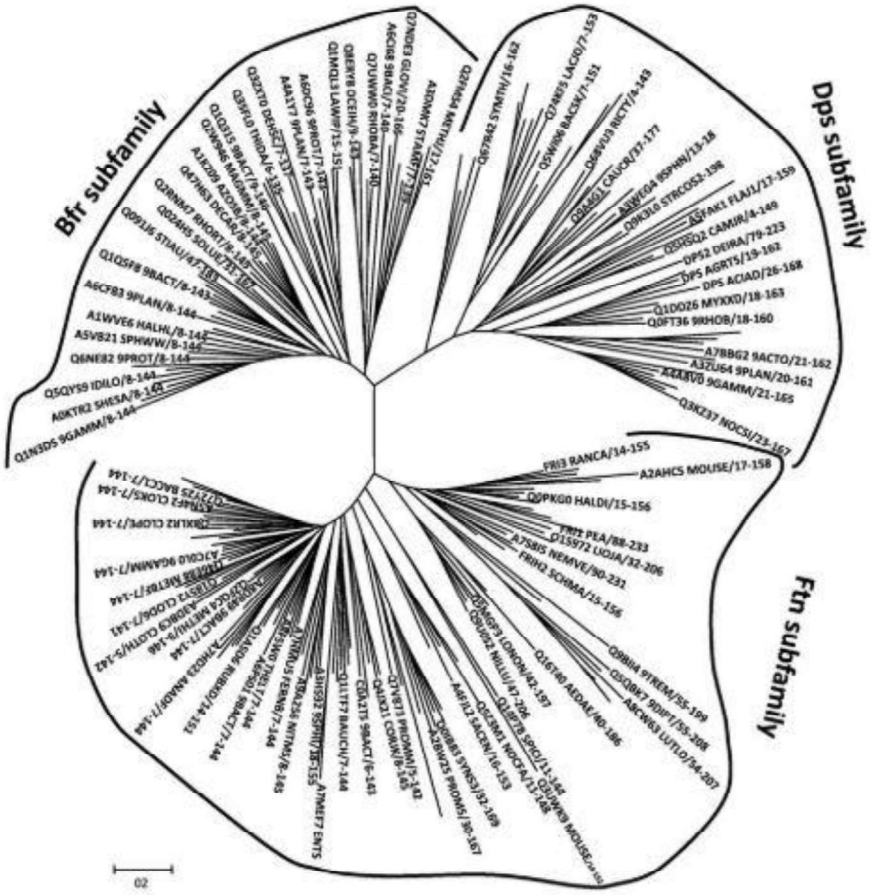

Figure 1-1: Phylogenetic tree of ferritin family of proteins, from [3]**.**

Proteins from this family share highly similar subunits with a four-helical bundle domain that act as building blocks for the assembly of a large hollow protein, almost spherical in shape [4]. This four-helical bundle is constituted by two homologous pairs of anti-parallel helices (***Figure 1-2***). When assembled, the hollow cage is permeated by channels at its axis points, and also through individual subunits, wherein iron and other molecules can enter or exit the protein. In the hollow centre an iron core of ferri-oxy-hydroxide mineral can be formed. This core can

exhibit autocatalytic properties, both in aerobic and anaerobic conditions [5], it however does not influence the electrophoretic or chromatographic properties of ferritins [3].

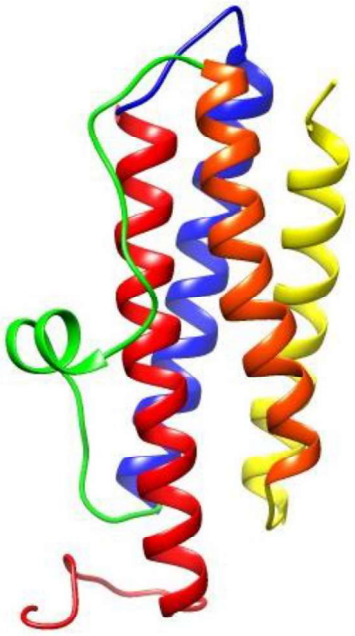

Figure 1-2: Structure of the four helical bundle of Dps monomers showing the two homologous pairs of anti-parallel helices (PDB code:1DPS), generated with USCF Chimera [6]

1.2 Reactive oxygen species

Partial reduced oxygen species are a threat to DNA structural integrity since they can act as reactive oxygen species (ROS). The reduction of molecular oxygen to water can produce three intermediates, superoxide ($O_2^{\bullet -}$), hydrogen peroxide (H_2O_2) and the hydroxyl radical ($OH^{\bullet}$). The hydroxyl radical is highly reactive and was postulated by Fenton [7] and by Haber and Weiss [8] to explain the iron dependent decomposition of H_2O_2 at acidic pH. The Fenton and Haber-Weiss reactions are described in equations 1.1 and 1.2.

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + OH^{\bullet} \qquad (1.1)$$

$$O_2^{\bullet -} + H_2O_2 \rightarrow O_2 + OH^- + OH^{\bullet} \qquad (1.2)$$

Free ferrous iron is also able to generate ROS in the presence of molecular oxygen, as described in equation 1.3. In turn, the superoxide radical ($O_2^{\bullet -}$) is can be converted into hydrogen peroxide (equation 1.4).

$$Fe^{2+} + O_2 \rightarrow Fe^{3+} + 2O_2^{\bullet -} \qquad (1.3)$$

$$2O_2^{\bullet -} + 2H^+ \rightarrow H_2O_2 + O_2 \qquad (1.4)$$

The formation of ROS is extremely harmful for cell machinery, including DNA and proteins [9][10]. In order to mitigate this effect, cells have found solutions, one of them the production of Dps (DNA-binding proteins from starved cells) which act in several fronts, physically by condensing DNA, and chemically by oxidising ferrous iron, in most cases using hydrogen peroxide.

1.3 Dps proteins

Among the different classes of proteins from the ferritin family, a subgroup stands out. DNA-binding proteins from starved cell (Dps), also called mini-ferritins, were identified in 1992 by Almirón *et al.* in cultures of *Escherichia (E.) coli* grown for 3 days [11]. More than a thousand of Dps-like proteins have been identified, around 97 % found in bacteria, and the remaining 3 % in archaea [12]. Dps, contrary to both Ftn and Bfr, only comprises 12 subunits instead of 24. This gives form to a smaller molecule with around 9 nm of outer diameter and 5 nm of inner cavity, instead of the 12 nm and 8 nm of the maxi-ferritins Ftn and Bfr. In turn, the iron-storage capacity is also smaller, with a capacity of 500, compared with 4500 iron atoms for the larger counterpart [12]. The Dps molecule has a different fold symmetry from the maxi-ferritins, wherein the six 4-fold symmetry of the later is replaced by 3-fold axes. Structurally, the subunit of Dps does not have the E helix and instead contains a short helical region in the B-C loop [13]. Dps proteins also present unique ferroxidase sites located at the interface between two fold related subunits (***Figure 1-3***) while maxi-ferritins have the catalytic site in the middle of the 4-helical subunits [14].

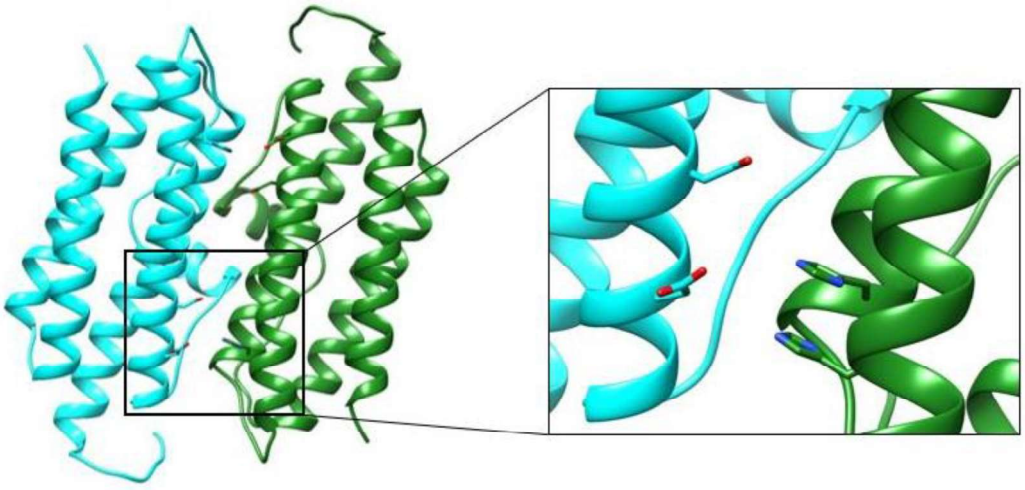

Figure 1-3: Dps ferroxidase sites located at the interface of a dimer, (PDB code:1DPS), generated with USCF Chimera [6]

Another striking difference is that Dps prefers to use hydrogen peroxide (instead of oxygen) to oxidise ferrous ions [15]. This fact is thought to be related with the essential different physiological function of mini-ferritins and maxi-ferritins, Dps is not only an iron storage protein, but also a cell protector against Fenton chemistry, particularly regarding DNA protection. Some Dps proteins also exhibit the ability to bind to the DNA molecule, in a non-specific manner, through the N- or C-terminal extensions [16][17]. In the ferroxidation reaction (fast iron oxidation at the ferroxidase sites), Dps proteins oxidise ferrous ions using hydrogen peroxide, releasing harmless water. DNA being negatively charged attracts Fe^{2+} ions, which can participate in Fenton reactions. Since Dps is stress-induced and associated with DNA, it performs both the removal of H_2O_2 and the sequestration of damaging Fe^{2+} ions [3].

1.3.1 Entry channels

Dps molecules present eight trimeric interfaces (***Figure 1-4***), which exist at the end of either N- or C-terminus of monomers, and create channels connecting the exterior to the interior of the protein. The channels at the N-terminal sites are hydrophilic and have a length of around 1 nm, with an outer opening of 9–17 Å and inner opening of 7–11 Å. These are similar to the typical tri-fold channels of ferritins and have been proposed as the principal route of iron entry, due to their negative charge, mediating the transport of Fe^{2+} cations to the ferroxidase site via an electrostatic gradient. The C-terminal channels are, in contrast with the N-terminal, hydrophobic and generally blocked by hydrophobic residues (called Dps-like tri-fold channels). Therefore, they are usually not considered the main point of entry of the cations to the inner cavity of the protein [14][18][19].

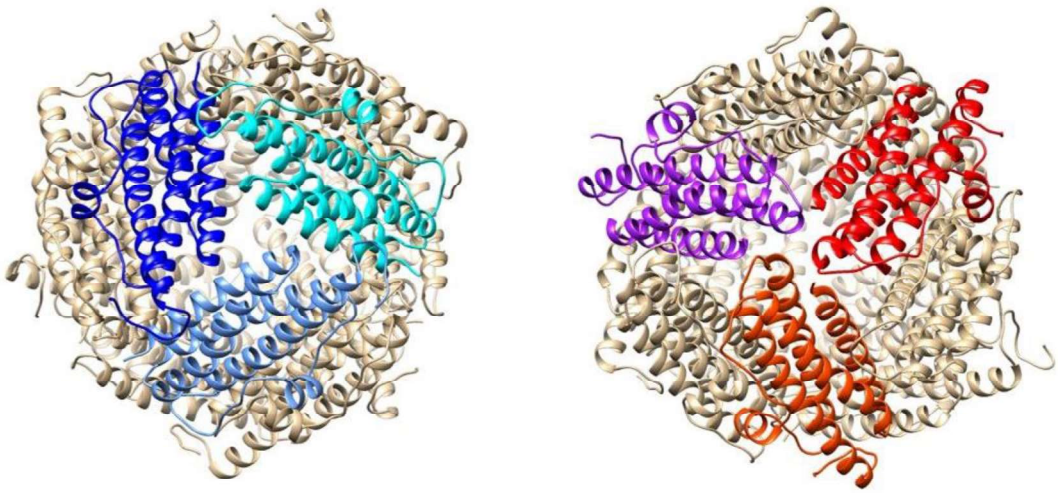

Figure 1-4: Dps-like tri-fold on the right and ferritin-like tri-fold on the left (PDB code:1DPS), generated with USCF Chimera [6]

1.3.2 Oligomerization of Dps

As stated before, Dps is a homododecameric protein. It has been shown that several conformations of Dps exist in solution and are dependent on the experimental conditions, namely pH, ionic strength, and temperature. It comes as no surprise that the different possible oligomeric states of Dps protein in solution affect its functionality, such as the iron mineralisation and DNA-binding properties. It is however not completely clear if the lack of ion mineralisation and DNA-binding properties arise from the different oligomers in solution or due to the experimental conditions tested.

Several amino acids seem to be involved in the folding of Dps protein. Brendan P. Orner *et al.* [20] have studied the effect of the interfacial residues known to be involved in the formation of salt bridges and also several aromatic residues on the folding of the *E.coli* Dps. It was shown that all protein variants studied presented a lower melting temperature, which when combined with TEM and size exclusion chromatography data, did not present a dodecameric structure, but instead either dimers or monomers.

Other works have indicated that the N-terminal tails are also essential for the dodecameric assembly of Dps. M. Vijayan *et al.* [17] demonstrated that when the 9 residues at the N-terminal tail of Dps from *Mycobacterium (M.) smegmatis* were deleted, only trimers were present in solution. In the case of the *Deinococcus (D.) radiodurans* Dps, deletion of the N-terminal tail resulted in the disruption of the dodecameric structure of the protein [21], however this specific Dps variant retained the DNA binding activity when only dimers were present in solution. The ionic strength of the solution was shown to have a role in the assembly of the protein. At low salt concentrations dimers were present and a dodecameric conformation was detected at higher salt concentrations [22]. This effect was also observed by Olga Ozoline *et al.* [23], where *E. coli* Dps monomers could be assembled into a dodecameric structure at high concentrations of salt. Chiancone *et al.* [24] have shown that Dps from *M. smegmatis* exhibits an oligomeric state dependence with pH, in which dimers were present at pH = 5.0 and pH = 8.5, shifting to the dodecameric form at pH = 7.0.

1.3.3 Iron oxidation and storage by Dps

The mechanism of iron uptake has been proposed for the Dps from *Listeria (L.) innocua* [25] and also for the *E. coli* protein [26]. Five steps have been identified: (1) iron entry to the dodecamer; (2) binding of iron to the ferroxidase centre; (3) oxidation of iron at the ferroxidase centre (ferroxidation reaction); (4) nucleation; and (5) formation of the mineral (mineralisation reaction).

The ferroxidase reaction in Dps, with H_2O_2 as the oxidant, is believed to occur as follows:

$$2Fe^{2+} + P^Z \rightarrow [Fe_2 - P]^{z+2} + 2H^+ \quad (1.5)$$

$$[Fe_2 - P]^{Z+2} + H_2O_2 + H_2O \rightarrow [Fe_2O_2(OH) - P]^Z + 3H^+ \quad (1.6)$$

$$2Fe^{2+} + H_2O_2 + 2H_2O \rightarrow 2FeOOH + 4H^+ \quad (1.7)$$

Wherein P^Z is the apo-form of the protein and $[Fe_2 - P]^{Z+2}$ the dinuclear iron bound at the ferroxidase centre. In equation 1.6, $[Fe_2O_2(OH) - P]^Z$ represents an oxidised iron complex at the ferroxidase site, that is subsequently mineralised as hydrous ferric oxide FeOOH (equation 1.7).

It is important to note that Dps proteins can utilise O_2 to oxidise iron, in a less kinetically efficient manner [25]:

$$2Fe^{2+} + P^Z \rightarrow [Fe_2 - P]^{z+2} + 2H^+ \quad (1.8)$$

$$[Fe_2 - P]^{Z+2} + {}^1/_2\, O_2 + H_2O \rightarrow [Fe_2(OH)_2 - P]^Z + 2H^+ \quad (1.9)$$

$$4Fe^{2+} + O_2 + 8H_2O \rightarrow 8FeOOH + 8H^+ \quad (1.10)$$

1.3.4 Ferroxidase sites

There are twelve ferroxidase sites in the Dps dodecamer, two between each dimer (***Figure 1-3***). These binding sites are highly conserved and composed of two histidine residues from one subunit and two carboxylate ligands from the second. At the moment it appears that only the *Thermosynechococcus (T.) elongatus* Dps is distinct from the remaining Dps proteins, wherein three histidine residues were found [16]. Several crystallographic structures have shown iron bound to these sites. Even though these iron coordinating residues are highly conserved, the occupancy of the same differs in the different crystal structures. In *L. innocua* a single iron atom was detected and the presence of a water molecule lying at 3 Å from the first could suggest a second iron binding site, giving rise to the bimetallic ferroxidase centre [14]. Same was observed for Dps proteins from other organisms such as *Helicobacter (H.) pylori* Dps [27], *D. radiodurans* Dps1 [28] and *Agrobacterium (A.) tumefaciens* Dps [29]. Contrarily, the structure of the *Bacillus (B.) brevis* Dps showed two iron atoms at the iron binding sites [18]. In the case of the *E.coli* Dps [13] or *T. elongatus* Dps [30] structures, only water molecules were found. This suggests that the amino acids which form the second metal coordination sphere are highly influential on the affinity of the ferroxidase centre for iron [16]. Once oxidized at the ferroxidase sites iron is stored as a polymorphic Fe_2O_3 mineral in the interior of the cage.

1.3.5 Iron core and nucleation

Two different nucleation sites have been identified, a type I located at the two-fold axis of the protein and a type II in the three-fold axis, also called Dps-like and ferritin-like nucleation sites (***Figure 1-4***). Three iron atoms at the type I nucleation site may form an initial seed allowing the growth of the mineral core. Two symmetry-related glutamate residues of two adjacent subunits are proposed as the structural basis for iron binding at the type II nucleation site; a similar configuration occurs with three glutamates [31].

The ferrihydrite mineral core is formed in the inner cavity of the protein. Two different types of cores have been described: a native core which seems to be present in some purified protein and an *in vitro* core that is formed upon iron addition to the medium in the presence of an oxidant. With respect to the first one, it is usually present in small sizes, in the order of tens iron atoms, with different iron:phosphate ratios. There also exist observations of other metals inside the protein cavity such as Ni, Co, Mn, Cu, Zn [32]. The second type of core, synthesised *in vitro*, has been widely studied [33][34][35]. It has been shown to have a more irregular structure than the native one and the iron atoms in the ferrihydrite are coordinated in an octahedral and tetrahedral geometry. The non-crystalline state of the ferrihydrite in the cavity of Dps makes the iron more readily available, relating with the physiological function of the protein. It has also been reported that the amount of phosphate affects the crystallinity of the cores. Moreover, there is evidence suggesting that oxidation of iron in the presence of Dps can occur under microaerophilic/anaerobic conditions [33].

1.3.6 Iron release

The release of iron atoms from the ferrihydrite mineral core present in the interior of the Dps, is not completely understood and much less is known than for the formation of the same. *In vitro*, the kinetics of the release is accessed by the release of iron from the core into the solvent, promoted by a reducing agent such has sodium dithionite and in the presence of an iron chelator. By measuring the formation of the Fe^{2+}-chelator complex in the medium, it is possible to characterize the kinetics of the reaction. It is important to note that this reduction release process is complex and involves several steps, such as the diffusion of the reducing agent into the cage shell up to the inner cavity of the protein, the electron transfer from the reducing agent to the Fe^{3+} solid mineral and the protonation of the hydroxide ions [36]. It has been shown in the *E. coli* Dps that this release is biphasic, with kinetic constants of 4.7×10^{-2} s^{-1} and 12×10^{-4} s^{-1} [36] that compare with 0.85×10^{-2} s^{-1} and 7.4×10^{-4} s^{-1} reported for the *A. tumefaciens* at pH 7.8 [29].

1.3.7 DNA-binding activity

The DNA-binding activity of Dps has been a motive of a lot of debate in the scientific community. Based on the fact that both DNA and Dps surfaces are negatively charged, it has been proposed that the positively charged residues at the N- or C-terminal tails are involved in the non-specific binding [13][37][38][39]. The non-involvement of these tails has also been defended [29]. Nevertheless, in most cases the N-terminal tail appears to be essential to the DNA binding. Among Dps homologues, the N-terminal tails vary in length and amino acid sequence, as well as in their structural arrangement, some protruding from the cage with high degrees of mobility, affecting the DNA binding activity.

In the Dps protein from *E. coli* it has been shown that the interaction between protein and DNA occurs due to three lysine residues in the mobile N-terminal region, which is extended from the cage into the medium and therefore able to interact with the phosphate groups of DNA [40]. It has also been demonstrated that protein variants lacking 2 out of 3 N-terminal lysine residues were able to bind DNA, however did not condensate it, and removal of all lysine residues abolished the DNA binding ability of the protein [16].

The requirement of a flexible N-terminal tail and the presence of lysine residues for DNA binding was proven for the *A. tumefaciens* Dps. As the *E. coli* protein, this Dps has 3 lysine residues, however contrary to the former does not bind or condense DNA. This is due to the fact that the N-terminal regions are immobilized on the dodecamer surface through hydrogen bonding and salt bridge interactions, and therefore unable to interact with DNA [29].

It has also been suggested that the interaction of the protein with DNA is mediated by divalent cations that form various bridges between Dps and DNA [40], although it is not clearly understood how the mechanism occurs.

Several techniques have been employed to characterise the DNA binding mechanism. Both electrophoretic mobility shift assay (EMSA) and atomic force microscopy (AFM) have been successful in showing the DNA binding ability and plasticity of Dps proteins. It has been observed that at pH greater than 7.5 Dps does not bind to DNA, however when pH is lowered to values closer to 6.5 several binding modes have been observed namely. At pH 7.5, the protein binds to DNA as beads on a string; looping and condensation of DNA as a result of Dps self-association, is seen at lower pH values (***Figure 1-5***) [41][42].

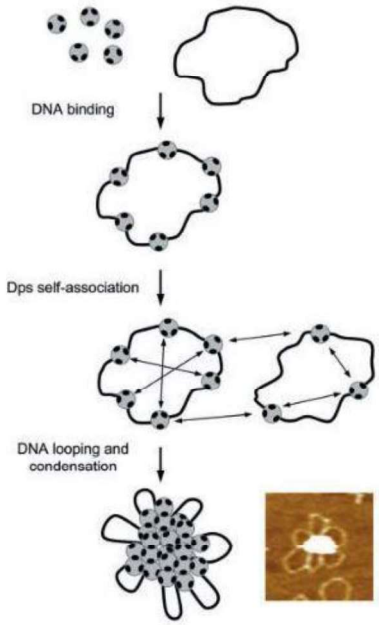

Figure 1-5: Scheme representation of the Dps – DNA binding modes with AFM micrograph of DNA looping and condensation shown, adapted from [42]

1.4. Biotechnology applications of proteins from the ferritin family

The ferritin family of proteins has been extensively studied for biotechnological applications. These are most of the time related with nanoparticle encapsulation, drug delivery or synthesis inside the hollow cavity of these proteins. A vast body of work has been conducted on the ferritin protein, taking advantage of the negatively charged channels and therefore their selectivity, which can harvest different divalent ions [43]. The common approach is to use the protein cage as a template/scaffold for mineralisation/synthesis. This can be achieved by self-assembly the protein in the presence of metal ions, chemically mediated redox reactions, or photochemistry. One of the main advantages when using the proteins from the ferritin family for nanoparticle synthesis or encapsulation of active agents is the fact that the end material obtained is homogeneous in size, which has been a major shortcoming in physico-chemical methods usually employed. Another advantage is the fact that the outer shell of the protein is biocompatible and can be subjected to different modifications to enhance selectivity for different biological compounds, either for sensors or *in vivo* drug delivery applications [44][45][46]. Ferritin has also been shown to be able to solubilize otherwise insoluble compounds, functioning has a vessel and also a detoxifying agent [47]; additionally it also has been reported that the thermal and photostability of molecules inside ferritin are enhanced [48][49]. Given the similarity between ferritins and Dps, similar synthesis approaches have been carried out to synthesize smaller nanoparticles inside the protein. This fact has expanded the range of applications of the synthesis

up to the quantum regime, which is attained when Dps smaller cage is used as the template for the synthesis [50].

There are several ways to take advantage of the ferritin family of proteins for uses in biotechnological applications. They could be broadly categorized utilizing its physiological function of mineralisation, as a nanoreactor, or in a disassembly/assembly process in order to encapsulate the molecule of interest in the interior of the protein [46].

1.4.1 Opening and closing the cage

One of the most well-established methods to open and close the ferritin protein cage takes advantage of pH changes to promote the disassembly and assembly of the protein, which is based on the reversibility of the process. Disassembling occurs at extreme pH values, either very acidic (pH 2-3) or very basic (pH 11-12) and it is then reversed to more neutral conditions, resulting in the assembly into the 24-mer structure [51]. This dynamic process has been used to easily convert the protein in a delivery vehicle. Among the areas that most benefit from this protein ability is Medicine, more specifically in cancer chemotherapy, taking advantage of the size, the hollow cage structure and its biocompatibility [52]. Ferritin nanocages have been used to encapsulate chemotherapeutic agents to fight cancer [47][48]. This type of approach has the advantage regarding the selectivity and also uptake of the cargo by the chosen cells, due to the possibility to conjugate the surface of the protein with molecules of interest [55]. Besides, ferritin has also been explored as a natural delivery agent for vaccines by fusing influenza virus haemagglutinin at the surface of the protein prior to mice immunization, which produces a better immunity response when compared with the usual immunization [56].

A different type of application that has also great potential for several industries, is to use the proteins from the ferritin family to solubilize compounds that in the free form are insoluble in water. Moreover, ferritins may also improve the thermal and photostability of compounds by providing a protective coating to the encapsulated compound. This has been exemplified for several molecules as curcumin [48] and β-carotene [56].

So far Dps proteins has not been used as extensively as ferritins to perform encapsulation of drugs or compounds with biological relevance, however *S.* Cacchi *et al.* [57] have been able to stabilize palladium inside the protein cage. One can expect that similar principles of encapsulation used in ferritin can be also applied to mini-ferritins.

1.4.2 Using proteins from the ferritin family as nanoreactors

Nanoparticles are fundamental in nanotechnology and are used in different applications, such as catalysts, fluorescence probes, drug delivery system, among many others. However, in order to be able to take advantage of nanoparticles, these must be homogenous in size,

morphology and content. Given this requirement bio-templated synthesis of nanoparticles has been performed using proteins from the ferritin family due to the size of their hollow cage. Since, and as aforementioned, they are able to oxidise and store iron making these proteins obvious candidates to mediate the synthesis of metal oxides [58].

Their usage as reaction vessels to produce nanoparticles or encapsulation of drugs with relevance to, for example, theragnostic applications has been documented. The self-assembly properties in combination with the inherent physiological functions of these proteins give rise to materials with enhanced properties when compared with the precursors [59].

Several types of proteins have been utilized with this in mind, such as viral capsids [60], however the proteins from the ferritin family has been like-wise used, taking advantage of the different synthesis or encapsulations routes that can be performed with this type of proteins (***Figure 1-6***) [61][46].

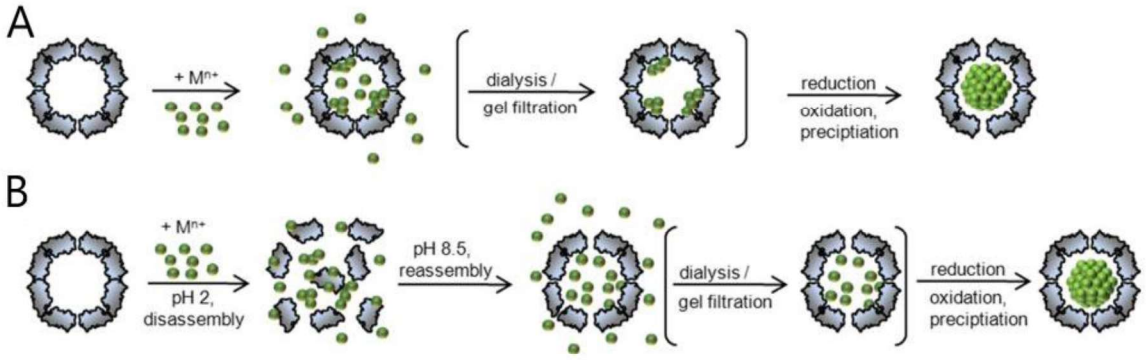

Figure 1-6: Possible routes for incorporation of synthetic cores inside ferritin nanocage. (A) Dissolved ions in the external medium are translocated through the ferritin shell via the entry channels to the inner cavity of the cage; and (B) Incorporation of ions in the solution are encapsulated by assembly and disassembly of the nanocage through pH changes, adapted from [46].

One of the firsts works that took advantage of ferritin to synthesise non-biological materials was reported by S. Mann *et al.* [62] that used apoferritin to synthesise manganese and iron oxides. The authors have demonstrated that biomineralisation of the material occurs inside the hollow cage and not at the surface, which is an indication of the viability of this methodology for the synthesis of different nanoparticles. Since then, several other nanoparticles have been synthesized, such as cobalt hydroxide [63], nickel and chromium hydroxide [64], cadmium sulphide [65], cadmium and zinc selenide [66], [67], lead sulphide [68], among many others. One great improvement of this approach is the fact that the synthesis can, most of the time, be made in a simple, affordable, one-pot reaction at ambient conditions. Diverse nanomaterials can thus be produced, with applications in a vast array of industries, such as semiconductors [69] and photovoltaics [70], which also validates the versatility of this method.

The potential of Dps to be employed for nanoparticle synthesis, although to a lesser extent, has also been explored. Yamashita *et al.* [71][50] were able to perform the synthesis of both cadmium sulphide and cadmium selenide nanoparticles in the interior of Dps from *L. innocua.* Owing to their smaller size, the CdS nanoparticles obtained were photoluminescent. This work showed that mini-ferritins can be used to produce smaller nanoparticles than those using ferritin, which can be applied, for instance, in the semiconductor industry.

1.4.3 Dps as a building block for the synthesis of supra-molecular complexes

The miniaturization of materials with the goal to reach the nanoscale has been a trend in the top-down scientific approach, mainly in the semiconductor industry. However, the possibility to generate nanomaterials with a bottom-up approach is a methodology that has been popularized during this century. The prospect of creating a synthetic assembly or self-assembly of biomolecules in ordered supramolecular structures is reaching a tipping point towards new types of biomaterials and their applications [72]. These new structures can be used in a variety of fields including, but not limited to pharmacy, regenerative medicine and diagnostics [59].

Biomolecules tend to be chemically well-defined and monodisperse, which present advantages when used as building blocks, either by themself or in conjunction with, for example, nanoparticles, giving rise to bio-(in)organic nanoparticles [73]. On the other hand, proteins allow site-specific chemical conjugation reactions and are produced under genetic control. This provides the ability to control the building blocks before the synthesis of the envisioned final supramolecular structure.

This protein can then be utilized as the building block for the synthesis of supramolecular polymer conjugates by means of non-covalent interactions or by site-selective protein modifications [74][75]. Click-chemistry as also been shown to be a possible pathway to produce protein polymers conjugates. Briefly, click-chemistry is a concept introduced by Sharpless *et al.* [76] that changed the paradigm of bonding design to a more capable use of reactions that inherently explore reactions that are efficient at binding two components. These reactions should be as stated by the authors, and quoting, “modular, wide in scope, high yielding, stereospecific, simple to perform, creating only harmless by-products (that can be removed without an additional chromatographic step) and requiring benign or easily removed solvents, preferably water.” One of the most used click-chemistry approaches applied to biomolecules, such as peptides, virus and oligonucleotides has been the copper catalysed azyde-alkyne click reaction [77].

So far, this elegant methodology has been successful and seems to be the most straightforward approach to synthesise biomolecular supramolecular complexes [77][78].

1.5. DNA origami

The DNA origami method has been introduced by Rothemund in 2006 [79], allowing the rapid synthesis of well-defined complex DNA nanostructures, based on the DNA unique self-assembly capability, which arise from the strong specificity of the Watson-Crick base pairing (***Figure 1-7***). In this technique a long single-stranded (ss) DNA scaffold (usually viral) is folded into a nanoscale shape by hybridization with several short synthetic oligonucleotides, the staple strands. Each of these strands is complementary to different separated segments of the scaffold strand which causes the folding of the structure during hybridization. This new DNA origami structure consists of completely double stranded (ds) stabilised by periodic crossovers of staple strands. One can therefore choose the shape of the final origami structure by choosing the sequences of individual staple strands [80].

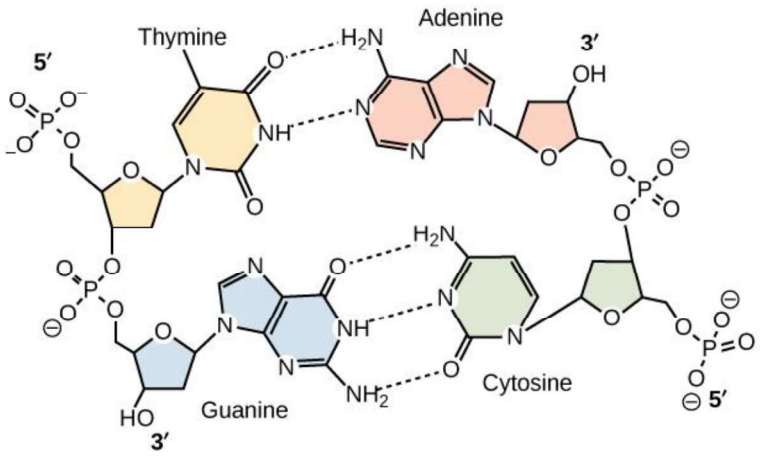

Figure 1-7: Watson-Crick base pairing in the DNA molecule, from [81].

The assembly of DNA origami is usually prepared with a high excess of staple strands in the presence of a buffer containing Mg^{2+}, which hinders the electrostatic repulsion of the negatively charged DNA strands. The solution containing the DNA scaffold and the staple sequences is rapidly heated above the melting temperature of the DNA and then slowly cooled down to room temperature. This slow process allows the staple strands to find their complementary sequence and fold the scaffold into the desired shape (***Figure 1-8***). The DNA scaffold is usually a M13mp18 single stranded (ss) DNA that is then folded with around 200 staple strands [79]..

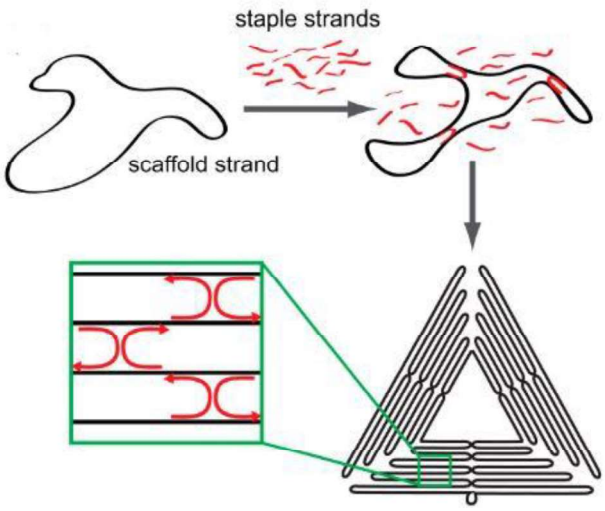

Figure 1-8: Schematic representation of DNA origami folding in triangular shape, adapted from [80].

One of the great advantages of DNA origami technique is the fact that the staple strands can extend from the DNA origami surface generating more than 200 unique sites that can be modified regarding length, nucleobase sequence and hybridization state, which can give rise to different shapes (***Figure 1-9***). This makes the origami templates useful as physical support for a variety of molecules, namely, nanoparticles [82], proteins [83], fluorophores [84] and quantum dots [85].

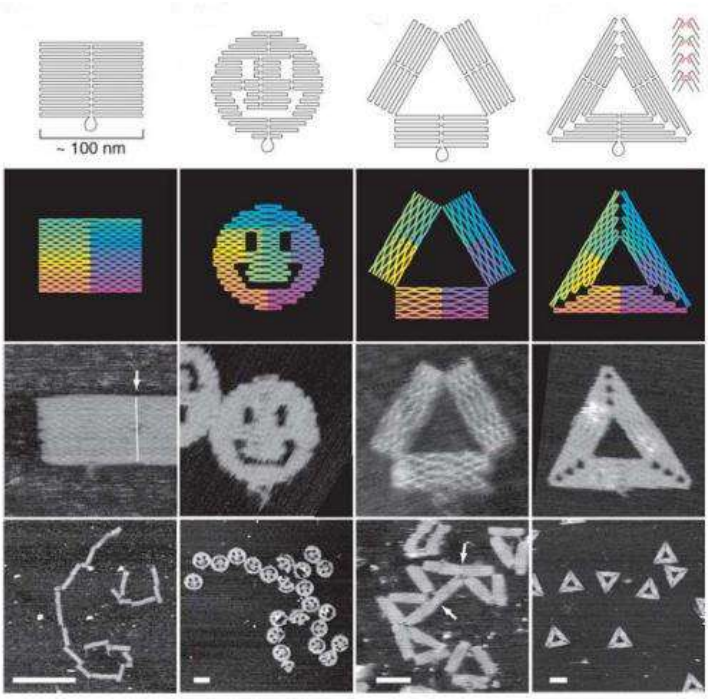

Figure 1-9: Different possibilities for DNA origami folding, all scale bars of 100 nm except bottom left of 1 μm, adapted from [79].

DNA origami nanostructures are usually probed with single molecule microscopy techniques such as atomic force spectroscopy (AFM) and transmission electron spectroscopy (TEM), although cryogenic electron microscopy (Cryo-EM) has also been applied to resolve complex 3D-origami structures [86]. Even though EM can yield very high-resolution images, when one considers the sample preparation requirements needed to perform a EM analysis, AFM is usually preferred. AFM can be operated either under atmospheric conditions or liquids, requiring very little samples preparation and measurement in different buffer conditions [87].

In AFM technique a cantilever with a sharp tip (radius of apex curvature 1-10 nm) is scanned over the surface under analysis. The interaction forces between the tip and the sample, such as long range and attractive van-der-Waals forces, electrostatic interactions, and Pauli repulsion lead to a deflection of the cantilever (in contact mode) or to a reduced vibrational amplitude (in dynamic mode), resulting in the recording of the sample topography, at lateral resolution of 1-10 nm and a height resolution of 0.1 nm. Under atmospheric conditions AFM is typically used to probe macromolecules such as DNA and proteins. AFM technique is also useful to determine other properties of the sample other than topographic, as mechanical properties and surface charge [80][87].

Other analytical tools are used to analyse functionalized DNA origami structures, namely fluorescence spectroscopy [88] and surface-enhanced Raman spectroscopy (SERS) [88][89].

So far the only work performed with DNA origamis and a protein from the ferritin family was related with DNA origami purification [91] and also using DNA origami to nanopattern proteins on mica surface [92]. DNA origamis have not been used to study the DNA binding mechanism by Dps proteins.

1.6 Overview of this thesis

The ferritin family of proteins are ubiquitous in all kingdoms of life. In this thesis the physiological role of Dps is studied. Two of the most important roles of Dps protein in the cell are DNA protection due to their ability to use H_2O_2 to perform ferroxidation, which results in cell detoxification, and physical protection of DNA thus shielding it from damages.

The main objective of this thesis is to study and characterize Dps protein from a halophilic bacterium, *Marinobacter hydrocarbonoclasticus* 617, to take advantage of its physiological role for future biotechnological applications.

Chapter 2 briefly delves into the expression and purification of Dps and general techniques used during the course of the work. This is followed by the study and characterization of the oligomeric assembly of Dps taking advantage of the comparison with a deletion mutant. The assembly process is still not completely understood. To exploit the biotechnological potential of Dps benefiting of its cage structure, the effect of the experimental conditions (pH, ionic strength, or protein concentration) on the oligomeric state of the protein must be elucidated.

In chapter 3 Dps protein is used as a nanoreactor for the synthesis of nanoparticles. A simple and versatile methodology is developed for nanoparticle growth inside the protein. This chapter demonstrates its versatility to produce different sulphur containing metal nanoparticles. A manuscript is also included, wherein Dps is used as a building block to produce long bioactivity fibres.

The DNA binding activity of Dps is studied in Chapter 4 with great emphasis on AFM and EMSAs (electrophoretic mobility shift assays) analysis. DNA origami technique is used to try to elucidate the DNA binding mechanism of Dps. A manuscript is also presented, reporting the main results and demonstrating the effect of a single charged alteration in the N-terminal tail on DNA binding.

The last chapter, chapter 5, summarizes the main conclusions of the work and discusses the possible future work.

2

Oligomeric forms of Dps

2.1 Overview

The expression and production of Dps protein from *M. hydrocarbonoclasticus* in *E. coli* focused on the present work is addressed in this chapter. The expression plasmid containing the Dps encoding gene was originally produced by M. Guilherme [93]. The protocol for overexpression and purification of the wild type (WT) Dps was first described by M. Guilherme and subsequently optimized by D. Penas [94]. The protein variants used in the course of this thesis were all expressed and purified by D. Penas [94].

WT Dps and the deletion mutant Δ15 (a protein variant in which the first 15 amino acid residues have been removed) were studied in order to understand the role of the N-terminal tail in the fully assembly of Dps as a homododecamer. The effect of the buffer composition and temperature was also investigated. In the case of the protein from *Deinococcus radiodurans*, the presence of different oligomeric forms in solution depending on the ionic strength of the buffer has been reported [22]. The relevance of the N-terminal tail in the assembly of the *M. smegmatis* Dps was also demonstrated [17].

Data presented in this chapter and acquired by the application of different spectroscopic techniques indicate that the N-terminal tail (size and composition) is of upmost importance to the overall stability and assembly of the protein in a 12-mer structure. This also provides insight into the assembly dynamics of the nanocage, paving the way for possible biotechnological applications.

2.2 Experimental procedure

2.2.1 Expression and purification of wild type Dps

Expression of WT Dps was carried out in *E. coli* BL21(DE3) cells harbouring the pET-21c-Dps-WT plasmid. The overexpression was performed in 1 L LB medium containing 100 μg/mL of ampicillin. First, 5 mL of LB containing 100 μg/mL of ampicillin was inoculated with a transformant and allowed to grow for 8 h at 37 ºC, under 220 rpm. 1 mL of this was used to inoculate 100 mL LB medium with 100 μg/mL ampicillin and grown overnight in the same conditions. Finally, 10 mL of the later inoculum were transferred to 1 L of LB medium with 100 μg/mL ampicillin, which was grown overnight at 37 ºC and 220 rpm. When the optical density at 600 nm reached 0.5, gene expression was induced by addition of 0.5 mM isopropyl β-D-1-thiogalactopyranosid (IPTG). Cells were allowed to grow for 3 h at 37 ºC, under 200 rpm and then harvested by centrifugation at 11,000 x g for 10 min at 6 ºC (Z 36 HK, HERLME LaborTechnik). The resulting pellet was resuspended in 10 mM Tris–HCl buffer pH 7.6. To enhance cell disruption and protein extraction, cell suspension was subject to 3 freeze-thaw cycles

and incubated with DNase I (Roche). The lysis was then performed by sonication using an ultrasonic homogenizer with a 7-mm probe (LABSONICE M, Sartorius). The sonication occurred with the cells on ice for 3 min, with 80 % amplitude followed by 90 s of rest, being the cycle repeated 4 times. DNase I was further added to reduce the viscosity of the resulting extract. For cellular fractionation, the suspension was ultracentrifugated at 181,000 x g for 70 min at 4 ºC (Beckman Coulter type 70 Ti rotor). The soluble fraction was dialysed overnight at 4 ºC against 10 mM Tris–HCl buffer pH 7.6. Before purification the soluble protein extracts were centrifugated at 10,000 x g for 10 min at 6 ºC (Z 36 HK, HERLME LaborTechnik) and filtered with a 0.45 μm membrane filter (polyethersulfone, VWR).

WT Dps was purified following a two-step protocol. Firstly, the soluble protein extract was loaded into a DEAE Sepharose Fast Flow column (2.6 x 30 cm, Cytiva, formerly GE Healthcare Life Sciences) equilibrated with 10 mM Tris–HCl buffer 7.6 and washed with the same buffer. The adsorbed proteins were then eluted with a discontinuous linear gradient of 500 mM NaCl in 10 mM Tris–HCl buffer pH 7.6 (elution buffer) using a flow rate of 5 mL/min and collecting fractions of 12 mL. The fractions with WT Dps protein were pooled and dialysed overnight at 4 ºC against 10 mM Tris–HCl buffer pH 7.6.

The second purification step was performed in a Resource Q column (6 mL, Cytiva) equilibrated with 10 mM Tris–HCl buffer pH 7.6. The protein fraction was loaded onto the column at a flow rate of 4 mL/min and washed with the equilibration buffer. The elution of protein was done with a discontinuous gradient as previously described. A flow rate of 1 mL/min was applied and 2 mL fractions collected.

To test the presence of catalase contaminations, the activity of this protein was tested by incubating 100 μL of protein with three drops of 30 % (w/v) hydrogen peroxide (H_2O_2). The presence of catalase is considered if bubbles are formed in the sample upon H_2O_2 addition. Pure WT Dps fractions were dialysed overnight at 4 ºC against 200 mM MOPS, pH 7.0, 200 mM NaCl buffer and stored at −80 ºC.

Protein concentration was determined by UV-Vis spectroscopy, using a molar absorption coefficient of 2.93 x 10^5 $M^{-1}cm^{-1}$ at 280 nm [93]. The protein was concentrated utilizing Vivacell 70 and Vivaspin 20 concentrators, both with 100 kDa MWCO (Sartorius).

2.2.2 Spectroscopic methods and techniques

Circular Dichroism (CD) is a spectroscopic technique, which has been used for more than 40 years. It is especially useful in the area of structural biology, where protein structures are studied [95][96]. CD spectroscopy is based in the differential absorbance of left and right circularly polarized light by chiral molecules. For proteins and their complexes, the UV

(ultraviolet) wavelength is important. In typical laboratory settings where a xenon arc lamp is employed the photon flux is not enough to acquire reliable spectra in the UV region. To overcome this setback synchrotron radiation circular dichroism (SRCD) can also be used. SRCD presents several advantages when compared with the conventional CD spectroscopy, these advantages derive from the increased photon flux. The higher number of photons allow an increase to the signal-to-noise data in the acquired data. Given this, in SRCD, one can use smaller amounts of material, perform faster and more accurate measurements down to wavelengths as low as 168 nm [97] (for more information see Appendix A).

CD measurements were obtained at the BioLab-UCIBIO, Biological and Chemical Analysis Facility, Universidade Nova de Lisboa, using an Applied Photophysics™ qCD Chirascan spectrometer. Spectra were recorded with 1 nm steps and a dwell time of 2 s per step, in triplicates, using 0.01008 cm quartz cells (SUPRASIL, Hellma GmbH, Germany), for the wavelength range of 185–280 nm. Samples were measured at 1 mg/mL in 200 mM MOPS pH 7.0 200 mM NaF.

SRCD measurements were obtained at the SRCD facility at the UV1 beamline on the storage ring ASTRID at the Institute for Storage Ring Facilities (ISA), University of Aarhus, Denmark. The light passed from the synchrotron through a CaF_2 window into a CD instrument purged with nitrogen. In the instrument an initial linear polarization is converted into alternating circularly polarized light with a Rochon polarizer and a photo-elastic modulator (PEM, model I/ CF50, Hinds, USA) operating at 50 kHz. The signal is detected with a photo-multiplier tube (PMT, 9402B, Electron Tubes, UK). Camphor-sulfonic acid served as a calibration material for the instrument. The obtained spectra were recorded with 1 nm steps and a dwell time of 2,725 s per step in triplicates, using 0.01008 cm quartz cells (SUPRASIL, Hellma GmbH, Germany), for the wavelength range of 170–280 nm. Samples were measured at 1 mg/mL in 200 mM MOPS pH 7.0 200 mM NaF and the buffer spectra were used as baseline and subtracted to the sample spectra. The temperature ramps were acquired from 5 ˚C to 90 ˚C, and the following equilibrium times in between step increases were used: 5 minutes from 5 ˚C to 40 ˚C, 2.5 minutes from 40 ˚C to 65 ˚C, 5 min from 65 ˚C to 90 ˚C. Before and after the temperature ramp a spectrum at 25 ˚C was measured.

In order to calculate the contribution of α-helix and β-sheet content in the secondary protein structure, suitable spectra were fitted using Dichroweb online server [98].The fitting was conducted using CDSSTR method and SP175 reference data set [99].

Dynamic light scattering (DLS) was used to estimate the hydrodynamic radius, R_H utilizing a SZ-100 nanopartica, Horiba Scientific. The hydrodynamic radius, and polydispersity index (PDI) were obtained as the mean ± standard deviations (SD) of triplicate measurements.

Small Angle X-ray Scattering (SAXS) measurements were conducted at the European Molecular Biology Laboratory (EMBL), Hamburg at the PETRA III, beamline, P12, using a size-exclusion chromatography (SEC) Superdex 200 5/150 GL column (Cytiva) coupled with the standard experimental setup [100]. Data analysis was performed with ATSAS software following the recommended guidelines [101]. The distance distribution function P(r) was determined using DATGNOM tool. DAMMIN program was used to generate the ab initio models using P1 symmetry generating 15 models, aligned with DAMAVER. The obtained model was refined with DAMMIN. Crysol was used to obtain the envelope of the crystal structure of Dps WT and the hypothetical oligomer and compared with analysed samples. The CHIMERA software was used for 3D graphics representation.

UV-Vis spectroscopy measurements were performed with Evolution 201/300 Thermo Scientific spectrophotometer connected to a computer.

2.2.3 Thermoanalytical methods and techniques

Differential Scanning Calorimetry (DSC) was performed at BioLab-UCIBIO, Biological and Chemical Analysis Facility, with a differential microcalorimeter Nano DSC (TA Instruments) equipped with a personal computer system. DSC data was analysed using NanoAnalyse software, developed by TA Instruments, wherein a two-state scaled model was employed to calculate the melting temperature, T_m, and enthalpy, ΔH. In order to improve the confidence of the results the statistics model of the software was utilized. The model was run 5000 times with a desired confidence of 95 %.

2.2.4 Biochemical characterization

The oligomeric state dependence with different buffer composition and protein concentrations were assessed using a FPLC system equipped with a UV detector (ÄKTAprime plus, GE Healthcare Life Sciences) and a Superdex™ 200 10/300 GL (Cytiva). The column was pre-equilibrated with a buffer containing 50 mM MOPS, pH 7.0 and different molar concentrations of NaCl (to equal the salt concentration of each protein sample to be analysed). 1 mL of each protein solution was applied onto the column at a concentration of 1 mg/mL, with a flow rate of 0.5 mL/min with 50 mM MOPS, pH 7.0 with either 50 mM, 100 mM, 200 mM, or 300 mM NaCl. To assess the protein concentration effect on the oligomerization process, three different concentrations of WT Dps, 1 mg/mL, 5 mg/mL and 10 mg/mL in 50 mM MOPS, pH 7.0, 50 mM NaCl were injected in column the pre-equilibrated with the same buffer at a flow rate of 0.5 mL/min.

2.3 Results and discussion

To study the influence of the N-terminal extension on the melting temperature, WT Dps and the deletion Δ15 variant proteins were characterized. The SRCD spectra of WT Dps and Δ15 mutant (***Figure 2-1***) are typical of α-helix secondary structure rich proteins, which is in concordance with the previous reports on ferritins [102] and Dps homologues [22].

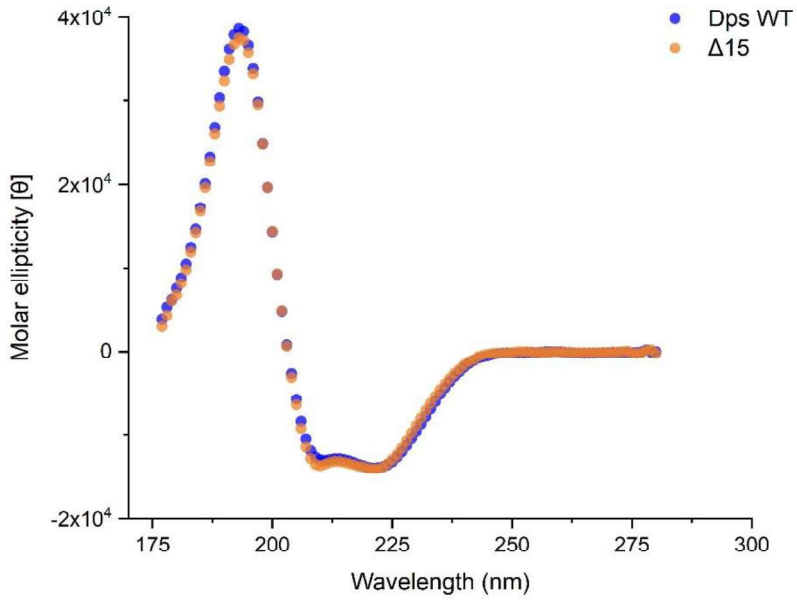

Figure 2-1: SRCD spectra of Dps WT and Δ15 proteins (1 mg/mL) in 200 mM MOPS pH 7.0 200 mM NaF at 25 ºC.

Table 2-1: SRCD spectra deconvolution using Dichroweb CDSSTR SP software (175-240 nm) of Dps WT and Δ15 mutant at 25 °C

	Helix1	Helix2	Strand1	Strand2	Turns	Unordered	Total	NRMSD
Dps-WT	0.45	0.20	0.060	0.030	0.090	0.17	1.0	0.0070
Δ15 mutant	0.61	0.23	0.040	0.010	0.060	0.050	1.0	0.0060

After Dichroweb data fitting, ***Table 2-1***, it is clear that the secondary structure of Dps WT is similar to the deletion mutant Δ15 ***Table 2-1***: , but in the former there is, however, a higher content of unordered structures. This is due to the fact that the first 15 residues present in Dps WT monomers are inherently unordered. Absent in the deletion mutant, this will be mainly reflected as an increase of the α-helix contribution.

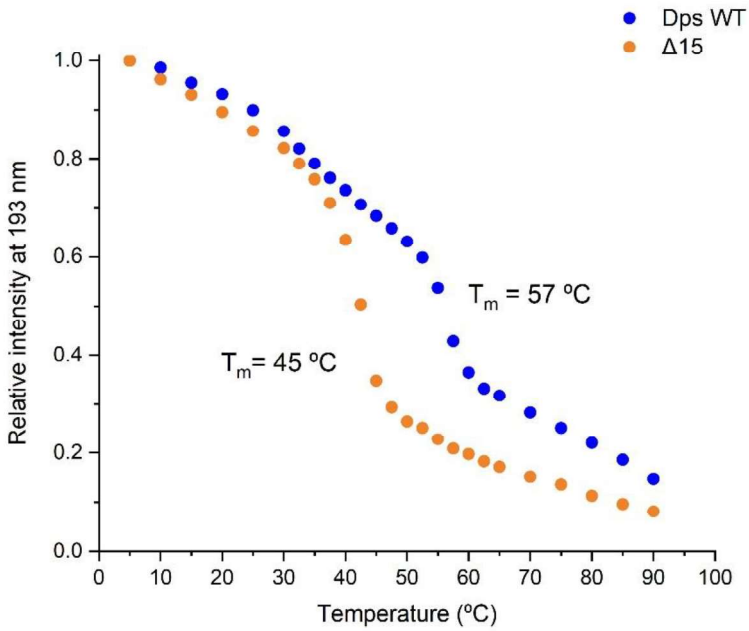

Figure 2-2: Dps WT and Δ15 mutant CD melting curves with melting temperatures.

Even though the initial secondary structure of Dps WT and Δ15 mutant are similar with exception to the unordered part, the thermostability of both proteins is different as ***Figure 2-2*** illustrates. After normalization of SRCD spectra and by monitoring the relative intensity at 193 nm, it is evident that the Δ15 mutant presents a lower melting temperature (T_m = 45 ºC) when compared with than Dps WT protein (T_m = 57 ºC). This difference of more than 10 ºC is indicative of a less stable oligomeric structure in the Δ15 mutant. It is important to bear in mind that both CD and SRCD only take into consideration the secondary structure of monomers, and since this particular system is a homododecamer, it is reasonable to conclude that the N-terminal tail is involved in the stability of the protein.

Figure 2-3 shows the profile of secondary structure during temperature ramps. Here it is evident that with the increasing temperatures, the α-helices have been converted into unordered structures. This is observed for both proteins and, once again, the transition occurs at lower temperatures values in the case of the Δ15 mutant relative to Dps WT. It is also interesting to note that in the mutant the conversion of α-helices occurs in a narrower temperature range, suggesting a more cooperative monophasic process. Also important is the fact that even though it starts with a less unordered structure when compared with Dps WT, at the end of the temperature ramp, 23% is in an unordered structure form, while Dps WT is 10% less in comparison at the end point, suggesting a stronger temperature stress resistance of the WT Dps protein.

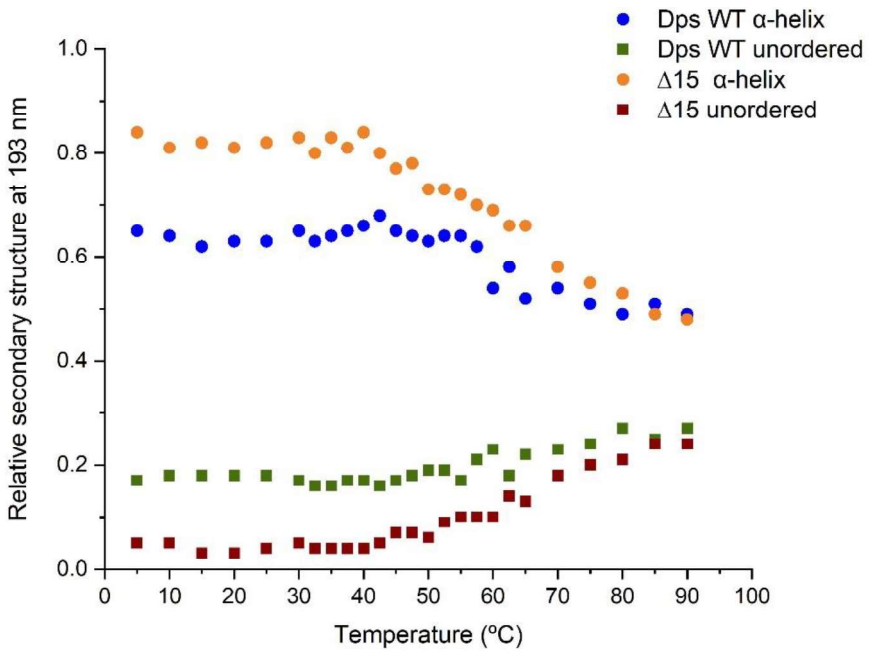

Figure 2-3: Dichroweb CDSSTR SP (175-240 nm) deconvolution of Dps WT and Δ15 mutant SRCD spectra with temperature ramps.

When samples were studied by DSC, ***Figure 2-4***, the Δ15 mutant also exhibited a lower T_m, of 46 ˚C, with a sharp peak and an enthalpy (area under each peak), ΔH_m around 1011 kJ/mol, suggesting a two-states unfolding process. However, the behaviour of Dps WT was distinct. A higher T_m value of 60 ˚C, in agreement with the SRCD data, and a ΔH_m of 949.6 kJ/mol was determined, within the same range of Δ15 protein. Besides, Dps WT shows other enthalpic processes during unfolding, one before and another after the T_m transition. The first at 48 ˚C with ΔH_m of 180.7 kJ/mol and the second at 74. ˚C with a ΔH_m equal to 367.5 kJ/mol. It is important to stress that DSC allows the study of heat changes of the whole 12-mer structure in solution and not just the mononomers. Therefore, we reasoned that the first heat exchange event at around 48 ºC occurs in both Dps WT and the Δ15 mutant. The absence of other processes in the Δ15 mutant protein curve led us to concluded that, either the mutant is not in a 12-mer state or Dps WT is present in different oligomeric states in solution.

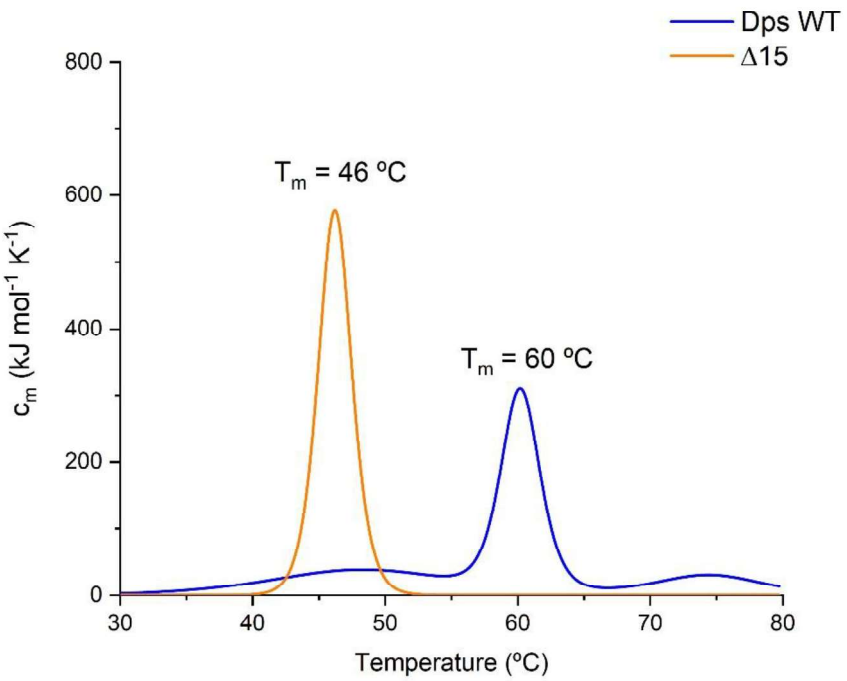

Figure 2-4: DSC thermogram of Dps WT and Δ15 mutant in 50 mM MOPS pH 7.0 50 mM NaF, data fitted with TwoStateScaled model.

When samples were assessed by DLS (***Table 2-2***) the most striking result was that while Dps WT displays, at room temperature, a hydrodynamic diameter of 15.0 ± 1.1 nm, the Δ15 mutant shows a 8.00 ± 1.20 nm, approximately half the size reported in the literature for Dps proteins [103]. This result is a clear indication that in solution both WT and Δ15 proteins assembly differently. One should mention that since the solvation sphere in DLS is also considered, the size is overestimated (in ~ 3 Å). When the temperature was increased, the behaviour of the proteins was not the same, as presented in ***Figure 2-5***. It was interesting to note that Dps WT starts to unfold and to form larger aggregates (increase of the hydrodynamic diameter) at a higher temperature than the Δ15 mutant, at 50 ˚C and 42 ˚C respectively, indicating that the mutant is less thermostable than the WT protein. DLS data are in concordance with both CD and DSC results.

Table 2-2: Diameter and PDI of Dps WT and Δ15 mutant, at room temperature, determined by DLS

	Diameter (nm)	**PDI**
Dps WT	15.0 ± 1.1	0.370
Δ15 mutant	8.00 ± 1.2	0.340

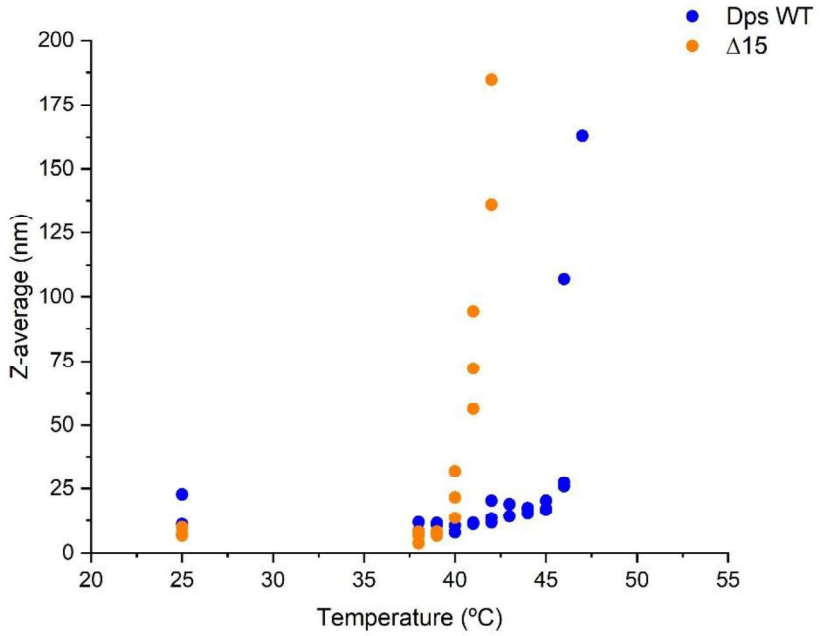

Figure 2-5: DLS analysis of Dps WT and Δ15 mutant proteins in 200 mM Mops pH 7.0 200 mM NaCl Z-average behaviour with temperature increase.

The oligomeric state of Dps WT and Δ15 mutant proteins was investigated by SAXS at different ionic strengths. Both proteins were studied in two buffers, 200 mM MOPS, pH 7.0, containing 200 mM NaCl (here considered as high ionic strength) and 50 mM MOPS, pH 7.0 with 50 mM NaCl (low ionic strength). The striking difference in the P(r) functions between both Dps WT and Δ15 mutant indicates a clear difference in their oligomeric state (***Figure 2-6*A**). This hypothesis was further confirmed by the difference in the radius of gyration (R_g), wherein the Δ15 mutant at a low ionic strength presents the smallest value of 2.41 nm, while Dps WT the highest with 3.86 nm.

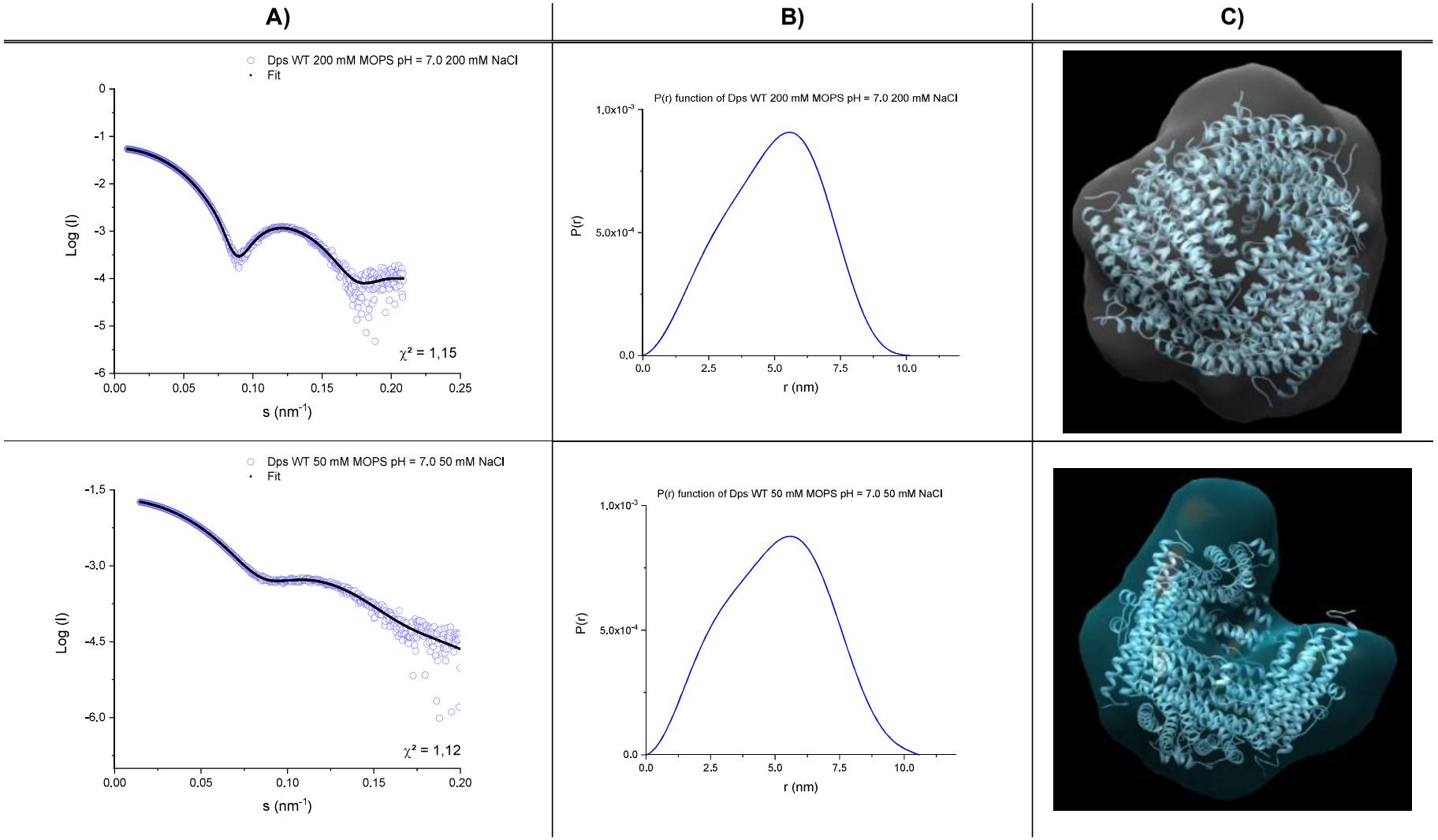

Figure 2-6:SAXS scattering profiles (A), P(r) functions (B), models and envelopes (C) of Dps WT at high (top) and low ionic strengths (bottom).

The molecular mass of Dps WT at high ionic strength, as estimated by SEC-MALS, was equal to 200.0 kDa. This would account for 11.2 monomers when one compares with the theoretical value deduced from the amino acid sequence. However, after generating of the model and superimposing it with Crysol data (***Figure 2-6C***), a full dodecameric structure with a χ^2 of 1.15 is fitted. Applying the same methodology for Dps WT at low ionic strength, the molecular mass of 179.5 kDa that would account for 10.1 monomers, according with the Crysol and the model. This appears to also explain why the R_g of this sample was slightly larger than the one for Dps WT at high ionic strength. Apparently, under these conditions Dps WT disassembles by losing two monomers (a dimer) in a blooming manner, exposing the hollow cage, and resulting in a slightly larger R_g and D_{max} relative to the folded state. The Δ15 variant does not exist in a fully assembled form in solution (***Figure 2-7***). The deletion of the first 15 amino acid residues in the N-terminal region seems to completely hinder the assembly capability of the system in a dodecameric structure. This is observed either at high or low ionic strength conditions. At high ionic strength SEC-MALS data suggest the presence of a hexamer, in agreement with Crysol and the model, with a χ^2 of 1.10. However, at low ionic strength a trimer seems to be the predominant form of the protein with Crysol and the model presenting a χ^2 of 0.95.

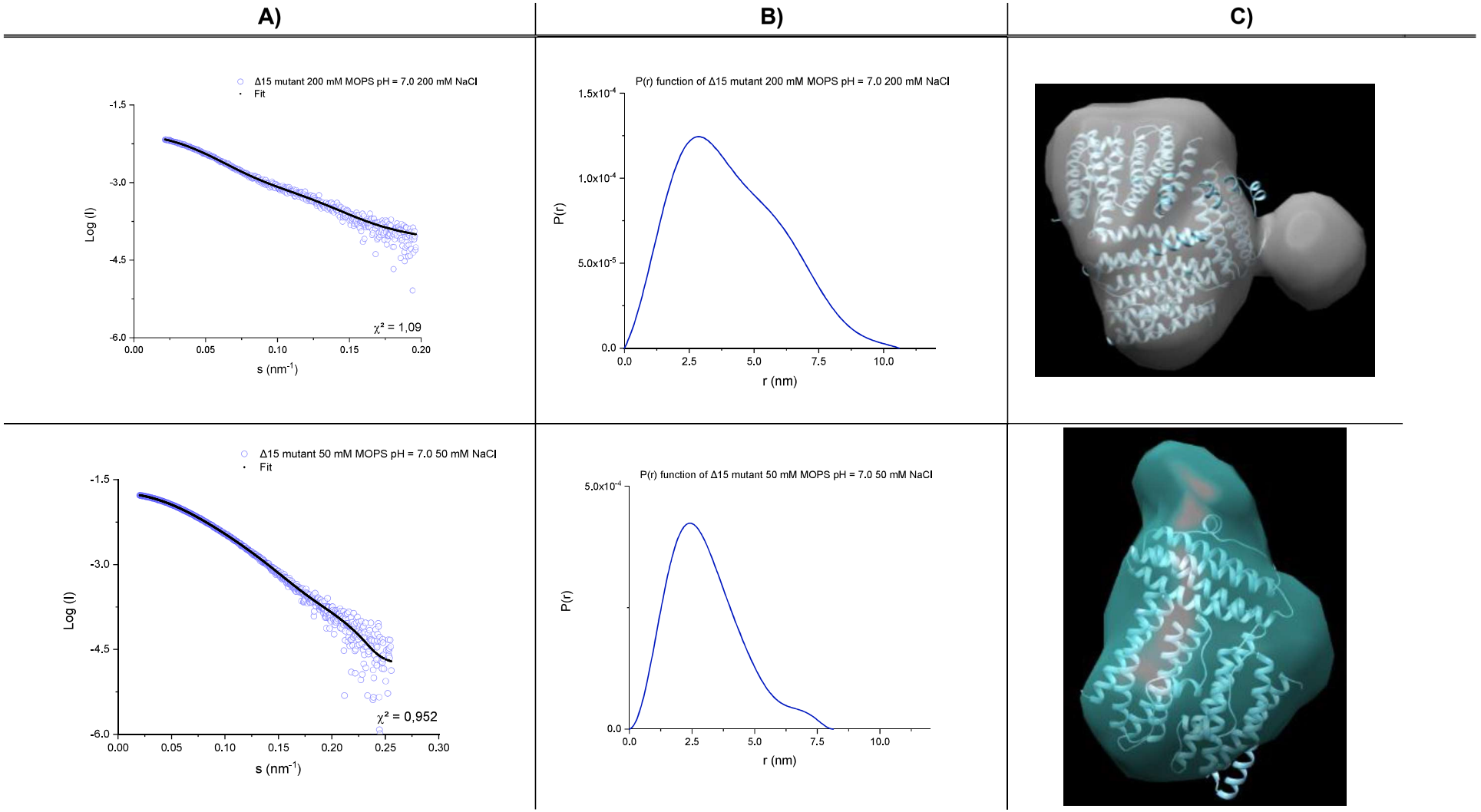

Figure 2-7: SAXS scattering profiles(A), P(r) functions (B), models and envelopes (C) of Δ15 mutant at high (top) and low (bottom) ionic strengths.

To understand the role of the N-terminal extension on the assembly process, the dynamics of the oligomerization was assessed by SEC (***Figure 2-8A***). The Δ15 mutant was eluted at 16 mL at low ionic strength (50 mM MOPS, pH 7.0, 50 mM NaCl), whereas the WT Dps chromatogram shows two elution peaks, at 13 mL and 16.4 mL of elution buffer. This result indicates that the WT co-exists in two different oligomeric forms in solution; the band eluted with 13 mL is attributed to the dodecameric form (12-mer) and the second one to a lower oligomeric form. Since this 3.4 mL difference in elution volume is not explained by the difference of the molecular mass of the 15 residues N-tail (that accounts for 20136 Da), one can conclude that the mutant does not fold into a dodecameric structure in these conditions.

An ionic strength dependence was carried out to assess if the change in ionic strength would favour an oligomeric form of Dps WT in detriment of the other. As ***Figure 2-8B*** shows, with the increase of the ionic strength, there is a shift in the intensity of the band correspondent to the lower oligomeric form to the 12-mer band in Dps WT. However, even at 300 mM NaCl there is still at least 10 % of the unknown oligomeric form (here designated X-mer). This result is interesting since it paves the way to control the assembly/disassembly of Dps WT cage by means of the ionic strength.

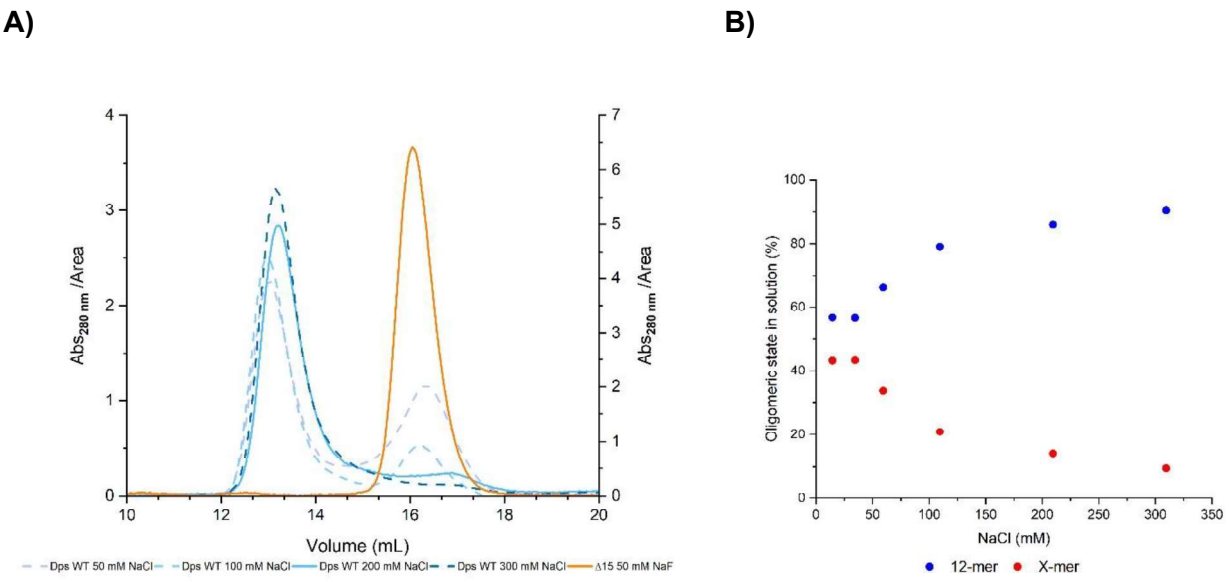

Figure 2-8: Oligomeric state dynamics of Dps WT and Δ15 mutant proteins (1 mg/mL) as a function of the ionic strength, in 50 mM MOPS, pH 7.0. (A) Chromatograms from the SEC step and (B) Dependence of the Dps WT oligomeric states with ionic strength.

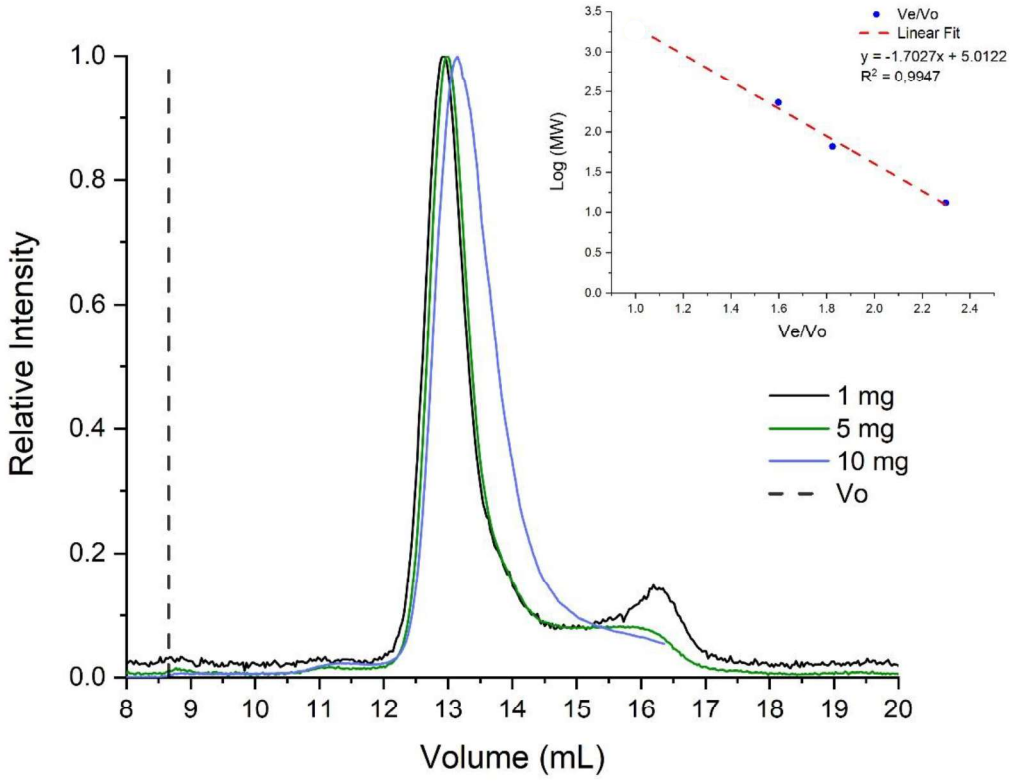

Figure 2-9: Effect of protein concentration on the oligomerization of Dps WT, at low ionic strength, analysed by SEC. V_o represents the void volume of the column.

To investigate the effect of protein concentration in the assembly dynamics at low ionic strength, a SEC step was conducted at three different concentrations of Dps WT. As shown in ***Figure 2-9***: Effect of protein concentration on the oligomerization of Dps WT, at low ionic strength, analysed by SEC. Vo represents the void volume of the column. a slight increase in the elution volume occurs with the increase of concentration of the injected sample. The absolute value of the difference between the elution peaks of the samples at 1 mg and 10 mg, 12.93 mL versus 13.14 mL, accounts for 2 monomers, taking into consideration the theoretical molecular mass of the protein, deduced from the amino acid sequence. Thus, the 1 mg-sample would be predominantly in a dodecameric form, shifting to a 10-mer form with a 10-fold increase in the concentration.

2.4 Conclusion

It was successfully shown that the absence of the first 15 residues at the N-terminal region hinders the assembly into a dodecameric structure at low ionic strengths. Depending on the ionic strength, the Δ15 mutant protein exits in an equilibrium between a trimeric and a hexameric forms, demonstrating that the N-terminal region is vital to a fully dodecameric structure. The assembly of the WT Dps protein is also dependent on the ionic strength, as well on the concentration, illustrating the importance of the buffer composition on the protein homogeneity. Based on these results one can reason that there is a dynamic equilibrium between different oligomeric forms. The physiological relevance of this dynamic process is yet to be unveil and more studies should be carried out to fully understand this mechanism. When taking into consideration the possibility to alter the structure of Dps WT in solution, future work can be envisioned to explore the encapsulation potential of different biotechnological relevant molecules inside the protein by means of ionic strength and concentration adjustments.

3

Biotechnological potential and applications of Dps protein

3.1 Overview

In this chapter the use of Dps as a nanoreactor for the synthesis of nanoparticles is addressed. In order to take advantage of Dps protein physiological functions as a detoxifier, a protocol published by Hansen et al [70] for lead sulfide quantum dots synthesis inside ferritin was adapted using Dps protein from *Marinobacter hydrocarbonoclasticus*. A reproducible method for the synthesis of nanoparticles in the interior cage of Dps, was established.

To validate the utilization of Dps as a versatile nanoreactor the protocol was tested for several metals, such as Fe, Cu, Pb, Cd, which were shown to be able to precipitate in the interior of the Dps protein cage when sulphur was present in the aqueous medium, forming nanoparticles of smaller than 5 nm in diameter.

The synthesis method was monitored by UV-Vis spectroscopy and the products of the different nanoparticle syntheses were characterized by the application of a wide range of techniques. These provide information regarding the maintenance of the structural integrity of Dps during synthesis, as well as information regarding the nanoparticle properties produced inside the cage, such as, size and crystal structure.

The manuscript "Supramolecular protein polymers using mini-ferritin Dps as the building block" is also included in this chapter. Herein a site directed mutant of Dps protein from *Marinobacter hydrocarbonoclasticus* was used as a building block to develop new supramolecular assembly complexes, with enhanced iron uptake capacity.

3.2 Experimental procedures

3.2.1 Nanoparticle synthesis the inside Dps protein nanocage

The synthesis of nanoparticles was carried out by sequential additions of aqueous solutions of the metal of interest (***Table 3-1***) and sodium sulphide ($Na_2S{\cdot}9H_2O$, Sigma Aldrich) to a 1 μM Dps buffered in 200 mM MOPS, pH 7.0, 200 mM NaCl, under moderate stirring at room temperature. After each addition of 100 atoms of the metal of interest per protein, the reaction mixture was incubated for 10 min, after which sodium sulphide was added (stoichiometry molar ratio to the metal) and let react for 10 minutes. Sequential additions were performed up to a molar ratio of 1500 metal atoms per Dps dodecamer.

The syntheses were followed by UV-Vis spectroscopy, using a Thermo Scientific™ Evolution 201, scanning between 220 and 900 nm; the absorbance at both 280 nm and 350 nm were measured.

The synthesis of Dps-NP complexes was performed in duplicate for each metal of interest.

Table 3-1: Precursor salts for nanoparticle syntheses

Salt	Chemical formula	MW (g/mol)
Copper nitrate (II)	$Cu(NO_3)_2 \cdot 2(5H_2O)$	232.59
Lead acetate (II)	$Pb(C_2H_3O_2)_2 \cdot 3H_2O$	379.33
Cadmium chloride	$CdCl_2 \cdot H_2O$	201.30
Ferrous sulphate	$FeSO_4 \cdot 7H_2O$	278.01

3.2.2 Spectroscopic methods and techniques

The synchrotron radiation SRCD spectra of both WT Dps protein and the Dps-PbS complex were obtained at the SRCD facility at the UV1 beamline on the storage ring ASTRID at the Institute for Storage Ring Facilities (ISA), University of Aarhus, Denmark, as mentioned before. The obtained spectra were recorded with 1 nm steps and a dwell time of 2,725 s per step in triplicates, using 0.01008 cm quartz cells (SUPRASIL, Hellma GmbH, Germany), for the wavelength range of 180–280 nm, using 200 mM MOPS, pH 7.0, 200 mM NaF as buffer. Two types of observation chambers were used: i) a standard, dual Peltier chamber for which samples are mounted vertically; and ii) a periscope chamber that features a horizontal mounting position. The use of the latter eliminates or reduces the effect of displacement due to aggregation/precipitation of protein samples or buffer density changes. In order to calculate the contribution to α-helix and β-sheet content in the secondary protein structure analysed suitable spectra were fitted using Dichroweb online server[98]. The fitting was conducted using CDSSTR method and SP175 reference data set [99].

The hydrodynamic radii, R_H of Dps WT and Dps-nanoparticle (NPs) complexes after synthesis were determined using DLS utilizing a SZ-100 nanopartica, Horiba Scientific. The buffer for all samples was 200 mM MOPS, pH 7.0, 200 mM NaCl, and protein-NP complex at 1 μM concentration for all samples, after 20 min centrifugation at 14500 rpm. The hydrodynamic diameter, and polydispersity index (PDI) were obtained as the mean ± standard deviations (SD) of triplicate measurements.

The structural analysis of the samples was done by X-ray diffraction (XRD, PANalytical, model X´Pert Pro) in Bragg-Brentano geometry with Cu-Kα line radiation (λ=1.5406 Å), from 10º to 100º with a step of 2º/min at the Analytical Laboratory, Department of Chemistry, Universidade Nova de Lisboa. After syntesis all samples after were dialyzed overnight against milli-Q H_2O and lyophilized before XRD measuring.

The Raman spectra of a 150 μL droplet of Dps-NP complexes in different MOPS buffer concentrations, pH 7.0, containing in 200 mM NaCl, were acquired on a clean glass slide, and

recorded using a Horiba Labram Raman microscope spectrometer equipped with a 9.1 mW 633 nm laser that was focused by a 60x Olympus immersion objective for measurements in solution. Data were processed using Origin 9.1. The Raman spectra (dried samples) were recorded by allowing 10 μL of Dps-NP solution in a buffered solution (pH 7.0) containing 200 mM NaCl to evaporate in a glass substrate and a 9.1 mW 633 nm laser that was focused by a 50x Olympus objective.

3.2.3 Biochemical characterization

The oligomeric form of Dps WT and Dps-NP complexes were assessed by size exclusion chromatography coupled to a FPLC system equipped with a UV detector (ÄKTAprime Plus) and a Superdex™ 200 10/300 GL column, both from GE Healthcare Life Sciences. The column was pre-equilibrated with 200 mM MOPS, pH 7.0, and an equal molarity of NaCl as each protein samples to investigate the effect of the ionic strength. 1 mL of each protein solution at 1 mg/mL was applied to the column with a flow rate of 0.5 mL/min, in 200 mM MOPS, pH 7.0, 200 mM NaCl.

3.2.4 Microscopy techniques - TEM

The transmission electron microscopy images were obtained in a JEOL JEM – 2100 – HT 80-200 kV, at International Iberian Nanotechnology Laboratory, Braga Portugal. A small drop of sample was placed on a carbon film supported copper grid, CF400-Cu Electron Microscopy Sciences, and allowing it to dry at room temperature analysed without staining. Image J [104] was used to measure the diameters of the analysed samples.

3.3 Results and discussion

The radius of the inner cavity of Dps, as per the 3D structure of Dps homologs is considered to be 2.25 nm [105]. Assuming a perfect sphere shape of the cavity this corresponds to a total volume of 4.77×10^{-20} cm^3. With this value it is then possible to estimate the number of metal atoms that could theoretical be fitted in the interior of the hollow cage, as presented in ***Table 3-2***

Table 3-2: Theoretical capacity of Dps WT to accommodate metal atoms based on the protein inner volume.

Metal	Density (g/cm^3)	MW (g/mol)	Weight ($g \times 10^{-19}$)	Moles ($\times 10^{-21}$)	Number of metal atoms
Fe	4.84	87.91	2.31	2.63	1582.18
Cu	4.76	95.61	2.27	2.38	1430.70
Pb	7.6	239.3	3.63	1.52	912.680
Co	2.51	60.07	1.20	1.99	1200.78
Cd	4.82	144.5	2.30	1.59	958.850

In all experiments (***Figure 3-1***) it was possible to observe a change of colour when sulphur was added to a metal pre-loaded protein sample. The control, prepared by adding sodium sulphide to Dps, as expected, did not change the absorbance in any of the monitored wavelengths. This is indicative of the formation of a mineral core inside the protein cage, similar to the iron mineralisation process catalysed by Dps [31]. It is also important to remark that if the order of additions was reversed, first sulphur and then the metal of interest, a precipitate was clearly visible followed by the formation of a deposit, due to the low solubility of sulphides. This observation demonstrates the capacity of WT Dps in solubilizing the reaction final products.

The synthesis of both Dps-CuS and Dps-CdS complexes (***Figure 3-1***) progressed in a similar manner, with a steady slope of metal incorporation reaching a plateau at higher metal: protein ratios. The observed shift in the absorbance is due to the diffraction of light given the different opacities of the solutions reached. In the case of Dps-PbS complex this is even more clear, and the maximum is reached at around 800 atoms per Dps, which is lower than all other metals studied, and in agreement with the lower number of these atoms that should theoretically fit inside Dps WT hollow cage (***Table 3-2***). It is important to remark that sulphur atoms were not considered in the calculations and that the stoichiometry between the metal of interest and sulphur inside Dps WT was not determined.

For the Dps-FeS complex, the shape of the spectra is also different. The dependence of the absorbance at 350 nm (or 280 nm) seems to be biphasic and sigmoid, distinct than the aforementioned ones. No plateau was reached. This behaviour could indicate that the metal is being incorporated into the interior cavity of Dps WT at different rates or by distinct mechanisms. As before, no precipitates were observed on solution, indicating the ability of the protein to solubilize otherwise insoluble precursors

Figure 3-*1* also shows that it is possible to synthesise mixed metal minerals inside the protein nanocage, such as CdPbS, by sequential additions of both metals in a alternating manner.

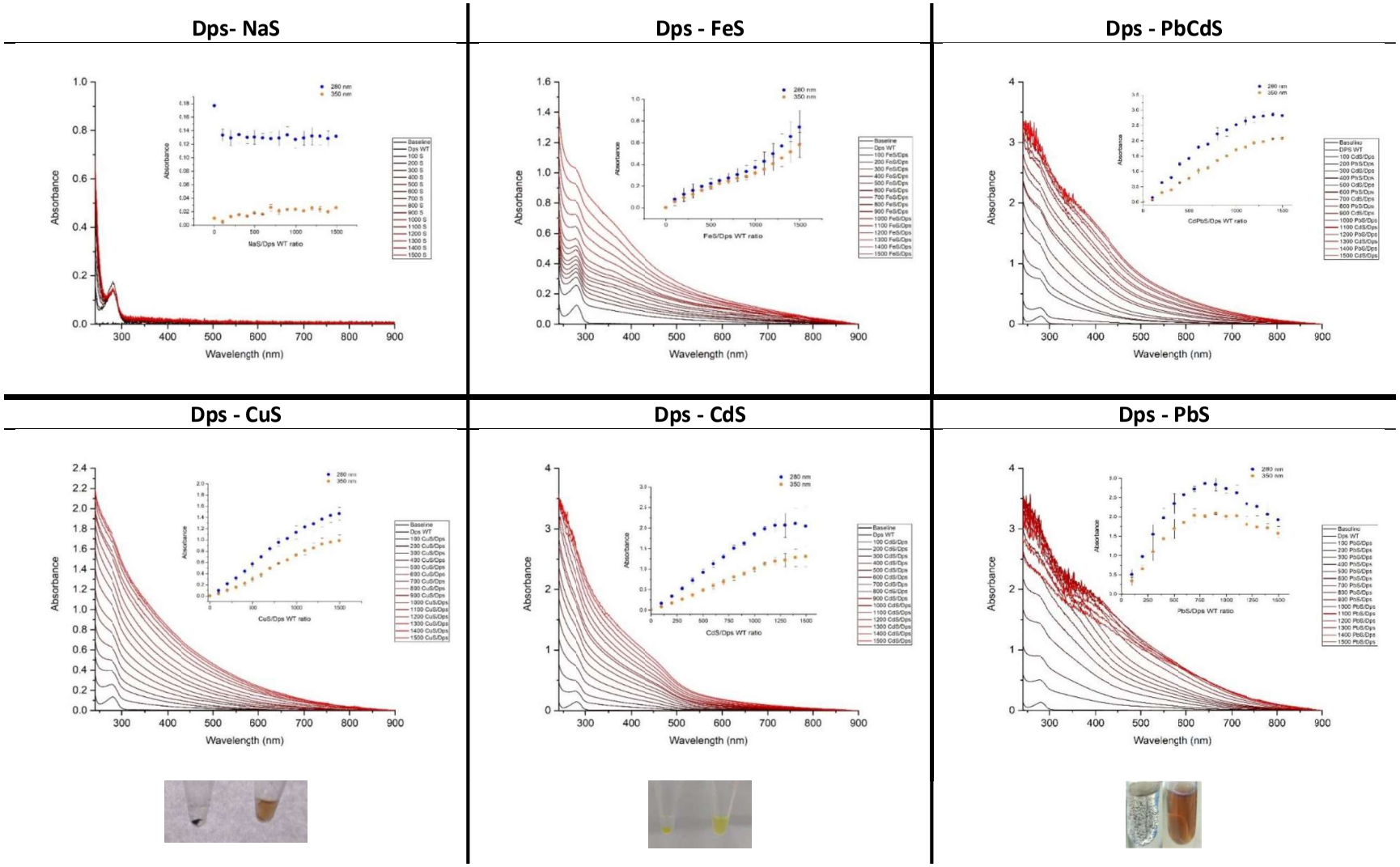

Figure 3-1: Syntheses of Dps-NP complexes monitored by UV-Vis spectra.

DLS was applied to infer about the effect of the formation of NP (metal sulphide complexes) in the interior of the protein. Samples loaded with 800 atoms per Dps molecule were analysed and compared with Dps WT, as shown in ***Table 3-3.***

Based on these results one can assume that there is no change in the size of the protein, during or after the NP synthesis procedure. However, as previously stated, the colour of the solution when metals are present changes dramatically which therefore indicates the formation of mineral core inside the nanocage.

Table 3-3: DLS data of Dps -NPs

Sample	Diameter (nm)	SD (nm)	PDI
Dps-WT	9.47	1.40	0.37
Dps-PbS	9.70	1.30	0.17
Dps-CdS	9.57	1.37	0.13

Comparison of the SRCD spectra of Dps WT, Dp-PbS, Dps-CdS and Dps-FeS complexes reveals that the content in secondary structure is mainly conserved after NPs synthesis, and that the melting temperature of the sample is not greatly affected in the case of Dps-PbS and Dps-FeS. Contrarily, a dramatic change in the CD spectrum of the Dps-CdS complex is noticeable. Here the T_m increased by 10 ºC and the SRCD spectrum at room temperature also shows lower molar absorptivity. It is clear that the Dps-PbS complex is the most disruptive for the secondary structure of the protein showing the lower content of α-helix and the larger increase in unordered structures, when compared with Dps WT. On the other hand, Dps-FeS complex presents the higher content in α-helix of all samples. The Dps-CdS complex secondary structure had to be fitted with a different data set than the remaining spectra, which can interfere in the α-helix content (***Figure 3-2*** and ***Table 3-4***)

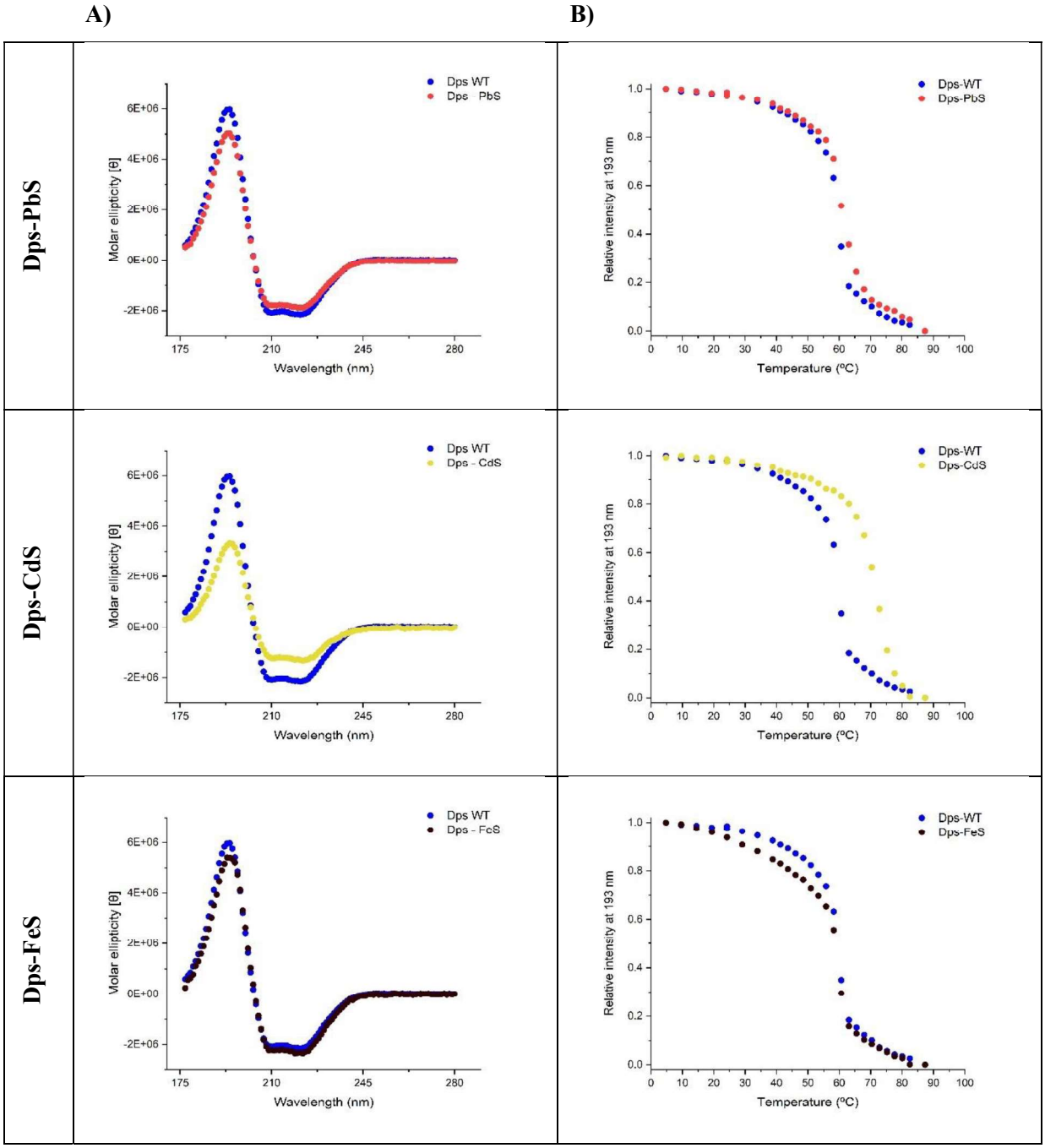

Figure 3-2: SRCD analysis of Dps WT, Dps-PbS, Dps-CdS and Dps-FeS. Spectra at 25 ºC (A) and melting curves (B) in 200 mM MOPS pH 7.0 200 mM NaF.

Table 3-4: Secondary structure content of Dps WT, Dps-PbS, Dps-CdS and Dps-FeS complexes at room temperature, determined using Dichroweb CDSSTR SP (175-240 nm) and CDSSTR SP (190-240 nm) for Dps – Cds.

	α-Helix	β-Sheet	Turns	Unordered	Total	NRMSD
Dps-WT	0.79	0.08	0.04	0.09	1	0.005
Dps-PbS	0.65	0.02	0.09	0.15	1	0.008
Dps-CdS	0.87	0.03	0.06	0.04	1	0.005
Dps-FeS	0.93	0.01	0.02	0.04	1	0.004

Microscopy analysis by TEM (***Figure 3-3***) allowed to estimate the size of the mineral core of Dps-PbS and Dps-FeS compexes. Both are monodisperse in size, which demonstrates the synthesis in a confined space as the interior of the nanocage protein. The average size diameter of the metal sulphide nanoparticles *(n = 50*) was estimated to be 3.33 nm, with a standard deviation of 1.08 nm for Dps-PbS complexes and an average size diameter of 2.66 nm with a standard deviation of 0.65 nm for Dps-FeS. These value are in agreement with the size of inner cavity of the protein (4.5 nm in diameter) [105], and therefore also underpins the synthesis of the material in the interior of the protein.

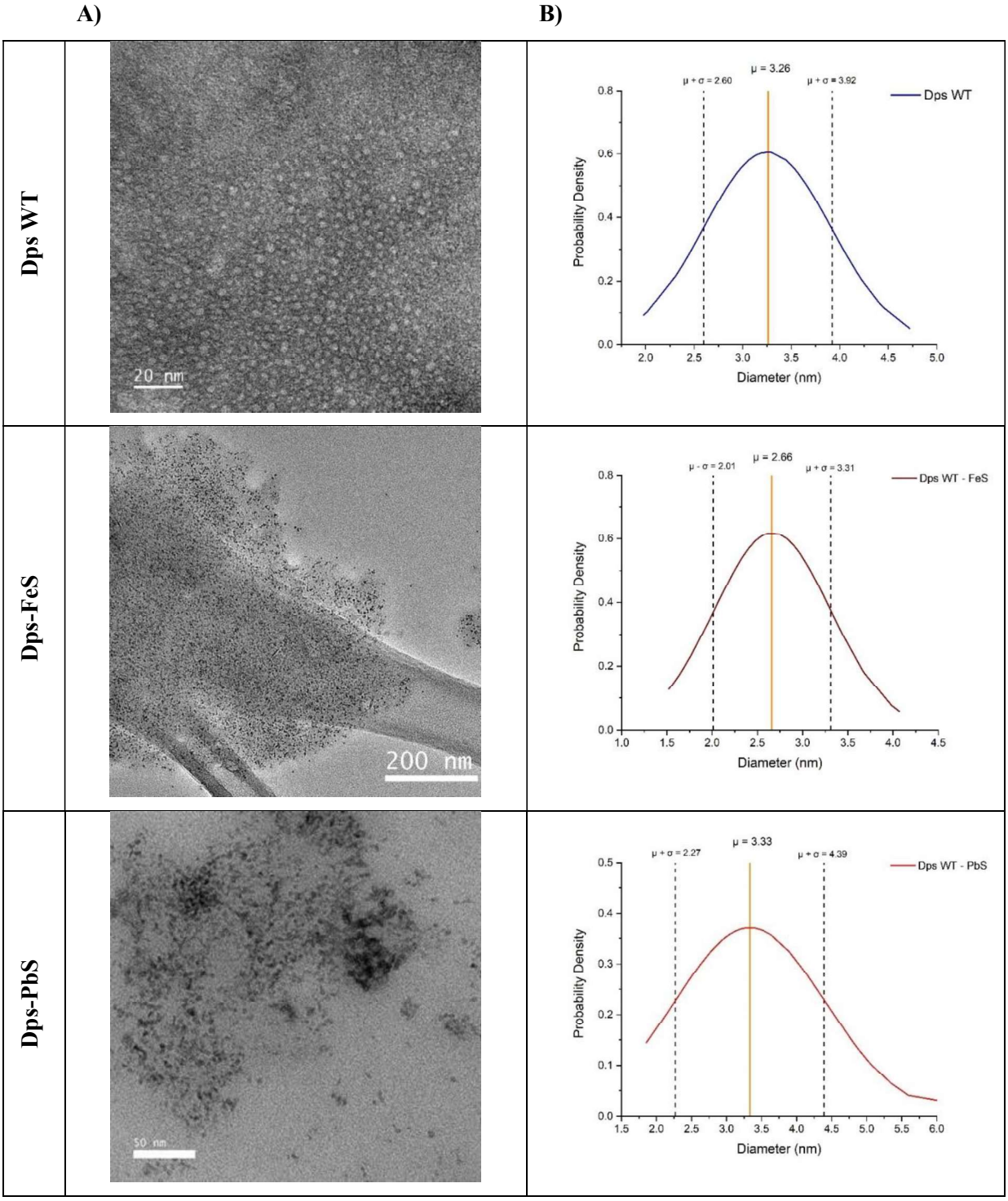

Figure 3-3: TEM images of Dps-PbS, Dps-FeS complexes without staining, and Dps WT uranil staining (A) with normal distribution of 50 cores (B).

To obtain data about the homogeneity and molecular mass of the Dps WT protein after NP synthesis, SEC was applied (***Figure 3-4***). A major chromatographic band was obtained at an elution (Dps-CdS and Dps-PbS complexes) that agrees with the dodecamer structure of the protein. Minor bands, however, with lower molecular mass (elution volume above 30 mL) that can be attributed to smaller oligomers of the protein.

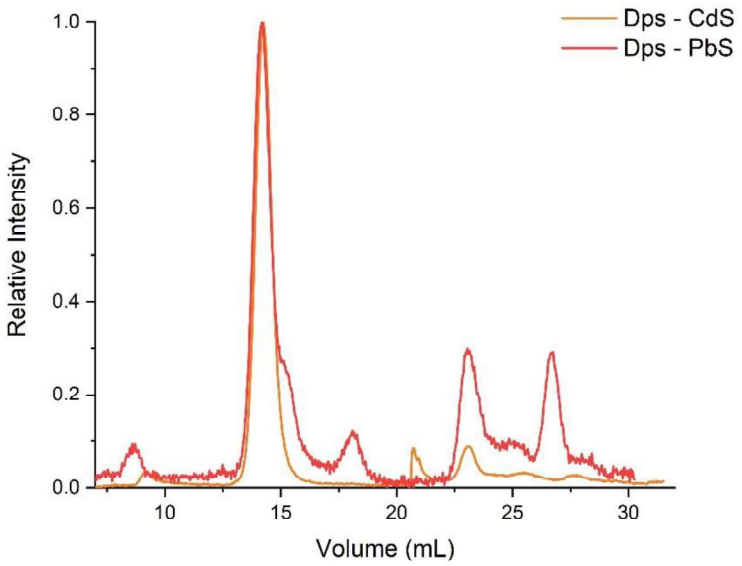

Figure 3-4: SEC (Superdex 200 10/300) Dps-Cds and Dps-PbS chromatogram in 200 mM MOPS, pH 7.0, 200 mM NaCl.

Crystallinity of Dps-PbS NPs was assessed by DRX. The narrow peaks suggest that the materials are highly crystalline. Three peaks at 25.89°, 31.64° and 45.15° were detected on the diffractogram as presented in ***Figure 3-5***. The closest related pattern found, could be attributed to eclarite, a mineral with a X-ray diffraction at 3.414(1), 2.014(0.8) and 2.893(0.7) and corresponding angles of 26.08°, 44.97° and 30.88°, with an orthorhombic dypiramidal geometry (***Figure 3-6***) [106][107]. The chemical formula of this mineral is $Pb_9(Cu,Fe)Bi_{12}S_{28}$, which is not expected to be present in Dps-Pbs samples due to the absence of bismuth. According to the Dana classification [108] eclarite is classified as 03.06.10, Sulfides - Sulfosalts where $2 < z/y < 2.49$ - $(A+)_i\ (A++)_j\ [B_y\ C_z]$, A = metals, B = semi-metals, C = non-metals. One can speculate that other semi-metal instead of bismuth might be present in the mineral composition and therefore cause the slight peak variations observed, or the most likely case is that we are in the presence of more than one type of mineral, which makes the XRD peak attribution complex.

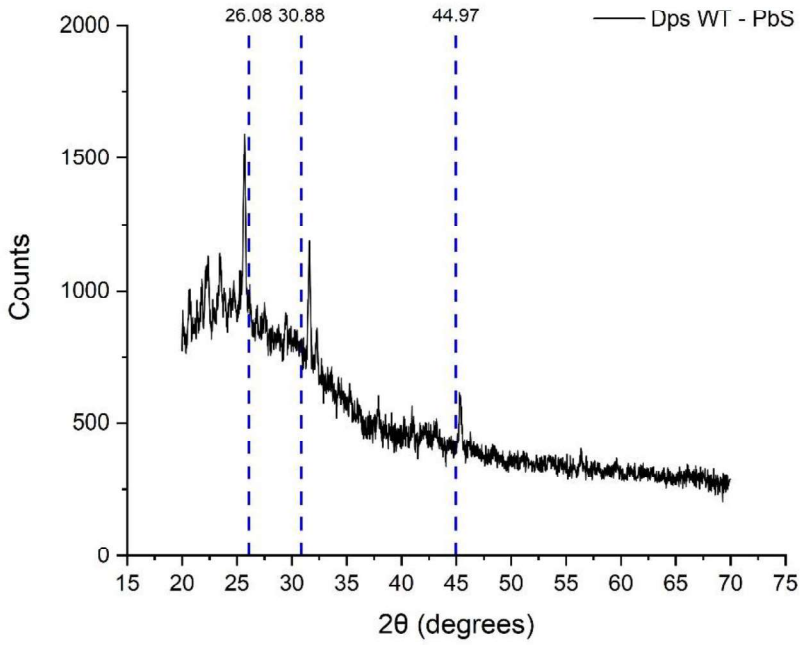

Figure 3-5*:* XRD diffractogram of Dps-Pbs nanoparticles with the closest related peak's position of eclarite traced.

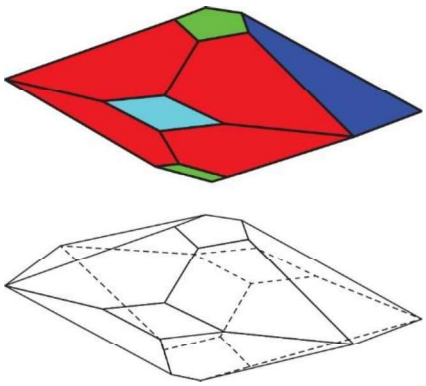

Figure 3-6: Orthorhombic dipyramidal crystal shape 2/m 2/m 2/m.

Raman spectroscopy was employed to characterize Dps-Fe and FeS complexes, both in solution, glass and silicon substrates, in parallel with Dps-PbS in solution and silicon.

The first question was to elucidate the contribution of MOPS buffer. MOPS absorbs in the region of 1050 cm^{-1} and between 800 and 900 cm^{-1} as shown in ***Figure 3-7***. Upon dialysis against milli-Q water, only the silicon peaks are detected between 600 and 700 cm^{-1} and 900 to 1000 cm^{-1}.

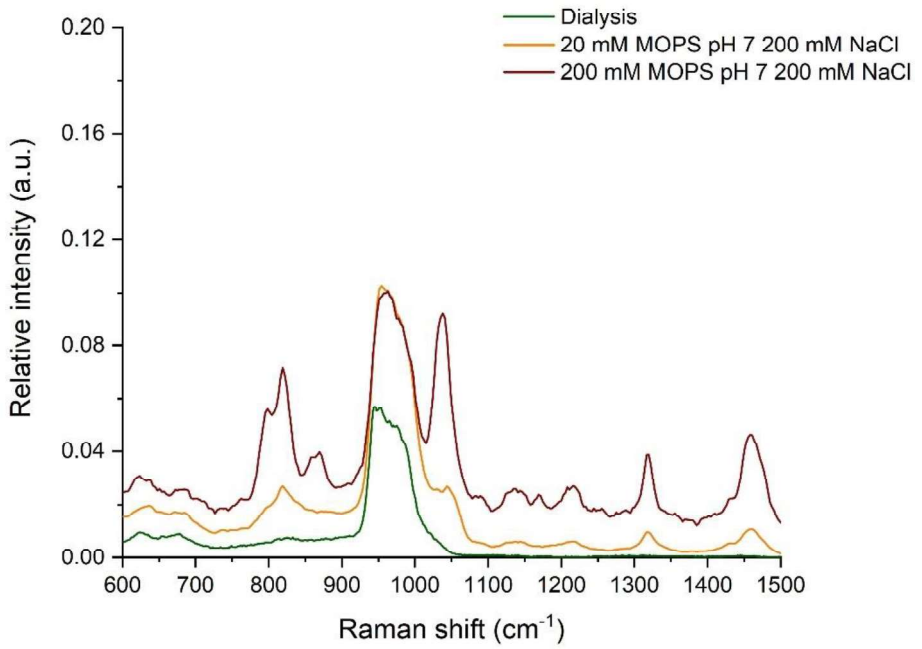

Figure 3-7: Raman spectra of MOPS buffer at pH 7.0 (200 mM MOPS 200 mM NaCl, red, 20 mM MOPS 200 mM NaCl , yellow, and after dialysis against 200 mM NaCl), green. 633 nm laser 1 mW power 10 acquisitions 30 s per acquisition. Each spectrum corresponds to the average of 3 measurements.

When Raman measurements of Dps WT in solution were carried out in 200 mM NaCl pH 7.0 (after dialysis) it was possible to identify the characteristic α-helix peaks of the protein, between 1300 and 1340 cm^{-1} (***Figure 3-8***).

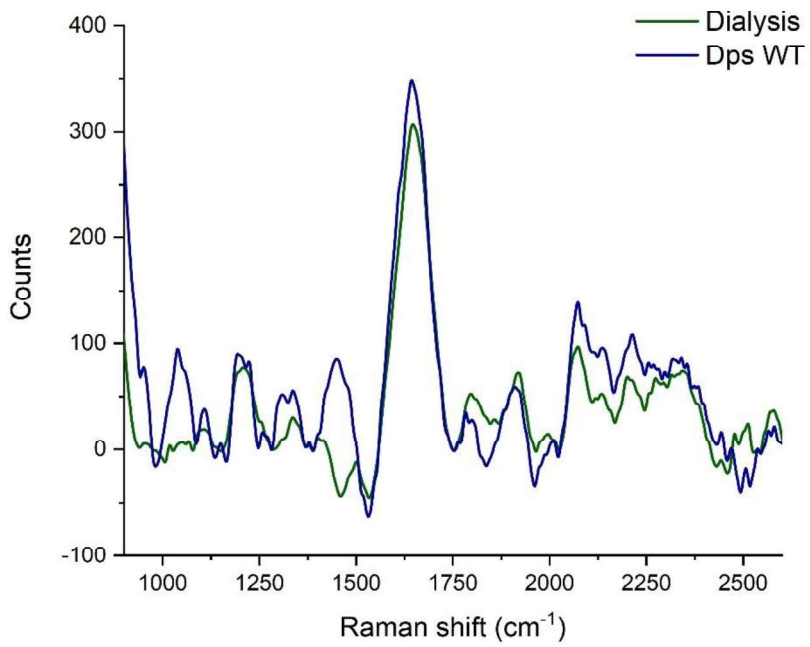

Figure 3-8: Raman spectra of Dps WT, blue. and of 200 mM NaCl pH 7.0 green 633 nm laser 5 mW power 10 acquisitions 120 s per acquisition. Each spectrum corresponds to the average of 3 measurements.

Reaction with Fe^{2+} ions in the presence of H_2O_2, changes the Raman spectrum. Characteristic peaks of goethite, α–FeO(OH), at 292 cm^{-1}, 384 cm^{-1} and 545 cm^{-1} (RRUFF ID: X050091) [109] seem to be present at low iron loads (12 Fe/Dps), that are no longer detected in the 256 Fe/Dps spectrum (***Figure 3-9***). This evidence suggests the formation of different minerals or rearrangements inside the Dps WT depending on the concentration of metal ions.

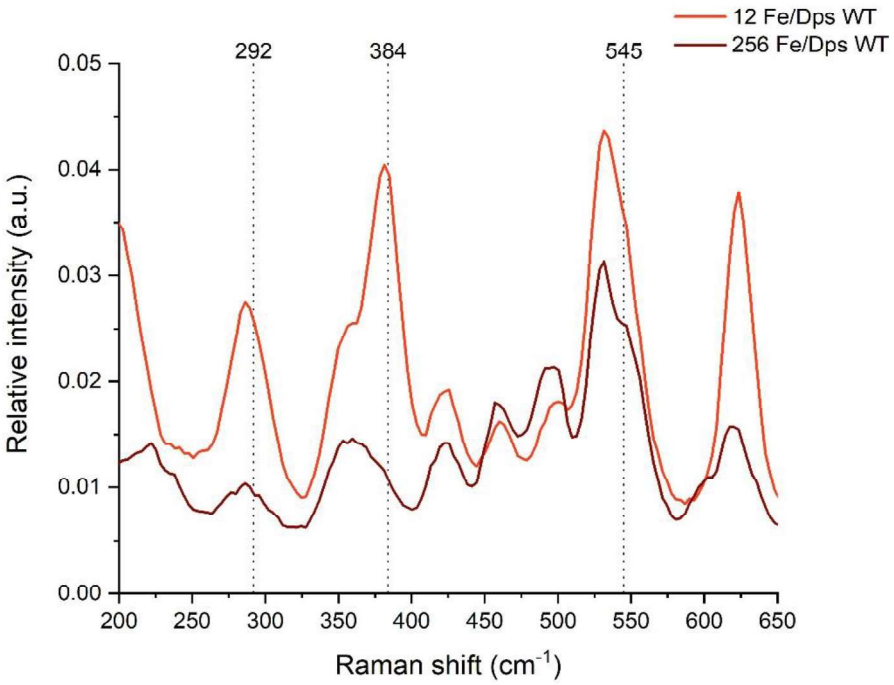

Figure 3-9: Average of 3 Raman spectra in silicon substrate of iron loaded Dps WT, 12 Fe/Dps, and 256 Fe/Dps, in 20 mM MOPS, pH 7.0, 200 mM NaCl. 633 nm laser 5 mW power 10 acquisitions 30 s per acquisition.

The peak at 1050 cm^{-1}-peak present in the Raman spectrum of the WT Dps is shifted to lower wavelengths upon reaction with ferrous iron in the presence of H_2O_2 (***Figure 3-10***). The new peak is typical of magnetite (RRUFF ID: R140861), a mineral with a chemical formula of $FeO{\cdot}Fe_2O_3$, suggesting the presence of a mixed valence Fe^{2+}/Fe^{3+} iron nanoparticles inside the Dps WT [110].

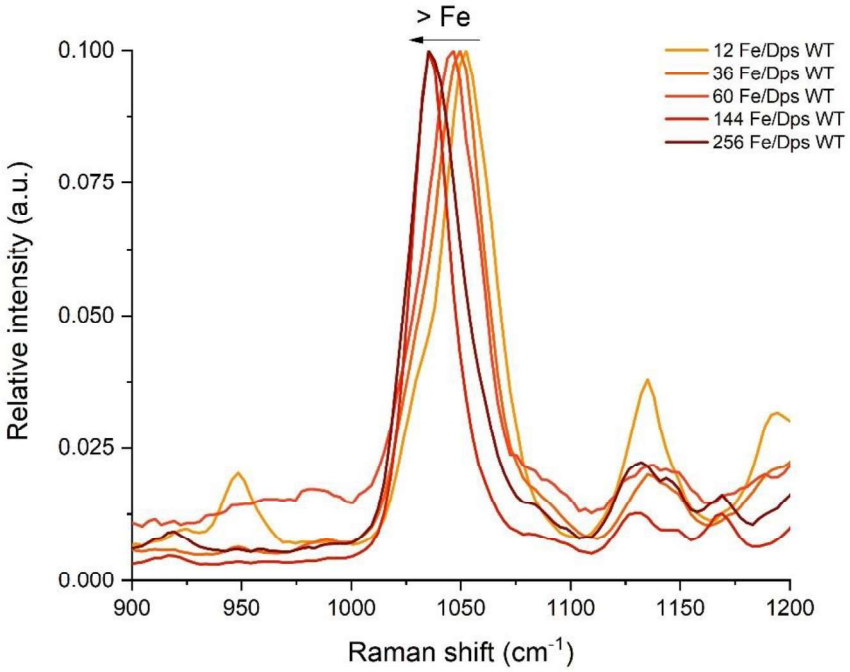

Figure 3-10: Effect of Dps WT iron loading in the presence of H_2O_2 on the Raman spectra in 20 mM MOPS, pH 7.0, 200 mM NaCl. 12 Fe/Dps (yellow); 36 Fe/Dps (light orange); 60 Fe/Dps (orange); 144 Fe/Dps (dark orange) and 256 Fe/Dps (brown). 633 nm laser 5 mW power 10 acquisitions 30 s per acquisition. Each spectrum corresponds to the average of 3 measurements.

When Dps WT with different loads of iron was measured using a glass substrate the different peak changes related to the increase of iron, detected between 300 and 600 cm^{-1}, which is the region characteristic of goethite mineral FeO(OH) [110], was evident. No changes were observed for the α-helix peaks at 1300 cm^{-1}, 742 cm^{-1}, 857 cm^{-1}, 940 cm^{-1}, 1005 cm^{-1}, 1176 cm^{-1} and 1211 cm^{-1}, which suggests a conservation of the protein structure with higher iron ratios (***Figure 3-11*** and ***Table 3-5***), in agreement with the CD data (***Figure 3-2*** and ***Table 3-4***)

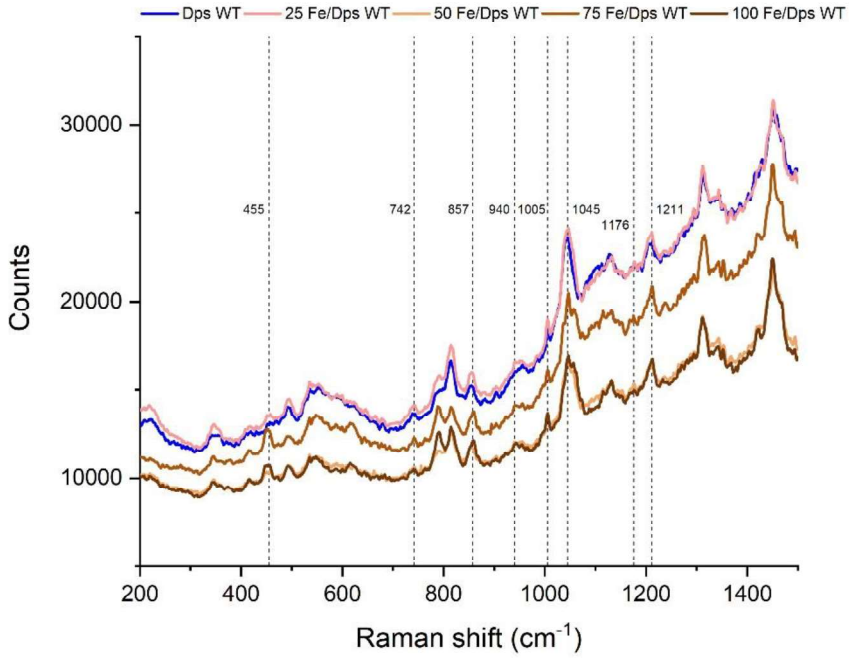

Figure 3-11: Average of 3 Raman spectra in glass substrate of Dps WT blue and from light orange to brown of: iron loaded Dps (25 Fe/Dps, 50 Fe/Dps, 75Fe/Dps, 100 Fe/Dps in 20 mM MOPS, pH 7, 200mM NaCl). 633 nm laser 5 mW power 10 acquisitions 60 s per acquisition.

Table 3-5: Assignment of peaks in the Raman of apo-Dps

Raman peak (cm^{-1})	Assignment	Reference
455	Fe-O, Fe-NO complexes	[111]
742	Tryptophane indole ring	[112]
857	Tyrosine H- bonding	[112]
940	α-helix, C–C, C–N stretch	[113]
1005	Phenylalanine	[114]
1176	Tyrosine, CH_3	[115]
1211	Tyrosine	[115]

When Dps is incubated with Fe^{2+} ions and sulphur for the synthesis of a nanoparticle of FeS inside the protein, the Raman spectrum acquired in solution in 20 mM MOPS, pH 7.0, 200 mM NaCl buffer, is changed at around 1340 cm^{-1}, which might indicate a loss of protein structure. This is also seen at higher wavelengths between 2000 and 2300 cm^{-1} (***Figure 3-12***).

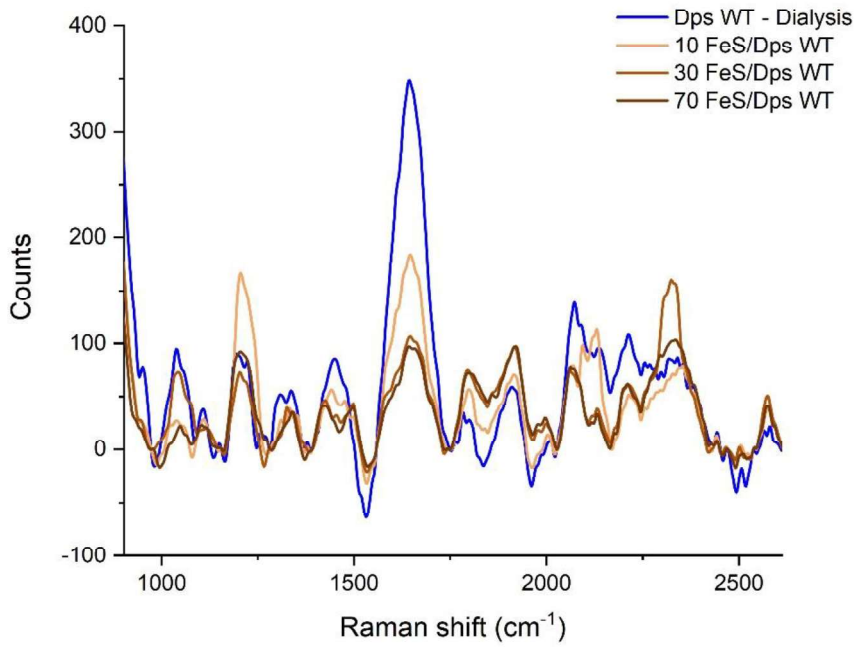

Figure 3-12: Average of 3 Raman spectra of Dps-FeS in solution (20 mM MOPS, pH 7.0, 200mM NaCl) effect of the iron loads from light orange to brown (10 FeS/Dps, 30FeS/Dps, 70 FeS /Dps). 633 nm laser 5 mW power 10 acquisitions 60 s per acquisition.

When Dps-PbS was analysed in the dry state using silicon as substrate it was clear that several changes are present in the Raman spectra mainly in the region between 300 and 450 cm^{-1} and at wavelengths above 600 cm^{-1} specially in the region of 750 to 900 cm^{-1}, where three clear peaks are visible in the spectrum of the protein with 512 PbS/Dps (***Figure 3-13***). These spectral features are closely related with the Raman spectrum of the mineral anglesite, $Pb(SO_4)$ (RRUFF ID: R050408) [110].

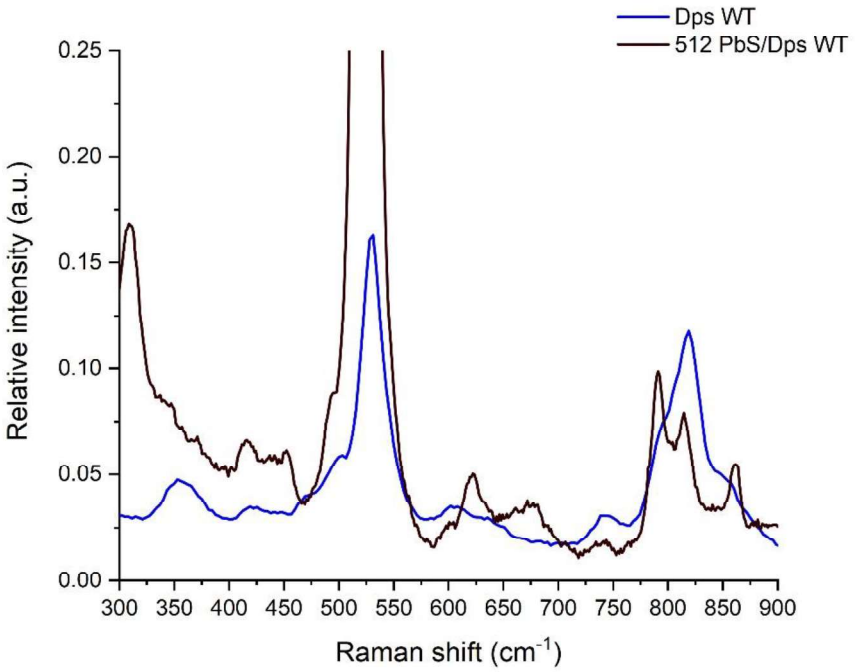

Figure 3-13: Average of 3 Raman spectra in the dry state in silicon substrate NaCl of: Dps WT and 12PbS/Dps in 200 mM MOPS pH 7.0 200mM, 633 nm laser 5 mW power 10 acquisitions 60 s per acquisition.

The in solution Raman spectra of the Dps-PbS complexes (various Pb/protein ratios) showed no significant differences, except in the region between 1100 and 1250 cm^{-1}, most likely due to the presence of tyrosine residues [115]. It was not possible to explain the peaks present between 1750 and 1900 cm^{-1} (***Figure 3-14***).

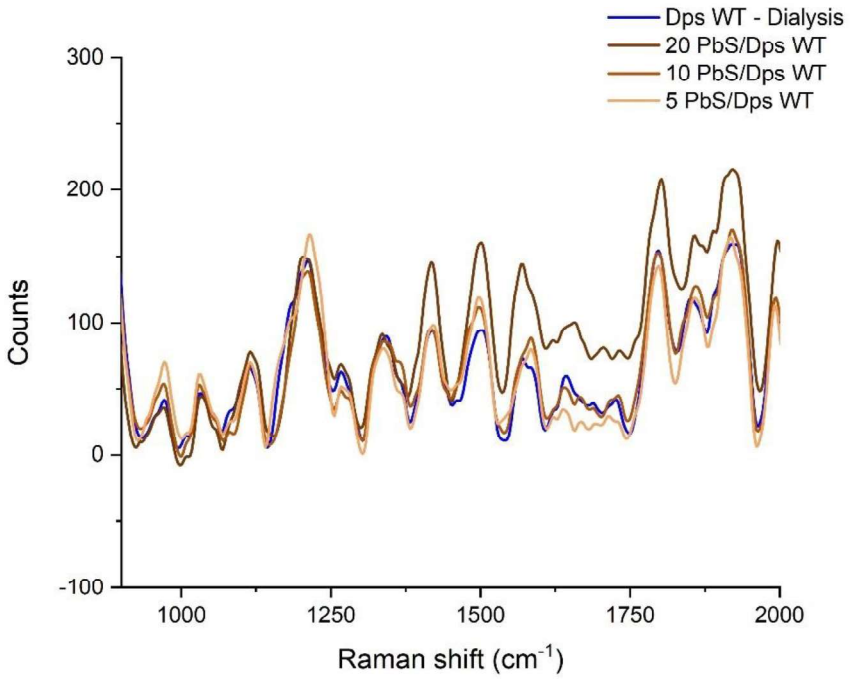

Figure 3-14: In solution average of 3 Raman spectra of Dps WT, 5PbS/Dps, 10PbS/Dps and 20PbS/Dps complexes in 20 mM MOPS pH 7 200mM NaCl from light orange to brown. 633 nm laser 5 mW power 10 acquisitions 60 s per acquisition.

Dps-CdS nanoparticles were also accessed by Raman spectroscopy, however due to high fluorescence of the sample no reliable spectra were acquired.

3.4 Conclusion

The simple methodology procedure developed and utilised for the synthesis of metal nanoparticles using Dps WT as a nanoreactor proved to be efficient. UV-Vis spectroscopy data allowed to monitor the synthesis reaction. The absence of precipitates, evaluated by visual observation and by UV-Vis spectroscopy, indicated that the nanoparticles were synthesised inside the Dps. SRCD data showed no drastic change of the secondary structure of the protein after the synthesis procedure. TEM imaging corroborated this observation.

SEC data indicated that after synthesis most of the protein was eluted in a 12-mer form further confirming the previous hypothesis. Characterization of the protein-nanoparticle complexes was not completely attained, and further work needs to be performed. XRD data showed that Dps-PbS contains a material most likely closely related with eclarite, and Raman spectroscopy indicated the present of anglesite. The most probable explanation is the existence of

several mineral forms inside Dps. The Raman characterization of Dps with oxidised iron indicated the presence of goethite and the probable magnetite, in line with the literature.

Further characterization of these different protein-nanoparticle complexes should be considered for further work, for instance the XRD characterization of the remaining Dps-NPs, also the fluorescence properties of these materials should be considered.

3.5 Manuscript: Supramolecular protein polymers using mini-ferritin Dps as the building block

M. Raquel Pacheco[‡a], João P. Jacinto[‡a], Daniela Penas[a], Tomás Calmeiro[b], Ana V. Almeida[a], Miriam Colaço[a], Elvira Fortunato[b], Nykola C. Jones[c], Søren V. Hoffmann[c], M. Manuela A. Pereira[d], Pedro Tavares[a] , Alice S. Pereira[a]

a Molecular Biophysics Laboratory, UCIBIO/Requimte, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

b CENIMAT/i3N, Departamento de Ciências dos Materiais, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

c ISA, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark

d Molecular Synthesis Group, LAQV/Requimte, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

‡ These authors contributed equally to this work.

The present work has been published on the 10th of November 2020 by the Royal Society of Chemistry (DOI: 10.1039/d0ob01702g).

Abstract

A missense mutant of a Dps protein (DNA-binding protein from starved cells) from *Marinobacter hydrocarbonoclasticus* was used as a building block to develop a new supramolecular assembly complex which enhances the iron uptake, a physiological function of this mini-ferritin. The missense mutation was conducted in an exposed and flexible region of the N-terminal, wherein a threonine residue in position 10 was replaced by a cysteine residue (DpsT10C). This step enabled a click chemistry approach to the variant DpsT10C, where a thiol–ene coupling occurs. Two methods and two types of linker were used resulting in two different mini-ferritin supramolecular polymers, which have maintained secondary structure and native iron uptake physiological function. Electrophoretic assays and mass spectrometry were utilized to confirm that both functionalization and coupling reactions occurred as predicted. The secondary structure has been investigated by circular dichroism and synchrotron radiation circular dichroism. Size and morphology were obtained by dynamic light scattering, size exclusion chromatography and atomic force microscopy, respectively. The iron uptake of the synthesized protein polymers was confirmed by UV-Vis spectroscopy loading assays.

3.5.1 Introduction

Protein cages are hollow nanoparticles which have a multitude of applications in nanobiotechnology.[43][116][117]They can be used as vessels for chemical reactions, [118][119] as precursors for the synthesis of nanostructured materials[120] or as drug delivery systems.[121][122] Dps (DNA-binding protein from starved cells) proteins, also named mini-ferritins, are bacterial proteins from the ferritin family that protect DNA from oxidative stress by

reversible storage of oxidized iron species.[31] Dps proteins are globular homopolymers, composed of 12 equal subunits folded into four-helix bundles that self-assemble into an almost spherical shape. The shell-like structure has a diameter of approximately 8–9 nm, with a spherical hollow cavity with approximately 4–5 nm in diameter.[123] Due to their well-defined structure and size, protein cages from the ferritin family are exceptional building blocks for the synthesis of supramolecular nanostructures that are non-toxic, biodegradable, and stable under a wide range of conditions such as temperature and pH.[124]

Contrary to synthetic nanocapsules or lipidic compartments, these proteins can also be chemically or genetically modified at specific amino acid residues.[125]Ferritin proteins have been used to produce hybrid biomaterials, such as a crystalline structure produced from classic ferritins (another subfamily of ferritins composed by 24 subunits) cross-linked by glutaraldehyde [126] or Au nanoparticles;[127]coordination polymers formed by modified Dps with chelating agents at the external surface and Fe(II) as the coordinated metal ion;[128] and filaments composed by classic ferritins and poly-(α, l-lysine).[74]Additionally, ferritin nanocages have been recently used to construct 2D superlattices mediated by disulfide bonds, or protein-based metal–organic frameworks.[129][130] In the first case, polymerization was triggered by oxidation reactions, whereas in the second study protein molecules were bound via metal coordination.

The design of novel supramolecular nanoarchitectures using proteins as building units is useful to expand and modulate their intrinsic natural activities,[124] in the case of proteins from the ferritin family, fast oxidation of free toxic ferrous iron and its storage in a ferric mineral form (biomineralisation) or cellular detoxification.[52] It is also important to mention that, in addition to Fe, ferritin proteins can also take up other heavy metals, such as Pb,[70] Cu [131] or Cd.[65] Protein-based functional supramolecular polymers may also be used as efficient factories for the synthesis of various monodispersed quantum dots, which can be applied in photocatalysis,[132] or several biotechnology applications,[43] as an alternative to synthetic organic or inorganic scaffolds/cages.[133]

In the present work, we demonstrate for the first time, that mini-ferritin protein can be used as building blocks to produce long linear supramolecular protein chains, longer than 1 μm, that may have interesting properties for development in emerging nanotechnology applications. Protein polymerization was achieved by cross-linking DpsT10C molecules with two different linkers using click chemistry, resulting in two distinct supramolecular protein structures. Both types of protein-polymer structures exhibited the natural Dps iron uptake activity. These supramolecular protein-polymers can be used as porous material with enhanced reactivity and selectivity, for different applications, as bioremediation or catalysis.[134][135]

Dps protein from *Marinobacter (M.) hydrocarbonoclasticus*[33] was first genetically modified to introduce a single cysteine residue at position 10 (here named DpsT10C), an exposed and flexible region at the N-terminus of each 12 subunits. This residue was used to selectively click ligate an alkene linker to form a thioether bond (thiol–ene coupling). In this reaction, a thiyl radical was generated by the photoinitiator lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP) upon UV light irradiation.[136][137] The radical was then added to the alkene double bond, creating a new thioether bond in the anti-Markovnikov product and a carbon radical, which abstracted a hydrogen atom from another sulfhydryl group and completed the radical cycle. Two synthetic approaches were used to functionalize Dps protein molecules and generate Dps polymers. In the first, the DpsT10C variant was S-allylated with 3-bromoprop-1-ene in a moderate basic buffer and subsequently coupled to DpsT10C by a thiol–ene coupling cross-linking two protein molecules through a $-(CH_2)_3-$ group. In the second, the DpsT10C variant was directly incubated with 1,2-bis(allyloxy)ethane and the photoinitiator, LAP, under UV irradiation, in a single step reaction, to form a $-(CH_2)_3O(CH_2)_2O(CH_2)_3-$ bridge between DpsT10C protein molecules. 3-Bromoprop-1-ene was previously used to derivatize cysteine and homocysteine residues of H3.3 histone short peptide constructs (residues 23–34),[138] while 1,2-bis(allyloxy) ethane was employed as staple between the two cysteine residues in the peptide YCKEACAL, using thiol–ene coupling.[135]

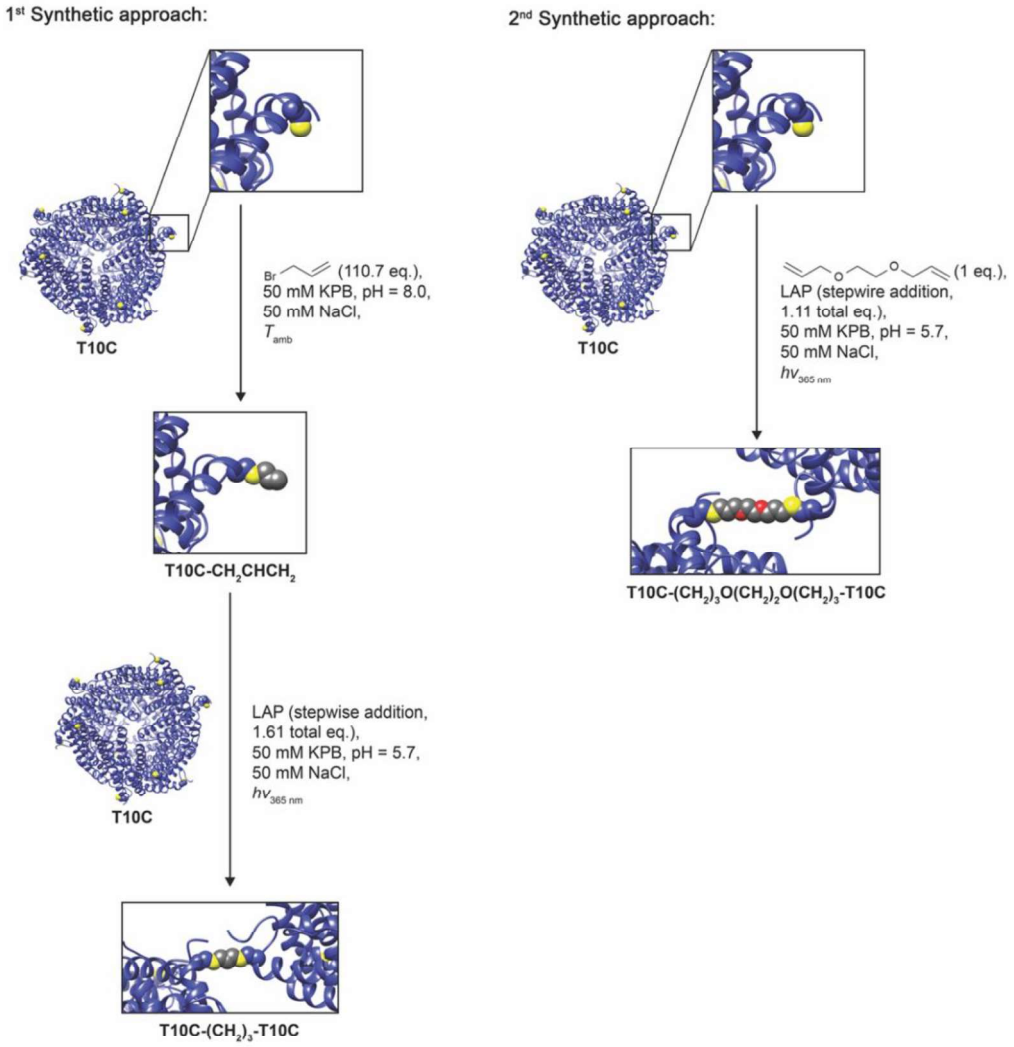

Figure 3-15: Schematics for the two synthetic approaches to generate supramolecular protein structures by coupling DpsT10C protein molecules with two different linkers. For illustration purposes, reactions are shown in only one of the 12 existing sulfhydryl groups in each protein molecule. Sulphur atoms of cysteine residues are shown in yellow, carbon atoms in grey and oxygen atoms in red (for simplicity hydrogen atoms are not shown). The DpsT10C structure was created using SWISS-MODEL server[139][140][141][142][143].

3.5.2 Results and discussion

S-allylation of DpsT10C with 3-bromoprop-1-ene aimed to introduce a double bond that allowed further thiol–ene coupling with the free sulfhydryl groups in DpsT10C protein molecules. Alternatively, DpsT10C protein molecules were coupled to each other, through a $–(CH_2)_3O(CH_2)_2O(CH_2)_3–$ linker in a one-step radical reaction with 1,2-bis(allyloxy)ethane. Both strategies are illustrated in ***Figure 3-15***.

Progress of S-allylation and thiol–ene coupling reactions was monitored by quantifying free sulfhydryl groups in the reaction mixture. The starting DpsT10C protein experimentally

accounted for 10.8 ± 1.4 sulfhydryls from a total of 12 groups present in the protein molecule, validating the experimental procedure to reduce and quantify –SH groups. In the first reaction, the S-allylation of DpsT10C with 3-bromoprop-1-ene, 84.2% of all –SH groups had reacted after 15 min, based on the quantification of 1.65 free sulfhydryl groups per DpsT10C protein molecule.

However, in the coupling reaction of T10C–CH_2CHCH_2 with DpsT10C protein to yield a T10C–$(CH_2)_3$–T10C cross-linked structure, the percentage of reacted sulfhydryl groups increased almost linearly over time up to 60 min (***Figure 3-16*** **–** black circles), reaching 41.9% of reacted groups after 120 min. The lower yield and longer reaction time can be due to steric hindrance of bulky protein units.

When DpsT10C proteins were reacted with the 1,2-bis(allyloxy) ethane in a single step to form T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C cross-linked structures, the percentage of sulfhydryl groups reached a plateau after 60 min, with 61.9% of reacted sulfhydryl groups at the end of the reaction (***Figure 3-16*** **–** white circles). This higher reactivity compared to the previous coupling reaction may be justified by the presence of a longer linker allowing a more stable conformation by reducing steric hindrance.

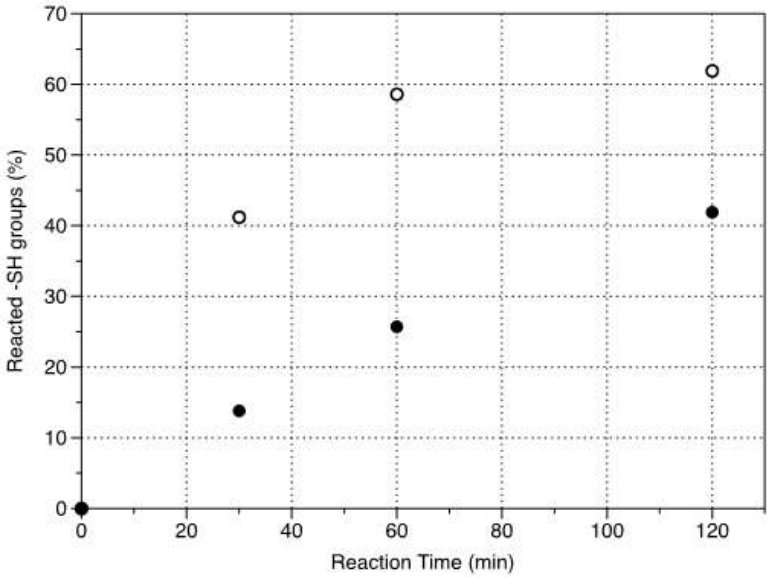

Figure 3-16: Progress of coupling reactions to form T10C-$(CH_2)_3$-T10C (black circles) and T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C (white circles) supramolecular structures, monitored by quantification of sulfhydryl groups.

In order to confirm S-allylation, coupling of DpsT10C protein molecules and formation of supramolecular protein structures, all products were analyzed by SDS-PAGE (***Figure 3-17***). In reducing denaturating conditions (lane labelled +ß in ***Figure 3-17*****A**), the DpsT10C multimer dissociates and a single protein monomer band, with an apparent molecular mass around 17 kDa,

was observed (band labelled M in ***Figure 3-17***). In the absence of β-mercaptoethanol (non-reducing conditions, lanes labelled −ß in Fig. 3), a residual dimer band with an apparent molecular mass of ~30 kDa was also visible (band labelled D in ***Figure 3-17***) and attributed to the formation of bonds between two monomers due to the oxidation of sulfhydryl groups or to thiol–ene coupling. For the DpsT10C protein (***Figure 3-17*****A**) the dimer band completely disappeared in reducing conditions (lane labelled +ß in ***Figure 3-17*****A**), indicating that disulfide bridges between monomers were reduced. The T10C–CH_2CHCH_2 protein showed a smaller amount of dimeric species (lane labelled −ß in ***Figure 3-17***B) probably due to the fact that fewer cysteine residues were now available to form disulfide bridges due to S-allylation. The dimeric species band was also lost upon the use of reducing conditions (lane labelled +ß in ***Figure 3-17*****B**). The results were strikingly different for the T10C–$(CH_2)_3$–T10C protein polymer that showed a stronger dimer band, that did not disappear when treated with β-mercaptoethanol (***Figure 3-17*****C**). In fact, ca. 60% of the dímer band intensity remained in reducing conditions corresponding to dimers that are bridged due to the successful coupling of DpsT10C protein with T10C–CH_2CHCH_2. The electrophoretic results for the coupling product T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C were similar to those described for the product T10C–$(CH_2)_3$–T10C, with formation of non-reducible dimers (data not shown).

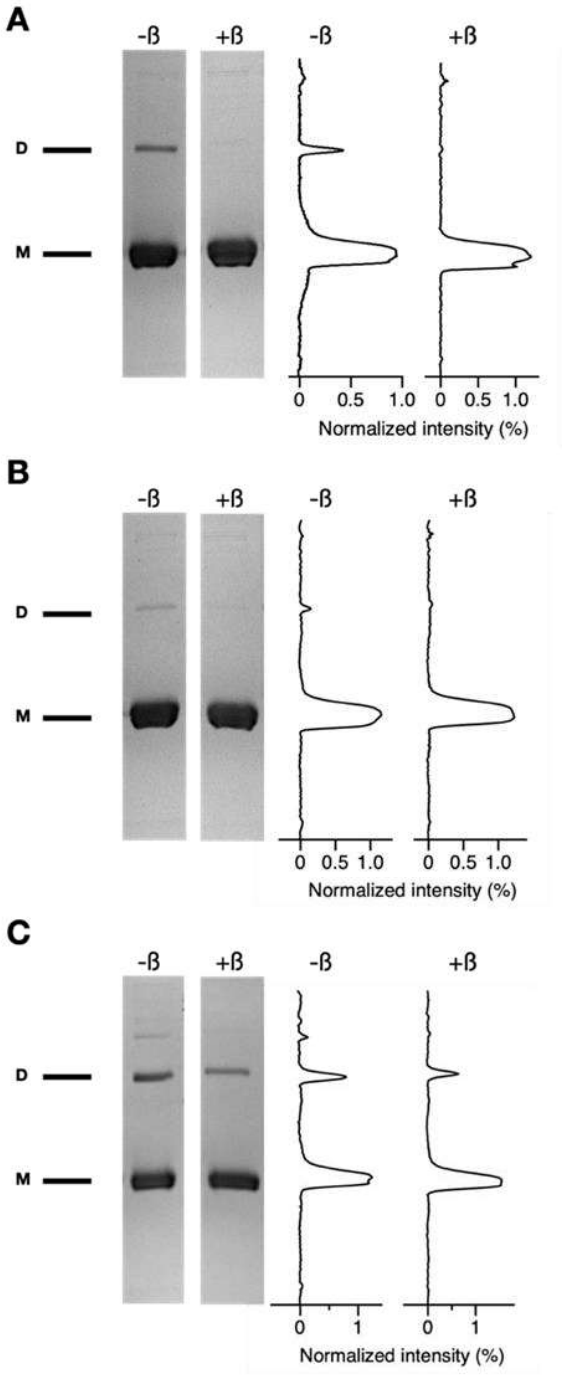

Figure 3-17: SDS-PAGE results for samples of DpsT10C (A), T10C-CH_2CHCH_2 (B) and T10C-$(CH_2)_3$-T10C (C) proteins. In each panel monomeric protein and dimeric protein bands (marked by letters M and D, left) and densitometric quantification (right) are shown. On top, -ß and +ß signals indicate if the result was obtained in the absence or presence of added β-mercaptoethanol.

Products of the S-allylation, T10C–CH_2CHCH_2, and coupling reaction, T10C–$(CH_2)_3$–T10C, were further characterized by MALDI-TOF MS to assess the allylation of sulfhydryl groups in DpsT10C and further linkage of two distinct protein molecules forming covalent dimers. As expected, the mass spectrum showed an increase by 40.07 Da factor due to the S-

allylation, bonding of the $CH_2CH=CH2$ group to the sulfhydryl group of the cysteine residue in both protein populations (see ESI†).

In order to assess the secondary structure and the thermal stability of protein polymers, a sample of T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C was analysed by circular dichroism spectroscopy and data compared with the DpsT10C protein (***Figure 3-18***). The spectra are identical in the acquired energy ranges and show negative peaks at 210 and 221 nm, as well as an intensive positive peak at 193 nm which are indicative of predominant α-helical secondary structure.[144] A more detailed analysis was carried out on data obtained by synchrotron radiation circular dichroism (SRCD) spectroscopy using the DichroWeb[98][145] server, using CDSSTR method and SP175 reference data set (optimized to 190–240 nm).[146][147] SRCD data is particularly valuable because it contains a higher information content with the possibility for data collection at high energies (e.g. in the range of 175 to 195 nm) as well as the high signal-to-noise levels obtained as compared to laboratory-based, benchtop circular dichroism (CD) spectrometers. DichroWeb analysis showed that the T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C polymer retains the secondary structure of the DpsT10C precursor protein with a modelled 80% of α-helices as compared to 76% as predicted by DSSP method using the 2struc server.[148] CD and SRCD data obtained during thermal denaturation studies showed a difference in denaturation profiles (***Figure 3-19***). Using measurements at 221 nm it was possible to obtain Tm values of 61.1 ± 0.5 °C and 53.3 ± 0.5 °C for DpsT10C and T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C polymer. These values emphasize the destabilization of the latter as related to the former. Another noticeable difference are the values obtained for ΔHm, which is almost reduced to half in the T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C polymer (58.2 ± 0.6 kcal mol^{-1} vs. 101.4 ± 0.4 kcal mol^{-1} obtained for DpsT10C). A possible interpretation for this difference is the expected collaborative effect on the disruption of hydrogen bonding, and other non-covalent molecular interactions,[149] in addition to a more favourable hydration of the denatured state for the polymeric T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C protein.

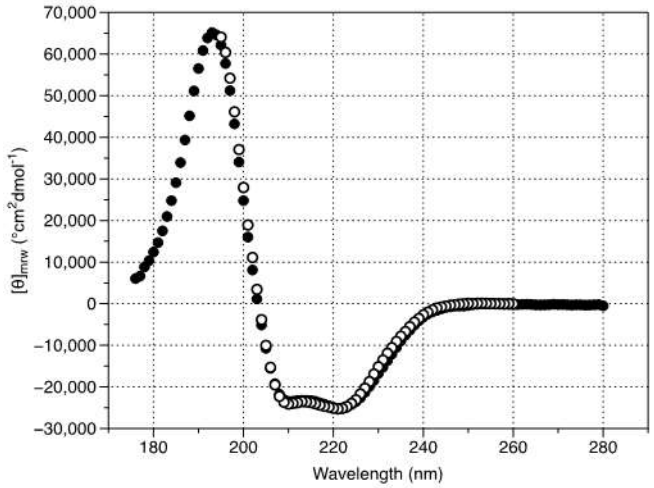

Figure 3-18: Circular Dichroism spectra obtained at 25 °C for DpsT10C protein (white circles, CD data) and T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C polymer (black circles, SRCD).

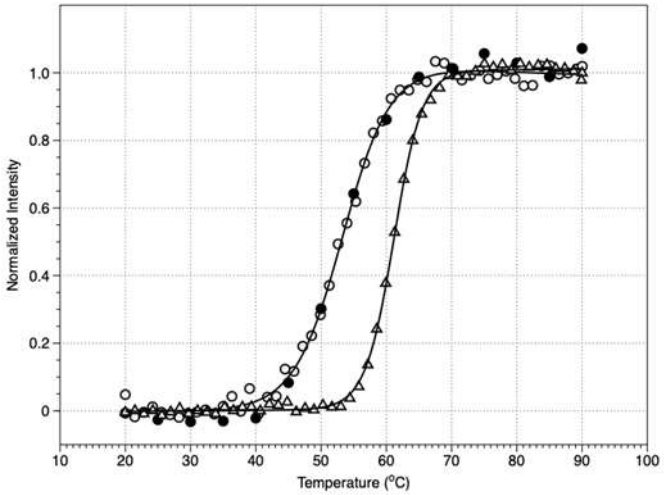

Figure 3-19: Thermal denaturation studies of DpsT10C protein (white triangles, CD data) and T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C polymer (white circles, CD data, and black circles, SRCD data). Thermal denaturation curves are derived from signal intensity at 221 nm versus temperature with proper baseline subtraction. Solid lines are the result of curve fitting procedure described in SI.

To measure the size distribution of the formed structures, the T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C protein polymer was analysed by dynamic light scattering (DLS) and compared with results obtained for the DpsT10C protein (***Table 3-6***). DpsT10C had a hydrodynamic diameter of 8.61 nm with a SD of 0.80 nm, which is in agreement with the ca. 9 nm diameter reported in the literature for these proteins.[150][151] The hydrodynamic diameter of T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C polymer shifted to larger sizes, ca. 35 nm, with a polydispersity index higher than the one observed for DpsT10C samples, indicating a heterogeneous size distribution of structures (see histogram representation of particle size distribution on ESI†).

Table 3-6: Hydrodynamic diameter and polydispersity indexes (PDI) of T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C and DpsT10C, determined by DLS

Sample	Hydrodynamic diameter ($\bar{x}$(SD), nm)	PDI ($\bar{x}$ (SD))
DpsT10C	8.61 (0.80)	0.353 (0.19)
T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C	34.60 (4.54)	0.614 (0.12)

Atomic force microscopy (AFM) images in liquid media were collected to analyze the morphology of T10C–$(CH_2)_3$–T10C and –$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C supramolecular structures and compared with DpsT10C (***Figure 3-20***). DpsT10C protein molecules appeared as spherical particles (upper images in ***Figure 3-20***). Cross-linking protein molecules with both linkers (–$(CH_2)_3$– and –$(CH_2)_3O(CH_2)_2O(CH_2)_3$–) resulted in the formation of long protein chains with globular DpsT10C molecules as tandem repeats. In the case of the T10C–$(CH_2)_3$–T10C protein structure, worm-like chains seemed shorter (ca. 0.5 μm) and wider than T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C ones that could reach micrometers in length. A repetitive pattern has also been observed in other high molecular weight engineered protein structures, namely the titin-mimetic protein macromonomers, produced by supramolecular polymerization linked by disulfide bonds.[152] These polymeric structures presented an average molecular weight of around 0.5 MDa, the largest protein structure so far developed. Another example is the long protein nanowires (up to 128 nm in length and a height of 4.8 nm) of glutathione S-transferase dimers linked by cucurbit[8]uril and a tripeptide FGG–tag that exhibited antioxidant activity.[153]

In our study the longest observed T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C protein polymer was composed of protein units (DpsT10C oligomers) of about 15 kDa nm^{-1}, a high-density protein structure, which accounted for up to 15 MDa in total molecular mass. The diameter, ca. 50–55 nm, of the long structures seems to indicate that chains can further interact with each other and self-assemble into large tridimensional structures. ***Figure 3-20*** also shows that due to their high concentration, both types of protein-based chains appeared as bundles, forming a widespread network.

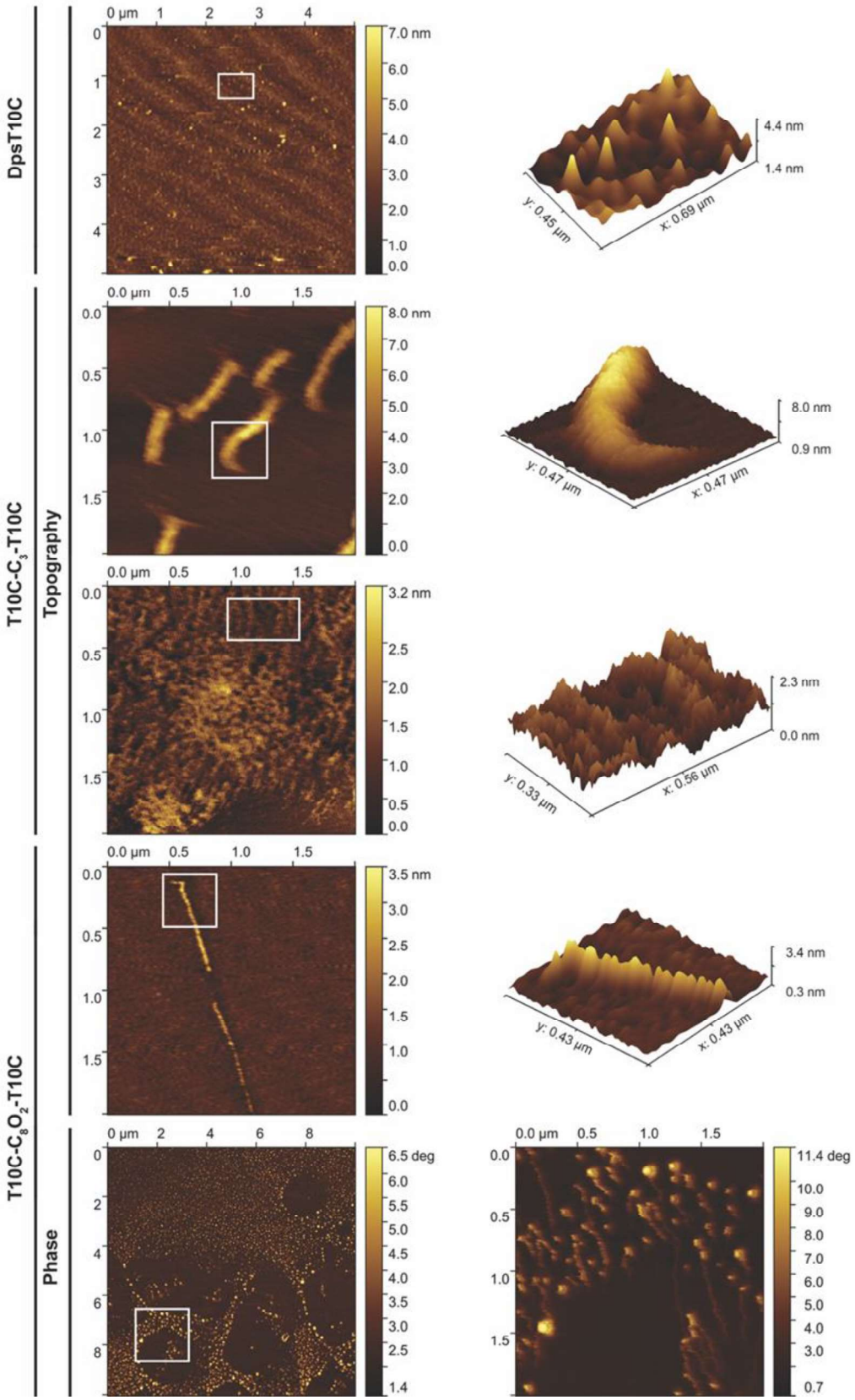

Figure 3-20: AFM topographic and phase images in liquid medium of DpsT10C protein and supramolecular structures, products of the coupling reactions, T10C-$(CH_2)_3$-T10C and T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C supramolecular structures. The images presented on the right correspond to three-dimensional representations and close-ups of the rectangular sections marked on the left images.

To confirm the high molecular mass of T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C, a sample was applied into a size exclusion chromatography (SEC) column (see ESI†). A major band was eluted with the void volume of the column (above 600 kDa), consistent with the expected high molecular mass of the long protein polymers.

The iron uptake enzymatic activity of T10C–$(CH_2)_3$–T10C and T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C protein structures was compared with DpsT10C, reacting the proteins with ferrous ions in the presence of an excess of H_2O_2, in 50 mM MOPS buffer, pH = 7.0, 50 mM NaCl as previously described.[154] The increase of absorbance at 350 nm (ΔAbs 350 nm), characteristic of oxo/hydroxo-ferric species, was plotted as a function of the molar ratio Fe(II)/dodecameric proteins in solution (Fe/protein) as presented in ***Figure 3-21***.

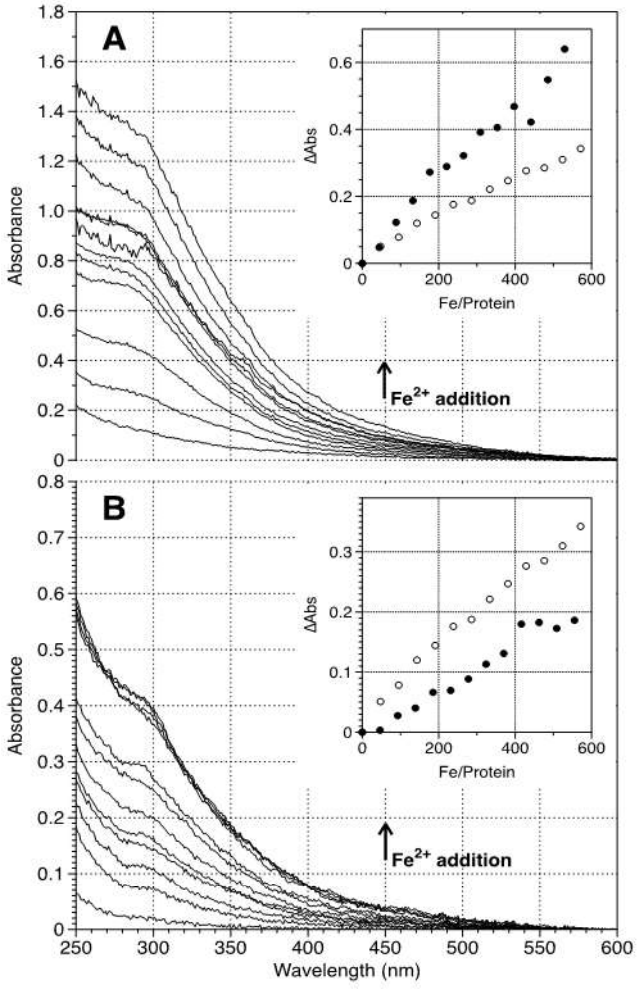

Figure 3-21: Iron uptake by T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C (A) and T10C-$(CH_2)_3$-T10C (B) protein structures in the presence of H_2O_2. Each spectrum represents an addition of 48 Fe(II)/protein. The insets show the increase in absorbance at 350 nm as a function of total Fe(II)/protein ratio. Black circles represent the increase for coupled products and open circles the corresponding increase for the DpsT10C protein.

The iron uptake assays revealed that both types of supramolecular structures, products of the coupling reactions, were enzymatically active, since the absorbance at 350 nm increased with increasing amounts of added iron, indicating the uptake and oxidation of ferrous ions in solutions (ferroxidation reaction).[154][155] Since no precipitation was observed during the assay, one can conclude that the oxidized iron was stored inside the protein nanocavity, in the form of a mineral (mineralisation reaction).[154][155] Comparing the iron storage capacity of both samples, it seems that while T10C–$(CH_2)_3$–T10C was able to store up to ~375 Fe per protein molecule in the supramolecular structure, less than free DpsT10C protein (white circles on ***Figure 3-21*B**), the product T10C–$(CH_2)_3O(CH_2)_2O(CH_2)_3$–T10C was able to oxidize and accommodate more than 530 Fe per DpsT10C molecule in the supramolecular structure, showing a quasi-linear dependence with added iron. This result can be attributed to a higher catalytic efficiency of this supramolecular protein structure (higher ferroxidation and/or mineralisation rate) or formation of different ferric species, with distinct molar extinction coefficients. Although densely packed, the polymeric structures are still porous, with different channels in each Dps protein building block, connecting the inner cavity to the outside. In both types of structures, we observed associations of the long protein polymers in complexes that show empty cavities between the chains (***Figure 3-20***). The substrates, Fe2+ ions and H_2O_2/O_2, can diffuse through these channels and cavities and bind to amino acids and be translocated to the active sites where the oxidation occurs. Once oxidized the ferric species are translocated to the inner cavity for storage in the mineral form.

3.5.3 Conclusion

To our knowledge, we report, for the first time, the development of enzymatically active high molecular weight functional polymeric structures with Dps protein used as a building block. In fact, these structures were able to oxidize ferrous iron and incorporate the ferric form inside Dps protein molecules. Such interpretation is supported by the fact that polymer synthesis did not disrupt the protein secondary structure, thus enabling the proper assembly for the native dodecameric nanocage structure. These protein-based linear polymers were synthesized through thiol–ene coupling of Dps molecules, a prokaryotic protein nanocage belonging to the ferritin family. The formation of well-ordered long polymeric structures formed by covalently bound globular proteins is challenging and limited.[75] The high degree of protein polymerization is observed in nature and is important for their physical and mechanical properties, as toughness, elasticity and strength.

The length of the synthesized protein polymers was dependent on the type of linker used on the coupling reaction, suggesting that a longer linker allowed the formation of longer protein polymers.

This work paves the way for the possibility to obtain materials that may be used for bioremediation or catalysis, or as a starting point to further explore the synthesis of protein based biomaterials with improved mechanical properties.

3.5.6 Supplementary Information

3.5.6.1 Materials and methods

All reagents and solvents commercially available were of pro-analysis grade (except when specified) and were used as acquired without further purification. Lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP) was synthesized according to Fairbanks et al.[156] and 1,2-bis(allyloxy)ethane was synthesized based on the procedure described by Abramov and collaborators.[157] Ultrapure water was used in all experiments. The solvents and buffers were deoxygenated with four 5 min-cycles of vacuum/argon in a Schlenk line.

In the DpsT10C allylation reaction, deoxygenated 50 mM potassium phosphate buffer (KPB), pH = 8.0, with 50 mM NaCl was used as solvent, while for the coupling reactions the buffer was deoxygenated 50 mM KPB, pH = 5.7, containing 50 mM NaCl.

UV irradiation of coupling reaction mixtures, under stirring at room temperature, was performed through a clear side of a 1.5 mL quartz cuvette with 1 cm optical pathlength (Hellma), with a 365 nm light using a bandpass filter and a 200 W Xe/Hg arc-lamp (Newport Oriel, model 66902, with a power supply Newport Oriel, model 69907, and a 365 nm filter Semrock, model MaxLamp^{TM}, with an outer diameter of 25 mm and transmission between 360 nm and 372 nm >93%).

3.5.6.2 Sulfhydryl groups quantification

Calibration curves and samples for sulfhydryl groups determination during S-allylation and coupling reactions were prepared according to the protocol described by C. K. Riener et al.[158] with modifications as stated below. For the calibration curves, a 1 mM L-cysteine stock solution, 0.01% HCl in water was freshly prepared prior to usage. Standard solutions with final concentrations between 0 μM and 30 μM in 50 mM KPB, pH = 8.0, 50 mM NaCl with 0%, 0.1% or 4.8% v/v dimethylformamide (depending on its content in the reaction mixture) were prepared by dilution of the stock solution, in order to construct a 9-point calibration curve. The stock solution was used immediately after preparation.

Standards and samples were prepared in 50 mM KPB, pH = 8.0, 50 mM NaCl with 0%, 0.1% or 4.8% v/v DMF, taking 30 μL of the solution to be quantified, unless otherwise stated. All standards and samples were prepared in a final volume of 980 μL. 20 μL of Ellman's reagent (7.8

mM 5,5′-dithio-bis(2-nitrobenzoic acid), 0.2 mM EDTA, 100 mM NaH2PO4, pH = 7.0) were added and homogenized by inversion. After 5 min sample incubation at room temperature, absorbance at 412 nm was measured against water. All samples were quantified in triplicate, except in the coupling reaction with 1,2-bis(allyloxy)ethane where duplicates were assayed.

3.5.6.3 Production of DpsT10C variant

The DpsT10C variant was produced by site-directed mutagenesis using the NZYMutagenesis kit (NZYTech) following the manufacturer's protocol. Oligonucleotide primers, forward (5'–g aac ttt ata ggt ctc gac tgc gac aaa acc cag aag ctg g–3') and reverse (5'–c cag ctt ctg ggt ttt gtc gca gtc gag acc tat aaa gtt c– 3') were designed and analyzed for optimization with the software Primer Premier 6.24 from PREMIER Biosoft, synthesized and purified by NZYTech. Primers annealing and extension were performed at 60 ºC and 68 ºC, respectively. PCR product was digested with DpnI to eliminate parental plasmid DNA and used to transform ultracompetent *Escherichia (E.) coli* cells, NZYStar (NZYTech). Selected clones were analyzed by DNA sequencing of plasmid DNA (Eurofins Genomics). Production of the DpsT10C protein in *E. coli* BL21(DE3) cells in LB medium containing 100 μg/mL of ampicillin was done as described before for the wild type protein,[33] inducing protein expression with 1 mM of isopropyl-β-D-thiogalactoside. After pelleting, cells were resuspended in 10 mM Tris-HCl buffer pH = 7.6, containing 1 mM dithiothreitol (DTT) and lysed by sonication. The soluble protein fraction, collected after ultracentrifugation, was dialyzed overnight at 4 ºC against the same DTT-containing buffer and applied into a DEAE-Sepharose Fast Flow column (2.6 × 30 cm, GE Healthcare) according with D. Penas, et al.[33]

3.5.6.3 Production of supramolecular Dps polymers

The DpsT10C protein was incubated with 10 mM DTT at room temperature for 2 h. DTT was then removed by ultrafiltration on a Vivaspin 6 with 100 kDa MWCO (Sartorius Stedim Biotech) and the buffer exchanged to the appropriate one, depending on the reaction. Before S-allylation, the protein was quantified by UV-Vis spectroscopy using the molar extinction coefficient at 280 nm (293 $mM^{-1}cm^{-1}$).

3.5.6.4 DpsT10C S-allylation with 3-bromoprop-1-ene

The DpsT10C protein at 15.5 μM (186.1 μM in sulfhydryl groups), in 50 mM KPB, pH = 8.0, 50 mM NaCl, was incubated with 20.6 mM 3-bromoprop-1-ene (97%, Sigma-Aldrich, 56.7 μL of a 1.67 M solution in DMF, Scharlau) in a molar excess of 110.7 eq., at room temperature, under argon atmosphere with magnetic stirring, in a total volume of 4589 μL. Consumption of free sulfhydryl groups was monitored at 0 min (just before the addition of 3-bromoprop-1-ene linker reagent) and after 15 min, by taking 30 μL aliquots of the reaction mixture. After 20 min, the reaction was stopped by two consecutive dialysis in a 12-14 kDa MWCO membrane tube

(Medicell Membranes Ltd) against 1 L of 50 mM KPB, pH = 8.0, 50 mM NaCl, at 4 ºC. This protein solution, termed T10C-CH2CHCH2, was centrifuged at 10,000 g at 8 ºC for 15 min and quantified by UV-Vis spectroscopy assuming the molar extinction coefficient at 280 nm as for the DpsT10C protein.

3.5.6.5 Coupling of the DpsT10C protein with T10C- CH_2CHCH_2 (production of T10C-$(CH_2)_3$-T10C)

Immediately prior to the coupling reaction, the buffer of the previous T10C-CH_2CHCH_2 mixture was exchanged to 50 mM KPB, pH = 5.7, 50 mM NaCl using a Vivaspin 6 with 100 kDa MWCO.

Reactive groups (reduced sulfhydryl groups in DpsT10C and S-allylated groups in T10C-CH_2CHCH_2) were at 101.3 µM, in a total volume of 219.7 µL (assuming 100% yield in the allylation reaction). The reaction mixture was prepared in a 1 mL quartz cuvette by transferring DpsT10C, T10C-CH_2CHCH_2 and 16.8 µL of a freshly prepared solution of the radical photoinitiator LAP (5.30 mM in 50 mM KPB, pH = 5.7, 50 mM NaCl). The reaction mixture was irradiated with UV light (365 nm), at room temperature, for 120 min. The second and third aliquots of LAP solution (16.8 µL) were added at 30 min and 60 min, respectively. After 120 min, the reaction was stopped by freezing at −20 ºC. The decay of the number of sulfhydryl groups was monitored taking 15 µL aliquots of the reaction mixture at 30 min, 60 min, and 120 min (before LAP addition, when applicable). For simplicity, the product of the coupling reaction was termed T10C-$(CH_2)_3$-T10C.

In order to perform the enzymatic activity assays and MALDI-TOF mass spectroscopy analysis, the protein buffer was exchanged to specific solvents, through two consecutive dialysis in a Slide-A-Lyzer MINI dialysis unit with 3.5 kDa MWCO (Thermo ScientificTM) against 1 L of the solvent, at 4 ºC.

3.5.6.6 Coupling reaction of the DpsT10C protein with 1,2-bis(allyloxy)ethane (production of T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C)

In this reaction, DpsT10C protein (160.5 µL of 277.3 µM protein solution in 50 mM KPB, pH = 5.7, 50 mM NaCl) reacted at room temperature with the linker reagent 1,2-bis(allyloxy)ethane (1.9 µL of a 0.144 M solution prepared in 10% v/v DMF, 1 eq.) in a total volume of 162.4 µL, with a sulfhydryls concentration of 3.29 mM. Protein coupling was performed as described before, except for the addition of the photoinitiator. 8.7 µL of LAP solution (5.10 mM) was added at 0, 5, 10, 35, 40, 45 and 65, 70 and 75 min of reaction incubation time. Sulfhydryl groups were quantified by taking 10 µL aliquots at 30, 60 and 120 min.

Over time, the coupling reaction product, here termed T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C, settled on the bottom of the liquid solution forming a sediment. The sediment was collected by centrifugation at 10,000 g, at room temperature for 15 min and resuspension was only possible with ultrapure water. The supernatant was free of protein, as assessed by SDS-PAGE.

3.5.6.7 Electrophoretic analysis

Products of both S-allylation and coupling reactions (T10C-$(CH_2)_3$-T10C and T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C), as well as DpsT10C protein, were analyzed by SDS-PAGE (12.5% or 15% polyacrylamide) as described by Laemmli,[159] by taking 20 pmol of dodecameric protein, in the presence or absence of β-mercaptoethanol. The NZYBlue (NZYTech) was used as protein marker. Gels were stained with BlueSafe solution (NZYTech) according to the manufacturer.

3.5.6.8 Circular Dichroism and Synchrotron Radiation Circular Dichroism

Synchrotron Radiation Circular Dichroism (SRCD) spectra of T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C were recorded on the AU-CD beam line at the ASTRID2 synchrotron radiation source (Aarhus, Denmark). The obtained spectra were recorded in triplicates with 1 nm steps and a dwell time of 2.1 s per step, using a 0.01008 cm pathlength quartz cells (SUPRASIL, Hellma GmbH, Germany), for the wavelength range of 170–280 nm. The T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C sample was at 0.68 mg/mL. Water spectra were used as baseline and subtracted from the T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C solution spectra collected at 25 ºC. The temperature ramp was acquired from 5 ºC to 90 ºC, at 5 ºC steps in the ranges 5-40 ºC and 65-90 ºC and in 2.5 ºC steps between 40 ºC and 65 ºC, measuring the full spectrum at each temperature in triplicate with a dwell time of 2.1 s per wavelength step. Before and after the temperature ramp, a spectrum at 25 ºC was obtained.

CD spectra were acquired at the BioLab, Biological and Chemical Analysis Facility, UCIBIO-LAQV/Requimte, FCT NOVA, Portugal, using an AppliedPhotophysics Chirascan™ qCD. The obtained spectra were recorded in triplicate with 1 nm steps and a dwell time of 2 s per step using 0.02 cm quartz cells (SUPRASIL, Hellma GmbH, Germany), for the wavelength range of 195–260 nm. The DpsT10C protein was at 0.90 mg/mL in 50 mM KPB, pH = 5.7, 50 mM NaCl, while the T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C solution sample was at 0.34 mg/mL. Spectra were corrected with the buffer, or water in the case of T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C, as mentioned before. Temperature ramps were acquired from 20 ºC to 90 ºC at 1 ºC/min and with a dwell time of 0.4 s. Before and after the temperature ramp, a spectrum at 20 ºC was obtained.

Melting temperature (Tm) was calculated by fitting the experimental normalized ellipticity data at 221 nm as a function of temperature (K) with the equation 3.1:

$$SI = \frac{(\alpha_N + \beta_N \times T) + (\alpha_D + \beta_D \times T) e^{-\frac{\Delta H_m \left(1 - \frac{T}{Tm}\right)}{RT}}}{1 + e^{-\frac{\Delta H_m \left(1 - \frac{T}{Tm}\right)}{RT}}} \quad (3.1)$$

In which SI is the normalized ellipticity, ΔHm is the enthalpy at the unfolding transition, Tm is the melting temperature, T is temperature in Kelvin, R is the universal gas constant, and α_N, β_n, α_D, and β_D are the baseline slope and intercept for pre-transition and post-transition baseline, respectively.[160]

3.5.6.9 Dynamic light scattering

The hydrodynamic radius, R_H, of both DpsT10C in 10 mM Tris-HCl, pH = 7.6, 500 mM NaCl and T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C solution, at 2.5 μM, was determined using dynamic light scattering with a SZ-100-Z nanoparticle analyzer, HORIBA Scientific. The hydrodynamic diameter and polydispersity index (PDI) were obtained as the mean (standard deviations, SD, were also determined) of triplicate measurements. Data were collected at 25 °C, with a measurement angle of 90°, for 5 min with 10 s between acquisitions.

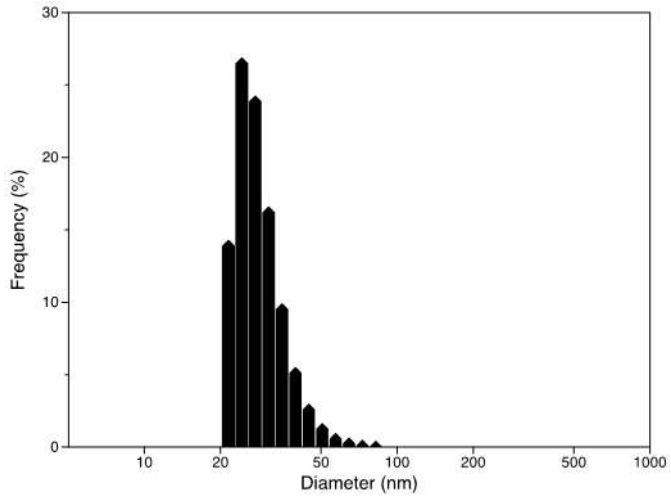

Figure 3-22: Particle size distribution of T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C.

The histogram representation of particle size distribution obtained for the T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C sample showing that structures are polydispersed, with particles diameters ranging from around 20 to 80 nm, ***Figure 3-22***.

3.5.6.10 Atomic force microscopy

Protein samples of DpsT10C (in 50 mM MOPS, pH = 7.0, 200 mM NaCl, 1 mM DTT), T10C-$(CH_2)_3$-T10C and T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C (both in 50 mM KPB, pH = 5.7, 50 mM NaCl) at 100 nM were applied on a freshly cleaved grade V1 muscovite mica (Ted Pella, Inc). The measurements were performed in liquid medium in an Asylum Research MFP-3DTM Standalone (Oxford Instruments) operated in alternate contact mode with commercially available silicon cantilevers (Olympus AC240TS, Olympus Corporation, f0 = 70 kHz; k = 2 N/m) at room temperature and atmospheric pressure. Images were subjected to a first-order flattening and were processed with Gwyddion modular program.[161]

3.5.6.11 Size exclusion chromatography

The resulting T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C protein polymers were fractionated by size exclusion chromatography (SEC) using a Superdex® 200 10/300 GL prepacked gel filtration column with a separation range between 10 and 600 kDa (Cytiva, formerly GE Healthcare Life Sciences). Blue Dextran (~2000 kDa) was mixed with the DpsT10C sample to determine the void volume of the column and eluted at 8 mL. The chromatogram shows that a major fraction of the protein polymers eluted at the same volume as Blue Dextran, above the separation range of the column, proving that its molecular mass is above 600 kDa, in accordance with the AFM data. The high molecular mass band presents some shoulders on the right side, indicating size heterogeneity of the polymers, in agreement with DLS data, ***Figure 3-23*****.**

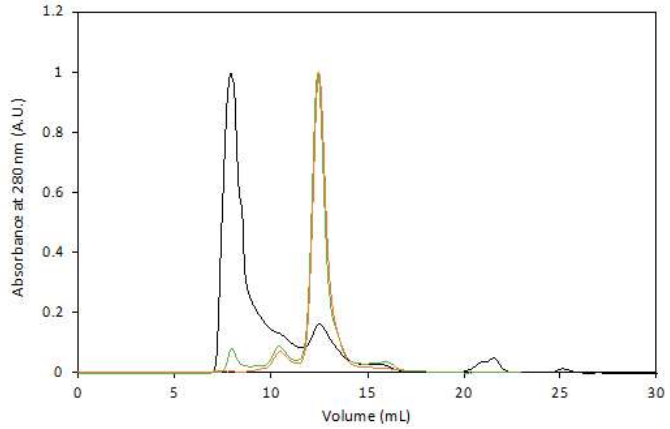

Figure 3-23: SEC profile of T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C protein structures (black line) in 50 mM KPB, pH = 5.7, containing 50 mM NaCl, at a flow rate of 0.5 mL/min. For comparison, DpsT10C with (green line) and without Blue Dextran (orange line) were applied in the same conditions.

3.5.6.12 Enzymatic activity

The iron uptake activity of DpsT10C, T10C-CH_2CHCH_2 and supramolecular protein polymers (T10C-$(CH_2)_3$-T10C and T10C-$(CH_2)_3O(CH_2)_2O(CH_2)_3$-T10C) was measured spectrophotometrically, monitoring the absorbance at 350 nm, adapting the protocol described by

D. Penas, et al.[33] Briefly, 48 Fe(II)/protein (molar ratio) were sequentially added to 1 μM of protein in 50 mM MOPS, pH = 7.0 containing 50 mM NaCl, in the presence of an excess of hydrogen peroxide, under magnetic stirring.

3.5.6.13 MALDI-TOF mass spectroscopy

Molecular masses of DpsT10C, T10C-CH_2CHCH_2 and T10C-$(CH_2)_3$-T10C were determined by matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry (MALDI-TOF MS) by the Laboratory for Biological Mass Spectrometry – Isabel Moura services, LAQV/Requimte, FCT NOVA, on a Ultraflex II MALDI-TOF TOF Bruker-Daltonics equipped with a LIFT cell and N2 laser. Briefly, 0.5 μL of 50 μM protein solution in water containing 0.01% TFA was hand-spotted onto a MALDI target plate previously treated with sinapinic acid. MS spectra were acquired in positive linear mode in the m/z ranges of 5-20 000 m/z and 20 000-130 000 m/z. A total of 1000 spectra were acquired at each spot position at a laser frequency of 50 Hz. Protein Calibration Standard I from Bruker was used for external calibration. ***Figure 3-24*** shows some of the typical results obtained for one of two replicates. For DpsT10C (Panel A) two prominent features averaging at 17 855 m/z and 17 296 m/z are discernible, which is in good agreement with the expect 17 858 m/z and 17 298 m/z for protein monomers with 156 and 151 amino acid residues. After the S-allylation reaction (Panel B), these features appear at 17 895 m/z and 17 334 m/z which is what one would expect for a successful reaction since the adduct should add 40.07 m/z to the total mass of each monomer. Finally, after the coupling reaction (Panel C) it was possible to find a feature averaging at 35 185 m/z. This can be interpreted as the mass of two monomers coupled together by the -$CH_2CH_2CH_2$ linker, for which a feature averaging 35 191 m/z is expected.

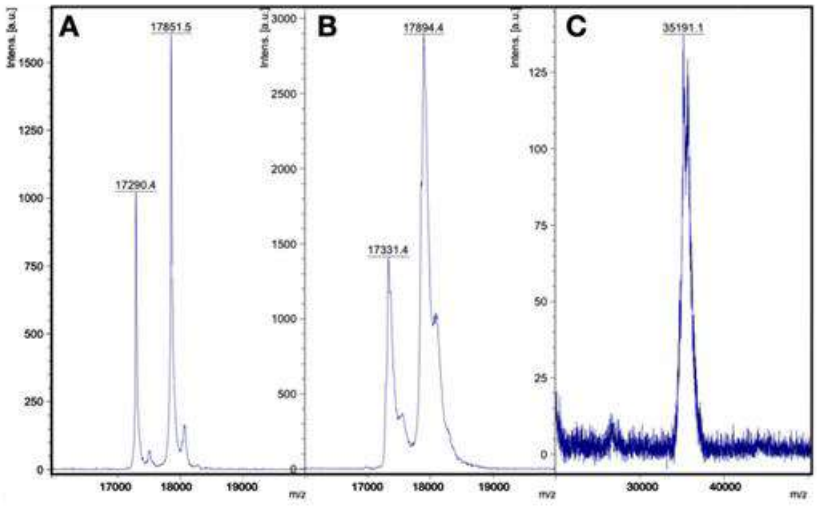

Figure 3-24: MALDI-TOF MS spectra of: A) DpsT10C; B) T10C-CH_2CHCH_2 after S-allylation, and C) T10C-$(CH_2)_3$-T10C, after coupling reaction.

4

DNA binding activity of Dps

4.1 Overview

One of the main roles of Dps is to protect DNA from external and internal agents being involved in the cellular response to multiple stresses such as oxidative, thermal, UV and extreme pH. As previous stated Dps can interact with DNA (supercoiled or linear, plasmid or genomic DNA) without sequence specificity, protecting it from chemical and physical damage

This chapter includes the manuscript "Dps-DNA interaction in *Marinobacter hydrocarbonoclasticus* protein. Effect of a single charge alteration", wherein CD and SRCD spectroscopies were applied to study how DNA (pUC19) binding would alter the secondary structure of Dps WT. The role of the N-terminal extension of the protein on DNA binding activity and protection was also investigated. Apo and holo forms of WT Dps and variant (Q14E), where a single negative charge was placed on the N-terminal region (first 15 residues), were incubated with supercoiled pUC19, a plasmid, and complex formation was analysed by Electrophoretic Mobility Shift Assays (EMSAs).

The secondary structure of WT Dps, apo and holo forms and Q14E variant proteins was assessed by Synchrotron Radiation Circular Dichroism (SRCD), in the absence or presence of pUC19 plasmid DNA.

AFM was another technique employed in order to understand the dependence of DNA binding activity of Dps with specific DNA strands. With this approach DNA origami structures were used as scaffolds for DNA strand specific positions and sequences and afterwards incubated with Dps WT under different buffer conditions.

4.2 Experimental procedures

4.2.1 DNA origami production

DNA origami assembly was performed by adding 5 μL of 100 nM viral scaffold strand M13mp18 to 202 short ssDNA 6 μL each in10 μL 10 x TAE (400 mM Tris, 10 mM EDTA, 200 mM acetic acid pH 8.3), 125 mM $MgCl_2$ and 46 μL deionized water. 3 μL of modified strands of interest (Appendix B1.2) at 100 nM are added to the mixture. The annealing process of the DNA structures is achieved by heating up to 80 ºC and cooling down stepwise over two hours to 6 °C in Primus 25 advanced ® Thermocycler. The final solution is filtered three times with 1 x TAE buffer (40 mM Tris-acetate buffer, pH 8.0, 0.1 mM EDTA) with 15 mM $MgCl_2$ by centrifugation for 6 min at 4,629 x g.

4.2.2 Binding of Dps WT to DNA origami

DNA origami was incubated with Dps WT at different pH values, between 4.0 and 9.0, adjusted with NaOH or HCl. Two different Dps buffer concentrations were used, 200 mM MOPS,

with 200 mM NaCl and 50 mM MOPS containing 50 mM NaCl. DNA origami and Dps WT were incubated in a 1:10 ration at 30 ºC, for 1 h with orbital shaking at 300 rpm.

4.2.3 Microscopy techniques – AFM

The assembled structures are immobilized in 10 x 10 mm^2 silicon substrates previously cleaned by air plasma. The sample is washed with 1 mL of 1:1 deionized water/ethanol solution and dried for 1 h in absolute ethanol to get rid of the remaining water.

The samples were either imaged on a Nanosurf Flex AFM or Bruker MultiMode 8, using a Tap 150 cantilever or a ScanAsyst-Air silicon tip, respectively, in tapping mode and analysed using the software Gwyddion 2.39.

4.3 Results and discussion

In order to access the structural integrity of the DNA origami triangles these were analysed by AFM in different buffer conditions, in the presence or absence of Dps WT, (***Figure 4-1*** and ***Figure 4-2***.)

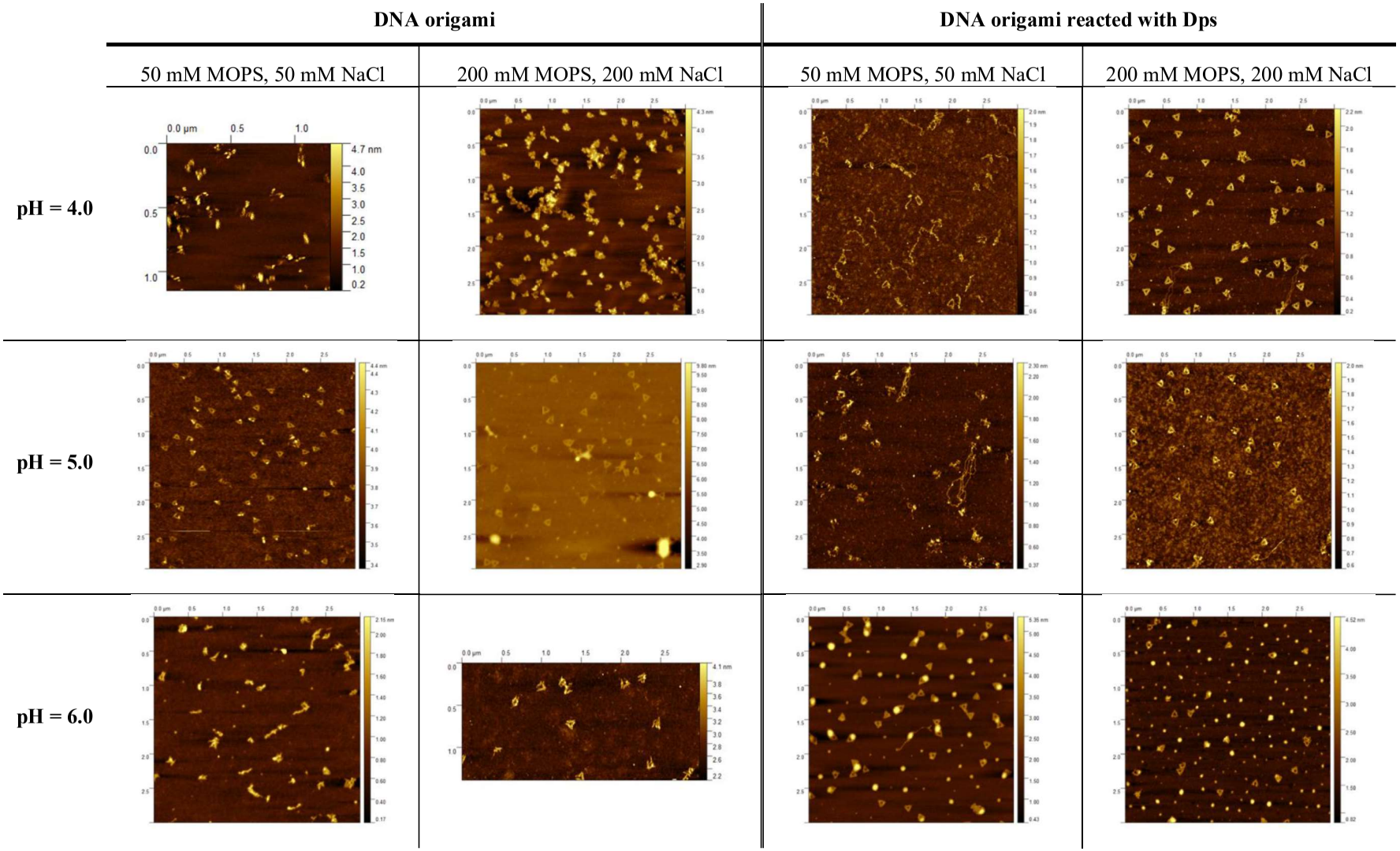

Figure 4-1: AFM micrographs of DNA origami and DNA origami reacted with Dps at pH = 4.0, 5.0 and 6.0, at low ionic strengths (50 mM MOPS, 50 mM NaCl) and high ionic strengths (200 mM MOPS, 200 mM NaCl).

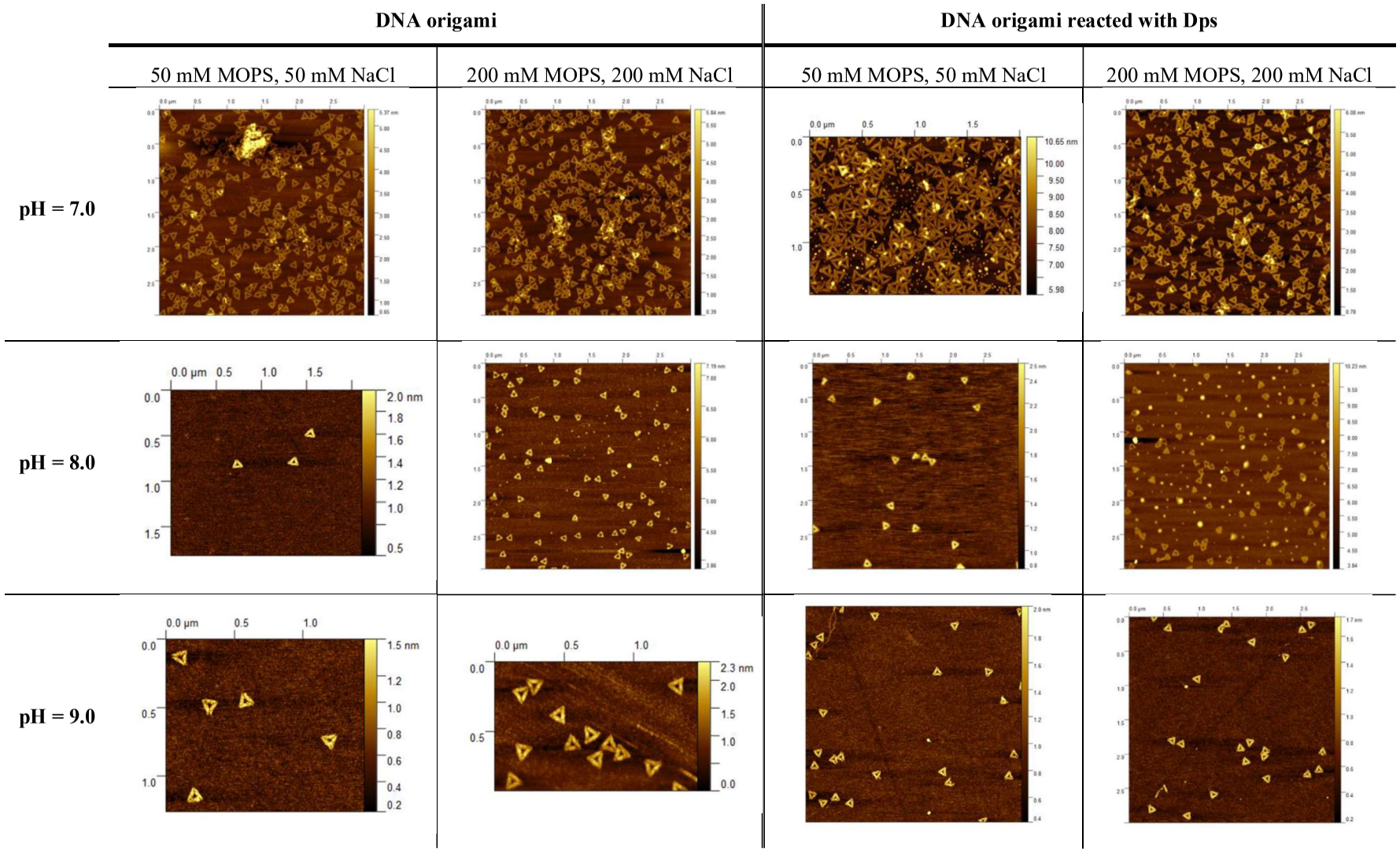

Figure 4-2: AFM micrographs of DNA origami and DNA origami reacted with Dps at pH = 7.0, 8.0 and 9.0, at low ionic strengths (50 mM MOPS, 50 mM NaCl) and high ionic strengths (200 mM MOPS, 200 mM NaCl).

At pH values over 7.0 the DNA origami triangles appear larger and in these experimental conditions there is no evidence that the Dps molecule is binding to the DNA origami. At pH 7.0 Dps protein is clearly detected near the DNA origami triangles. At lower pH values, below 6.0, the DNA origami starts to disassemble, a process that seems intensified in the presence of the protein. At pH 6.0 Dps appears to bind DNA origami. Further work was conducted at pH 5.0, 6.0 and 7.0 using 50 mM MOPS with 50 mM NaCl buffer. The structural alteration of DNA origami has been described on the literature, showing that higher concentrations of NaCl result in irregular height of the triangle, and that under pH = 4 degradation of the structure occurs. Basic pH seems to maintain the overall shape of the origami but details of the structure are blurred, probably due to hydrolysis and dehybrization [162].

Based on the location of the strands of interest in the DNA origami, the binding of Dps protein was expected to occur at 55 nm on each triangle edge, when one measures from the farthest location to the strand position, as depicted in ***Figure 4-3***.

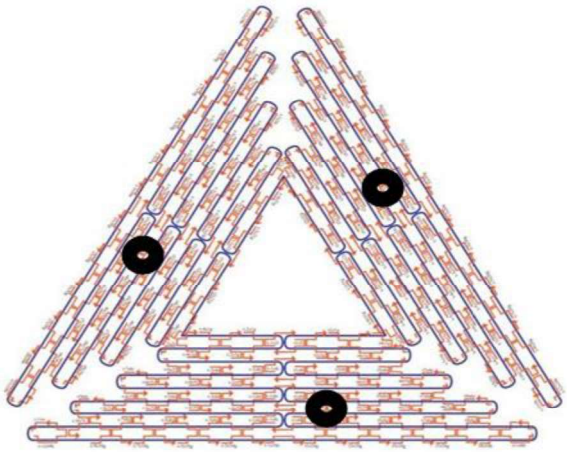

Figure 4-3*:* DNA origami map with strand locations highlighted.

The binding of Dps proteins to the DNA origami was considered when a Dps molecule was at the outer part of the triangle and its distance to the outer edge was measured relative to the closest edge. When more than one event occurred in the same triangle the distance was considered as if the triangle had 390 nm of length, the sum of its three sides lengths, and clockwise measurement iterations were performed, see ***Figure 4-4*** as an example.

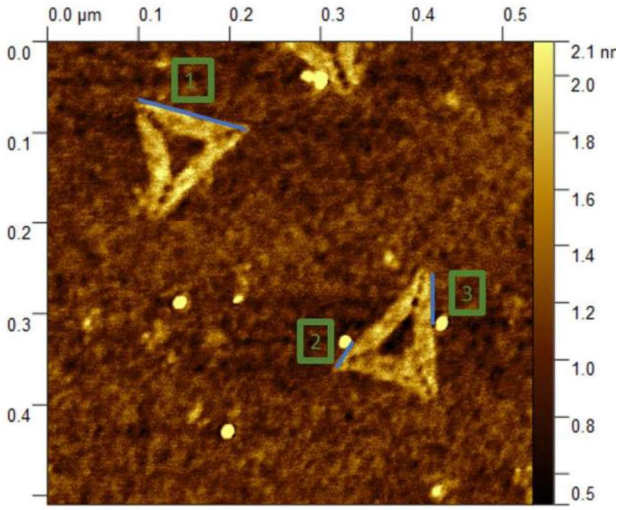

Figure 4-4: Example of the methodology followed to measure the binding of Dps to the DNA origami: 1 – DNA triangle origami with a 130 nm long side; 2 – Dps DNA origami binding event at 30 nm; and 3 – Dps DNA origami second binding event at 199 nm.

Dps WT was incubated with DNA origami with both double (dsDNA) and single stranded (ssDNA) DNA fragments of interest. However, no significant differences in binding were apparent, independently of the pH and temperature of incubation or DNA type, ***Figure 4-5***.

All conditions were taken together in a total of 201 events measured, being concluded that ca. 64 % of all binding events occur between 20 and 45 nm of the triangle side and that the vast majority of events occurs in a ratio of 1 Dps WT per triangle. The data show that the protein has distinct affinity to the different sites on the origami, with the region between 20 and 45 nm of the triangle side as the preferential binding site, with no correlation with the strands of interested being double or single stranded, ***Figure 4-5*** and ***Figure 4-6***.

It is known that Dps binds to plasmid DNA and that at pH between 6.5 and 7.5 binding and condensation is possible at pH values closer to 6.5 [163]. Due to the geometry constrains of the triangle, the experiments performed with DNA origami might hinder condensation. Nevertheless, the binding happened at acidic and neutral pH values, which is in line with the expected activity already known for Dps WT.

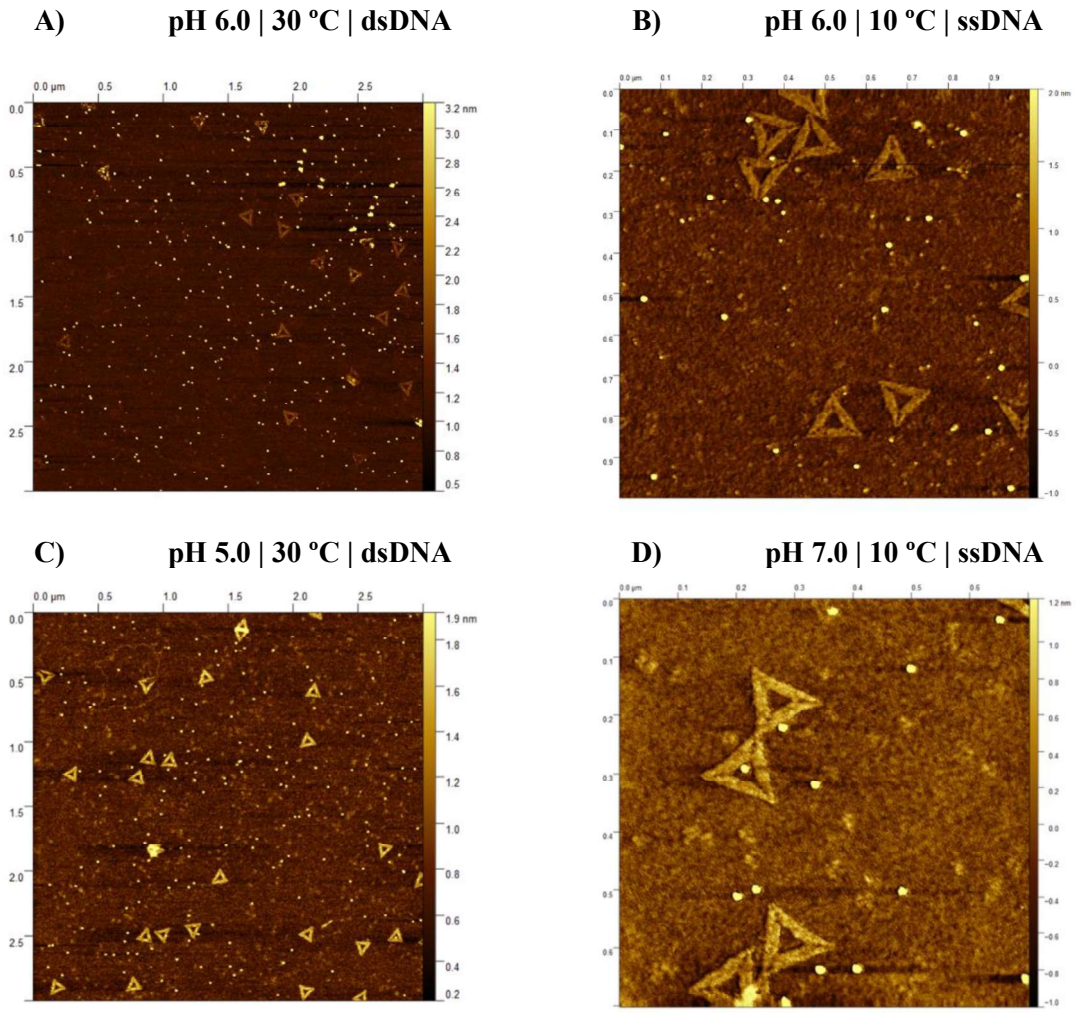

Figure 4-5: AFM micrographs of: DNA origami reacted with Dps. All experiments were performed in 50 mM MOPS, 50 mM NaCl buffer.

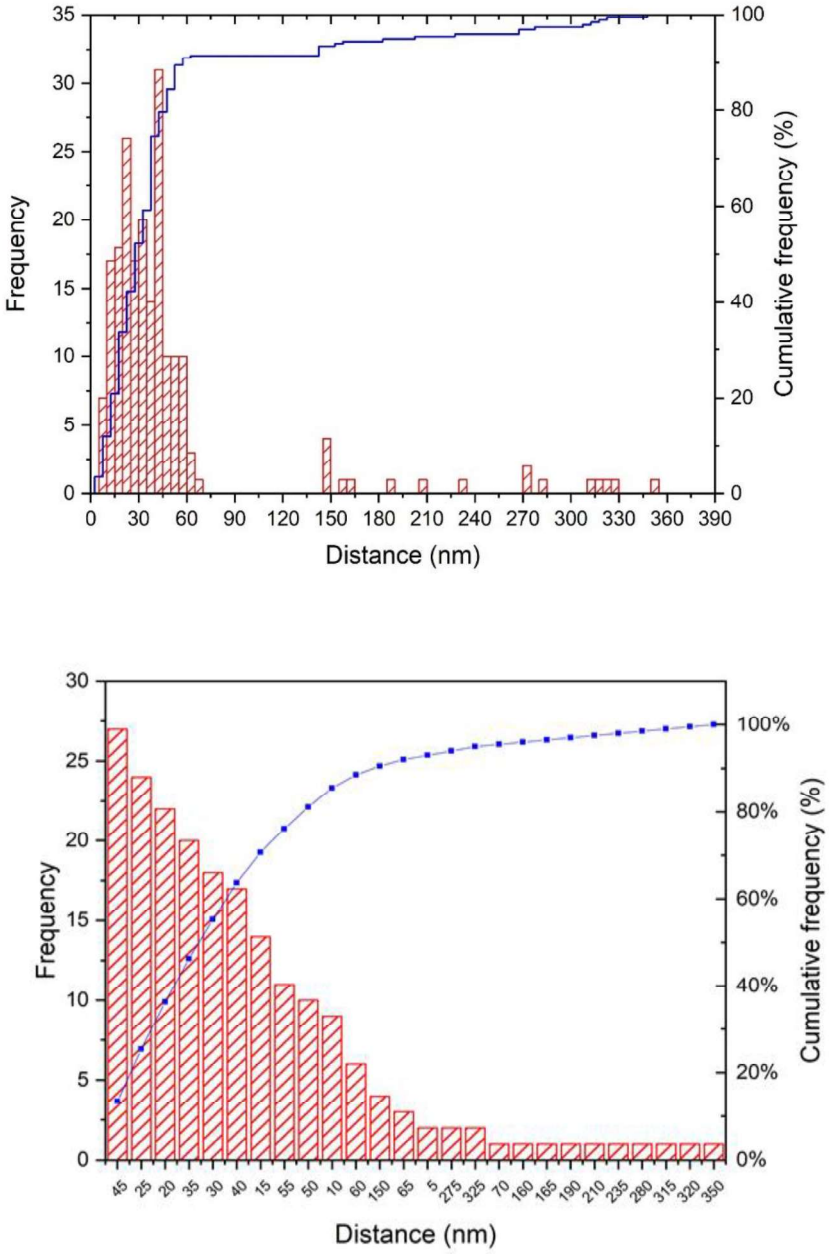

Figure 4-6: Pareto representations of the binding of Dps to DNA origami (n=201).

4.4 Conclusion

The binding of Dps WT to DNA origami triangles was analysed at different pH and buffer conditions. It was clearly shown that at low ionic strength (50 mM Mops, 50 mM NaCl) and more acidic conditions, the origami structures were destabilized. This destabilization was more expressive in the presence of the Dps protein. This elucidates the binding activity to DNA, and in this specific case to the DNA origami triangles, which on the Dps side is known to be enhanced at low ionic strengths and pH values.

Although it was clear that Dps bound to the DNA origami, more evident at pH 6.0 and low ionic strength, characterization of the binding to the specific DNA fragments was not possible. In fact, the AFM data suggest that the protein did not bind to those, directly interacting with the origami triangles.

One possible explanation could be that Dps has a greater affinity to the silicon substrate than to DNA. It is known that Dps binds non-specifically to DNA through electrostatic interactions, therefore a competition with silicon substrate might be occurring [164].

Experiments with different DNA origami shapes should be carried out in the future in order to study the binding effect of Dps WT to these structures. Another possible path to understand if this specific origami system is suitable for this goal would be to conduct the AFM measurements in solution, which would minimize the contribution of the drying process and be closer to the physiological state of Dps.

4.5 Manuscript: Dps-DNA interaction in *Marinobacter hydrocarbonoclasticus* protein: effect of a single charge alteration

João P. Jacinto‡[1], Daniela Penas‡[1], João P. L. Guerra[1], Ana V. Almeida[1], Nykola C. Jones[2], Søren V. Hoffmann[2], Pedro Tavares[1], Alice S. Pereira[1]

1 - Molecular Biophysics Laboratory., UCIBIO/Requimte, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.

2 - ISA, Department of Physics and Astronomy, Aarhus University, DK-8000 Aarhus C, Denmark

‡ These authors contributed equally to this work.

The present work has been published on the 26th of April 2021 by the European Biophysical Journal DOI: 10.1007/s00249-021-01538-0

Abstract

DNA-binding proteins from starved cells (Dps) are members of the ferritin family of proteins found in prokaryotes, with hollow rounded cube-like structures, composed of twelve equal subunits. These protein nanocages are bifunctional enzymes that protect the cell from the harmful reaction of iron and peroxide (Fenton reaction), thus preventing DNA damage by oxidative stress. Ferrous ions are oxidized at specific iron binding sites in the presence of the oxidant and stored in its cavity that can accommodate up to ca. 500 iron atoms. DNA binding properties of Dps are associated with the N-terminal, positive charge rich, extensions that can promote DNA binding and condensation, apparently by a cooperative binding mechanism. Here, we describe the binding and protection activities of *Marinobacter hydrocarbonoclasticus* Dps using Electrophoretic Mobility Shift Essays (EMSA), and Synchrotron Radiation Circular Dichroism (SRCD) spectroscopy. While no DNA condensation was observed in the tested conditions, it was possible to determine a Dps-DNA complex formation with an apparent dissociation constant of 5.9 ± 1.0 μM and a Hill coefficient of 1.2 ± 0.1. This interaction is suppressed by the inclusion of a single negative charge in the N-terminal region by point mutation. In Dps proteins containing a ferric mineral core (above 96 Fe/protein) DNA binding was impaired. SRCD data clearly showed that no significant modification existed either in secondary structure or protein stability of WT, Q14E variant and core containing proteins. It was, however, interesting to note that, in our experimental conditions, thermal denaturation induced protein aggregation that caused artifacts in thermal denaturation curves, which were dependent on radiation flux and vertical arrangement of the CD cell.

4.5.1 Introduction

Natural protein cages are profuse, including virus-like particle (VLP) chaperones, heat shock proteins and ferritins, among others, fulfilling a diversity of functions [165], [166]. Dps (DNA-binding proteins from starved cells), one of the smallest protein cages, are homopolymers

belonging to the ferritin family involved in the cellular response to multiple stresses (oxidative, thermal, UV and extreme pH) during the stationary phase [167][168]. It is able to form crystalline DNA-protein complexes to adapt to unfavourable conditions [169]. Their hollow rounded cube-like structure is composed of 12 subunits, folded into four-helix bundles (each monomer has an additional small helix)[170][171]. The Dps cage has an outer diameter of approximately 8-9 nm and an inner cavity of 4-5 nm, which can accommodate up to ∽500 ferric ions [12]. In most proteins, the inner cavity is connected to the exterior through two types of channels, a 3-fold symmetric channel formed by the junction of N-terminal tails (similar to the ferritin ones and thus designated ferritin-like channels) and a second type created by the C-terminus of monomers (due to its uniqueness, these were termed Dps-like channels)[4][172]. Iron traffic to and from the cavity occurs through the negatively charged ferritin-like channels directed by electrostatic fields [4].

The main function of Dps is iron detoxification, rather than iron storage. Sequestration of ferrous iron minimizes the formation of iron-mediated reactive oxidative species by Fenton reactions and, thus, cell damage. To fulfil this function Dps proteins contain unique ferroxidase sites located at the interface between two anti-parallel subunits (two ferroxidase sites per dimer, ∽2 nm apart, in a total of 12 sites), where ferrous ions are rapidly oxidized to the ferric state. The ferroxidase site is believed to be binuclear, with two non-equivalent iron binding sites, where two ferrous ions are converted into ferric species with reduction of one molecule of hydrogen peroxide (H_2O_2)[14][173]. Molecular oxygen can also oxidize ferrous ions, but at a much slower rate when compared with the co-substrate H_2O_2. Once oxidized, the ferric-oxo species are translocated to the cavity for mineral formation (biomineralisation reaction)[174].

While all Dps are able to sequester Fe2+ ions, oxidize and store them in the ferric mineral form, not all have the capacity to bind DNA. Some can interact with DNA (supercoiled or linear plasmid or genomic DNA) without sequence specificity, protecting it from chemical and physical damage. In the cell, Dps proteins condense DNA in the stationary phase, allowing for DNA protection against mechanical stress and enzymatic degradation [31][175]. The DNA-binding activity of Dps has been the motive for intensive debate in the scientific community. Given the fact that both DNA and Dps surfaces are negatively charged it has been proposed that the positively charged residues, variable in number, at the N- or C-terminal tails are responsible for DNA binding [37][3][175]. In a few cases, as *in Helicobacter (H.) pylori* Dps (also known as neutrophil-activating protein), the polyanion DNA molecule interacts with the positive surface of the protein [41].

Dps proteins that condense DNA have a positive (presence of lysine and/or arginine residues) and protruding N- or C-terminal extensions [176][177]. The absence of a protruding mobile tail nullifies the condensation activity. It has been shown that in the *Escherichia (E.) coli*

Dps, the interaction with DNA occurs due to three lysine residues in the mobile N-terminal extension, that interact with the phosphate groups on DNA[40]. Deletion of 2 or all 3 lysine residues impairs the DNA condensation activity in the first case and abolishes DNA binding in the second [16].

The requirement of a flexible N-terminal extension, besides multiple positive residues, for DNA binding and condensation was demonstrated for the *Agrobacterium (A.) fabrum* (formerly *Agrobacterium tumefaciens*) Dps. Despite the presence of 3 lysine residues in the N-terminal region, this protein, contrarily to the *E. coli* one, is unable to bind DNA because the N-terminal regions are immobilized through weak interactions to the dodecamer surface, namely hydrogen bond and electrostatic interactions [29]. Contrasting with other prokaryotic nucleoid-associated proteins or eukaryotic histones, binding of Dps to DNA does not affect the initiation of transcription and slightly alters the transcriptome of the *E. coli* stationary phase, but blocks the activity of endonucleases, in vitro [178].

With this work we aim to elucidate the role of the N-terminal extension of Dps from *Marinobacter hydrocarbonoclasticus* on DNA binding activity and protection. Apo and holo forms of wild-type (WT) Dps and variant (Q14E), wherein a single negative charge was placed on the N-terminal region, were incubated with a supercoiled plasmid, pUC19, and complex formation was studied by Electrophoretic Mobility Shift Assays (EMSAs).

DNA protection tests against hydroxyl radicals and proteolytic hydrolysis were also conducted. The secondary structure of wild-type Dps, apo and holo forms, and Q14E variant proteins was assessed by Synchrotron Radiation Circular Dichroism (SRCD), in the absence or presence of pUC19 plasmid DNA. Importantly, the use of a horizontally oriented sample holder in a non-conventional CD set-up, denoted the periscope chamber, is described to properly study the thermal stability of the protein.

4.4.2 – Materials and methods

4.5.2.1 Protein production and purification

Wild-type Dps protein was produced and purified as described by Penas et al. [33]. The DpsQ14E variant (in which glutamine 14 was replaced by a glutamic acid residue) was produced by site-directed mutagenesis using the pET21c-Dps as DNA template, applying the NZYMutagenesis kit from NZYTech. The oligonucleotide mutagenic primers used to produce the point mutation were designed and analysed for optimization with the software Primer Premier 6.24 from (Premier Biosoft International, Palo Alto, CA) and synthesized by NZYTech. Nucleotide sequence of mutagenic primers for point mutations (in bold) were 5'-GAC ACA GAC

AAA ACC GAG AAG CTG GCA GAC G-3' (forward) and 5'-C GTC TGC CAG CTT CTC GGT TTT GTC TGT GTC-3' (reverse). Briefly, each PCR reaction contained 35 ng of plasmid pET21c-Dps, 125 ng of each primer (forward and reverse), dNTP mix and 2.5 U of NZYProof DNA polymerase in the appropriate buffer, in a total volume of 50 µL. The PCR was performed in a thermocycler (Biometra, Invitrogen) following the manufacturer instructions using an annealing temperature of 60 °C. After analysis of the efficiency of the amplification on a 1% agarose gel, the remaining PCR mixture was incubated with 5 µL of DpnI for 1 h at 37 °C for elimination of the parental methylated and hemi-methylated DNA template. DpnI hydrolysis products (7 µL) were used to transform ultracompetent *E. coli* cells, NZYStar (NZYTech), plated on LB-Agar medium plates containing 100 µg/mL ampicillin and incubated overnight at 37 °C. Selected clones were analysed by DNA sequencing of plasmid DNA (sequenced by Eurofins Genomics). Dps-Q14E variant protein was produced in *E. coli* BL21(DE3) cells (NZYTech), in LB medium, at 37 °C for 3 h, following the Dps-WT protocol. After pelleting, cells (2.6 to 3.0 g/L of culture) were resuspended in 10 mM Tris-HCl buffer pH 7.6. The soluble extract collected after ultracentrifugation was dialyzed overnight at 4 °C against the same buffer prior to purification. The purification procedure of DpsQ14E variant was similar to the one described for the Dps-WT [33] using two chromatographic steps with ionic exchange columns (DEAE-Sepharose Fast Flow column followed by a Q-Resource column, both from GE Healthcare Life Sciences. The efficiency of all experimental steps was assessed by SDS-PAGE (12.5% acrylamide). Pure protein fractions were dialyzed overnight at 4 °C against 50 mM MOPS buffer pH 7.0, containing 200 mM NaCl and stored at 80 °C.

4.5.2.2 Electrophoretic mobility shift assays (EMSAs)

All electrophoretic mobility shift assays (EMSAs) were performed using supercoiled plasmid pUC19, in 50 mM MOPS buffer pH 7.0, containing 50 mM NaCl (designated as EMSA buffer). Dps proteins (at variable concentration), wild type and variant, apo and holo forms, were incubated with 10 nM or 15 nM of pUC19 in EMSA buffer in a total volume of 10 µL, for 10 min, at room temperature. EMSAs were analysed in 1 % agarose gels in 1x TAE buffer (40 mM Tris-acetate buffer pH 8.0, 0.1 mM EDTA), run at 80 V for about 1 h 30 min at room temperature. Gels were stained with SYBR Safe solution (Invitrogen) for 30 min and imaged with Safe Imager™ (Invitrogen) in a Gel Logic 100™ Imaging System (Kodak). Gel electrophoresis images were processed and analysed (free DNA and DNA-protein complex bands) using ImageJ [179]. For control assays, Dps protein was replaced by bovine serum albumin (BSA) (3 µM).

The DNA binding ability of Dps-WT loaded with iron (containing 12, 24, 48, 96, 192 and 384 iron atoms per protein) at a concentration of 13 µM was also evaluated by EMSA. Iron-loaded proteins were prepared using either H_2O_2 or O_2 as co-substrate and a ferrous sulphate solution

freshly prepared in milliQ water at pH 3.0, and H_2O_2 solution in 50 mM MOPS buffer pH 7.0, 50 mM NaCl. Dps-WT protein was loaded with 12 Fe^{2+} per protein molecule, incubated for 5 min at room temperature and reacted with a 15-fold molar excess of H_2O_2 to the concentration of ferrous iron, for 5 more minutes. Stepwise addition of iron and hydrogen peroxide were performed to achieve Fe2+/protein ratios of 12, 24, 48, 96, 192 and 384. After the last addition, the protein was incubated for 30 min at room temperature. When oxygen was used as co-substrate, the reaction mixture was saturated with oxygen gas following iron addition and every 12 h for 5 days at 4 ºC.

The relative DNA–Dps complex formation was plotted as a function of Dps protein concentration and fitted to the Hill equation 4.1:

$$f = f_{\max}[Dps]^n/(K_D + [Dps]^n) \quad (4.1)$$

where f is the fractional complex, f_{max} corresponds to 100% complex formation, [Dps] is the concentration of the binding protein, n is the Hill coefficient and K_D represents the macroscopic apparent dissociation constant and is a measure of the affinity of the protein to DNA.

4.5.2.3 Protease protection assays

The proteolytic resistance of the DNA-Dps-WT complex against Proteinase K was studied incubating Dps-WT (0.5 to 13 μM) with supercoiled plasmid pUC19 (10 nM), for 10 min at room temperature, in 50 mM MOPS buffer pH 7.0, 50 mM NaCl, followed by incubation with Proteinase K (20 mg/mL) for 30 min at 37 °C. A control was prepared by first incubating Dps-WT with Proteinase K for 30 min at 37 °C, and then with pUC19 for 10 min at room temperature. Additionally, Dps-WT was substituted with BSA protein (13 μM). The ability of Dps-Q14E protein (13 μM) to protect DNA was also assessed.

4.5.2.4 DNA protection assays against hydroxyl radicals

DNA protection assays against hydroxyl radicals were performed using supercoiled plasmid pUC19 and Dps-WT, iron loaded Dps-WT (96 Fe/protein) and Dps-Q14E proteins, in 50 mM MOPS buffer pH 7.0, 50 mM NaCl. First, a DNA binding reaction was prepared with pUC19 (10 nM) and protein (13 μM). Afterwards, 2.04 mM FeSO4.7H2O and 30.7 mM H_2O_2 were added to promote Fenton reactions during 8 min at room temperature. The reaction was stopped by addition of 10% SDS and incubation at 85 °C for 2 min. In parallel, and after the DNA binding reaction, Proteinase K (20 mg/mL) was added to the DNA−Dps-WT mixture (for 30 min at 37 °C), following reaction with ferrous sulphate and H_2O_2. As before, all reactions were analysed by electrophoresis in agarose gel.

4.5.2.5 Synchrotron Radiation Circular Dichroism (SRCD)

The synchrotron radiation SRCD spectra of Dps-WT protein were collected at the SRCD facility at the AU-CD beamline on the synchrotron radiation facility ASTRID2, at ISA, Department of Physics and Astronomy, Aarhus University, Denmark. The light passed from the synchrotron through a CaF_2 window into a CD instrument purged with nitrogen. In the instrument, a Rochon polarizer initially ensures linear polarization which is then converted into alternating circularly polarized light with a photo-elastic modulator (PEM, model I/ CF50, Hinds, USA) operating at ~50 kHz. The signal was detected with a photo-multiplier tube (PMT, 9402B, Electron Tubes, UK). Camphor-sulfonic acid served as a calibration material for the instrument. The spectra were recorded with 1 nm steps and a dwell time of 2.1 s per step in triplicate, using a nominally 0.01 cm quartz cells (SUPRASIL, Hellma GmbH, Germany), for the wavelength range of 170–280 nm. The actual pathlength of the cell was determined via an interference technique [180] to be 0.01008 cm. Protein samples were in 50 mM MOPS pH 7.0, 50 mM NaF at a final protein concentration of 1 mg/mL. Buffer spectra were used as baseline and subtracted from the protein spectra. Temperature scans were acquired from 5 ºC to 90 ºC using the following equilibrium times in between each increment: 5 minutes from 5 to 40 ºC, 2.5 minutes from 40 ºC to 65 ºC, 5 minutes from 65 to 90 ºC. Triplicate spectra at 25 ºC were obtained before and after the temperature scan.

Two types of observation chambers were used: i) a standard, dual Peltier chamber for which samples are mounted vertically; and ii) a periscope chamber that features a horizontal mounting position. The use of the latter eliminates or reduces the effect of displacement due to aggregation/precipitation of protein samples or buffer density changes. In order to calculate the α-helix and β-sheet content in the analysed secondary protein structure, suitable spectra were fitted using the Dichroweb online server [98], [181]. The fitting was conducted using the CDSSTR method and SP175 reference data set [181]. CD data was acquired at the BioLab, Biological and Chemical Analysis Facility, UCIBIO-LAQV/Requimte, FCT NOVA, Portugal, using an Applied Photophysics Chirascan™ qCD, and a UXL-150S Xenon short arc lamp, Hamamatsu. Spectra were recorded in triplicate with 1 nm steps and a dwell time of 2.7 s per step using the same pathlength calibrated 0.01008 cm quartz cells as in the SRCD measurements.

4.5.2.6 Dynamic light scattering (DLS)

DLS measurements were performed using a HORIBA SZ-100 equipped with a 10 mW 532 nm laser at a detector angle of 90º. Dps-WT samples were prepared at 1 mg/mL in 200 mM MOPS pH 7, 200 mM NaCl and centrifuged for 15 min at 14000 g before each measurement. Temperature scans from 25 to 60 °C were performed by measuring the Z-average at least three times at each temperature or repeatedly until three stable readings were obtained. Raw data results

were analysed using the equipment built-in software assuming a standard monodisperse form of distribution, using a particle refractive index of 1.6 (organic sample) and water settings as dispersion medium, with refractive index of 1.333 and viscosity given by the following equation 4.2:

$$\eta = 2.633x10^{-8} \cdot T^4 - 3.610x10^{-5} \cdot T^3 + 1.863x10^{-2} \cdot T^2 - 4.293 \cdot T + 3.736x10^2 \quad (4.2)$$

4.5.3 Results and discussion

4.5.3.1 DNA binding

The DNA binding capacity of the apo form (iron free) of *M. hydrocarbonoclasticus* Dps-WT protein was evaluated by EMSA (***Figure 4-7***). The results showed that the protein was able to bind to supercoiled plasmid pUC19. Binding of Dps-WT to pUC19 produced larger protein-DNA complexes which migrated slower than the free form of supercoiled pUC19 (***Figure 4-7*****A**, lanes 3 to 13). In the conditions tested, the intensity of the complex band increased with protein concentration, with an apparent saturation at around 10 μM Dps-WT. The observed binding was specific for Dps-WT, since substitution with BSA (a protein that does not bind DNA) did not show any shift of the free form of supercoiled pUC19 (***Figure 4-7*****A**, lane 2). The fractional complex formation was plotted as a function of protein concentration and fitted with the Hill equation, 4.1 (***Figure 4-7*****B**).

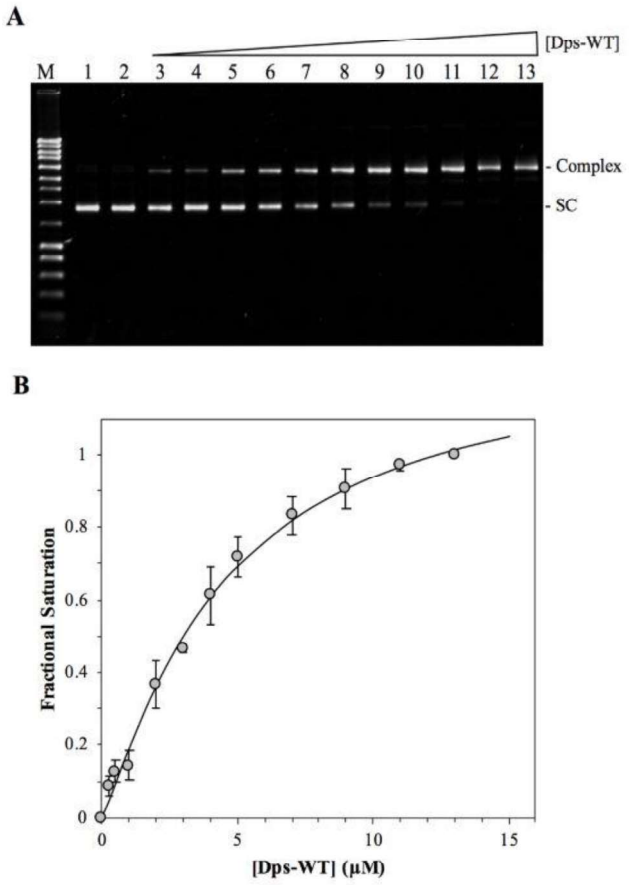

Figure 4-7: Binding of *M. hydrocarbonoclasticus* Dps-WT to supercoiled plasmid pUC19 in 50 mM MOPS buffer, pH 7.0, 50 mM NaCl, analysed by EMSA. (A) EMSA on 1% agarose gel. 1 – Supercoiled plasmid pUC19 (10 nM); 2 – Binding reaction of BSA (3 μM) to supercoiled pUC19 (10 nM); 3 to 13 – Binding of Dps-WT (0.25, 0.50, 1.0, 2.0, 3.0, 4.0, 5.0, 7.0, 9.0, 11 and 13 μM) to supercoiled pUC19 (10 nM). M – NZYLadder III. The free supercoiled plasmid pUC19 band (SC) and the pUC19-protein complex band (Complex) are identified on the right. (B) Hill plot of DNA binding by Dps-WT from four sets of EMSA experiments using increasing concentrations of protein (between 0.25 μM and 13 μM) and supercoiled plasmid pUC19 (10 nM). Less than 8% uncertainty is estimated for the fractional saturation.

In our experimental conditions, we calculated an apparent dissociation constant, K_D of 6 ± 1 μM and a Hill coefficient, *n* of 1.2 ± 0.1. The Hill coefficient indicates a moderate positive cooperativity on DNA binding and is comparable with the values reported in the literature for the binding of *Deinococcus radiodurans* Dps-1 to a 26-bp DNA duplex (n of 1.3 ± 0.2) or to the binding of *Desulfovibrio vulgaris* Dps-like bacterioferritin to a supercoiled plasmid DNA (n = 1.3 ± 0.2) [154] [182]. In turn, the K_D value obtained for *M. hydrocarbonoclasticus* Dps is similar to the one reported for *H. pylori* Dps binding to a linearized long plasmid DNA (3100 bp) [41]. It has been reported that most DNA-Dps interactions are mediated by a mobile N- or C-terminal extension, dependent on the presence of positively charged residues (lysines and/or arginines) [42][12][183]. The N-terminal tail of *M. hydrocarbonoclasticus* Dps (MGKNFIGLDTDKTQK)

contains 3 lysine residues positively charged at pH 7.0, that can establish electrostatic interactions with DNA.

In the experimental conditions tested, *M. hydrocarbonoclasticus* Dps did not promote condensation of the DNA. The presence of a single sharp band and the electrophoretic migration of the complex band suggest that *M. hydrocarbonoclasticus* Dps binds to supercoiled circular DNA changing its topology to a more relaxed form. Other Dps proteins are unable to condense DNA, namely the Dps-1 from Mycobacterium smegmatis, Dps from *Mycobacterium abscessus* subsp. *massiliense* and Dps4 and Dps5 from the *cyanobacterium Nostoc punctiforme* [41][184][185]. This behaviour was attributed to the presence of 4 negatively charged amino acid residues in the C-terminal extension that would compensate (or partly compensate) the 5 positively charged ones. This could explain the inability of *M. hydrocarbonoclasticus* Dps of DNA condensation, since 3 lysines and 2 aspartate residues are present on its N-terminal extension, composed of 15 amino acid residues.

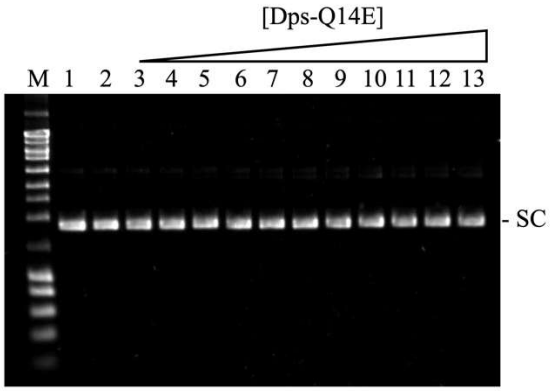

Figure 4-8: EMSA of the binding reaction of variant Dps-Q14E with supercoiled plasmid pUC19 in 50 mM MOPS buffer pH 7.0, 50 mM NaCl. 1 – Supercoiled plasmid pUC19 (15 nM); 2 – Binding reaction with BSA (3 μM); 3 to 13 – Binding reaction with Dps-Q14E (0.21, 0.375, 0.75, 1.05, 1.5, 1.8, 2.7, 3.75, 4.5, 6.0 and 7.5 μM) and pUC19 (15 nM). M – NZYLadder III. The free supercoiled plasmid pUC19 band (SC) is indicated.

To probe the effect of negatively charged residues on the N-terminal extension, a variant was prepared, replacing glutamine 14 by a glutamate residue, adding a negative charge at pH 7.0. The EMSA gel is presented in ***Figure 4-8***. As it is clear DpsQ14E variant was not able to bind DNA, demonstrating that a single charge residue substitution was enough to alter the binding affinity of the protein towards the DNA. In fact, the mutation resulted in a substantial decrease of the pI of the first 15 residue fragment, from 8.25 for the wild type to 5.88 for the Q14E variant. Alternatively, the non-condensation capacity of *M. hydrocarbonoclasticus* Dps could be due to the fact that it does not self-aggregate. Effectively, the ability of *Escherichia coli* Dps of forming large DNA-Dps complexes and DNA condensation has been attributed to its self-aggregation

propensity[42]. To investigate the affinity for DNA of iron-loaded Dps protein, samples of holoprotein were prepared with different amounts of ferric mineral core inside its inner cavity, using H_2O_2 or O_2 as co-substrate. The DNA binding capacity of the holo-protein was assessed by EMSA. ***Figure 4-9*** shows that the iron-loaded (holo) forms have lower affinity to DNA than the apo-Dps (iron-free), decaying with higher iron loads. When the protein was loaded with substoichiometric amounts of Fe^{2+} ions (relative to the putative iron binding sites), between 12 and 48 Fe^{2+}/protein, some complex was still observed (lanes 4 to 6 in gels in ***Figure 4-9***), that had become null for proteins with iron loads above 48. This observation had led us to suggest that when the iron was mainly in the mineral core form (ferric mineral in the protein inner cavity), the protein no longer has the ability to bind DNA.

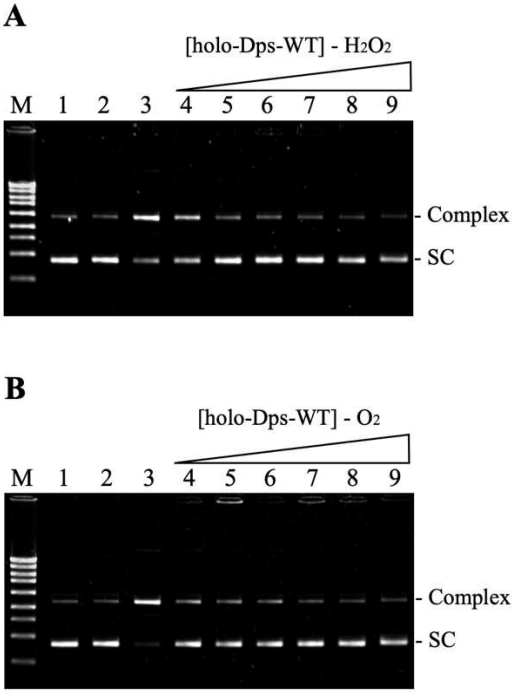

Figure 4-9: DNA binding analysis of iron-loaded Dps in 50 mM MOPS buffer pH 7.0, 50 mM NaCl. Binding reaction with iron loaded Dps-WT produced in the presence of H_2O_2 (A) or O_2 (B). 1 – Plasmid pUC19 (10 nM); 2 – Binding reaction of BSA (3 μM) to plasmid pUC19 (10 nM); 3 – Binding control of apo-Dps-WT (13 μM) to plasmid pUC19 (10 nM); 4 to 9 – Binding of Dps-WT loaded with 12, 24, 48, 96, 192 and 384 Fe^{2+}/protein; M – NZYLadder II (NZYTech). The free supercoiled plasmid band (SC) and the protein-DNA band (complex) are indicated on the right.

Tosha and colleagues have proposed a role for the N-terminal extensions of M ferritin from frog, a maxi-ferritin composed of 24 subunits [186]. According to the authors the N-terminal extensions would function as pore gates for ferrous ions entry/exit to/from the protein inner cavity through the ferritin-like channels. The pore gates would participate in a network of interactions (hydrogen bonds and salt bridges), involving residues Q6 and N7 (frog M ferritin numbering)

from the N-terminal, R72 (in helix 2) and R122 (from the loop between helices 2 and 3), conserved in all eukaryotic ferritins.

Although these residues are absent in Dps both types of ferritins share the same 3-fold symmetry channel, the so-called ferritin-like channels, and possess protruding N-terminal extensions. As mentioned before, the channels create an electrostatic potential, due to the presence of negatively charged residues (aspartates and glutamates) that direct the Fe^{2+} ions traffic to and from the protein cavity, down electrochemical gradient [4]. While absent in some Dps, we can envisage a similar role for amino acid residues N4 and Q13 in the N-terminal extension of *M. hydrocarbonoclasticus* Dps, gating the ferritin-like channels. In the presence of a ferric mineral core inside the protein, subtle structural changes would occur at the channel gate (closing the gate to avoid leakage of Fe^{2+} ions) inhibiting DNA binding, allowing the protein to act as an iron storage enzyme.

Nevertheless, binding of Dps to DNA seems to occur in different typologies, depending on the protein and microorganism. Although the most common mechanism involves the positive N- or C-terminal extensions or surfaces, the requirement of ferrous ions and H_2O_2, has also been described. In the case of Dps from *Campylobacter jejuni*, DNA binding seems to be induced by the presence of Fe^{2+} ions or H_2O_2 [187]. In turn, iron loading (400 Fe/protein) did not affect the DNA binding capacity of *H. pylori* Dps proteins [41]. Despite the extensive number of studies on the topic, the mechanism of Dps binding to the different topologic forms of DNA still not completely understood and further studies are needed to clarify this important process.

4.5.3.2 Assessing the secondary structure of proteins and DNA-Dps complex

SRCD was used to evaluate if the Q14E point mutation and the presence of a mineral core in the wild-type protein inner cavity (holo form with 96 Fe/Dps protein), affected DNA binding by altering the secondary structure of *M. hydrocarbonoclasticus* Dps. SRCD spectra are presented in ***Figure 4-10*** and are consistent with proteins with a high content of α-helices as expected from crystallographic structures of homologs. It should be noted that SRCD data is significant because it can reveal subtle changes in protein secondary structure. This is due to the high energies and photon flux achieved by the synchrotron beam, which makes possible the use of an extended spectral range from 175 to 195 nm when compared to conventional benchtop spectrometers.

Table* 4-1** shows the calculated secondary structures percentages for each sample (see ***Figure 4-16 in SI for the theoretical fits of each spectrum). Data obtained for the apo (***Figure 4-10***, white circles) and holo forms of Dps-WT (***Figure 4-10***, blue circles) showed that the

presence of 96 Fe/Dps mineral core did not induce any significant differences at the secondary structure level (displaying a slight increase in α-helices).

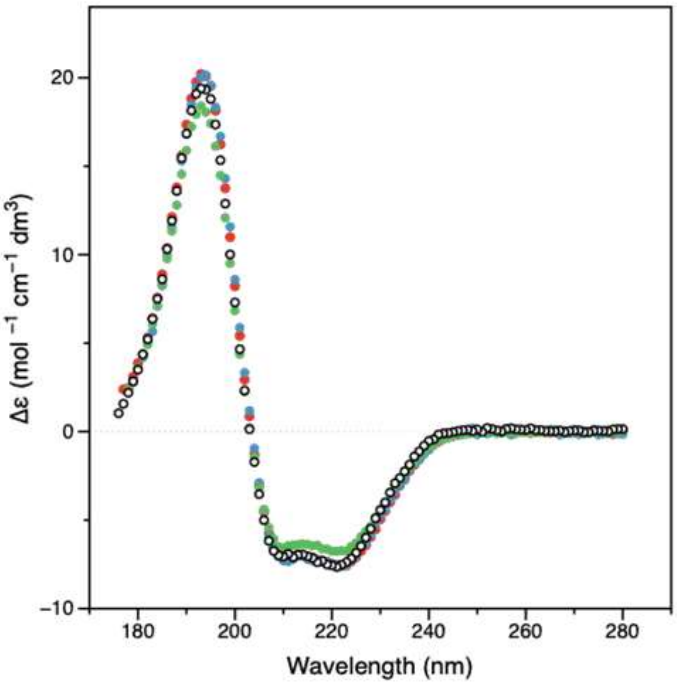

Figure 4-10: SRCD spectra obtained at 25 °C for apo Dps-WT (white circles); Dps-Q14E variant (red circles); Holo form with 96 Fe/Dps-WT protein (blue circles) and Dps-WT (4 μM) incubated with supercoiled pUC19 (125 nM) (green circles).

The Dps-Q14E variant (***Figure 4-10***, red circles) also maintained the structural features of Dps-WT. A more significant difference was observed for the protein incubated with pUC19 ***Figure 4-10***, green circles**)**, where spectral analysis showed a reduction of 10% in α-helices and an increase of 6% of unordered component. This suggests a more flexible structure in agreement with the rationale leading to the formation of DNA-Dps complex observed in the EMSA experiments described above. Thermal stability was probed for all protein forms, with similar results being obtained. The data obtained for Dps-WT displayed a characteristic behaviour observed for all tested samples and will be used here as an example of such peculiar behaviour. Comparable temperature scans between 5 and 90 °C were obtained for different experimental setups and, as shown in ***Figure 4-11A***, the denaturation profile was setup dependent. In both SRCD datasets acquired using the standard dual Peltier chamber (***Figure 4-11,*** green circles) and the Chirascan standard chamber (***Figure 4-11,*** orange circles) an inflexion around 40 to 45 °C was noticeable, albeit with differences in the overall profile. Such inflexions were not observed in SRCD data acquired with the periscope chamber, which clearly suggests aggregation leading to precipitation of protein samples.

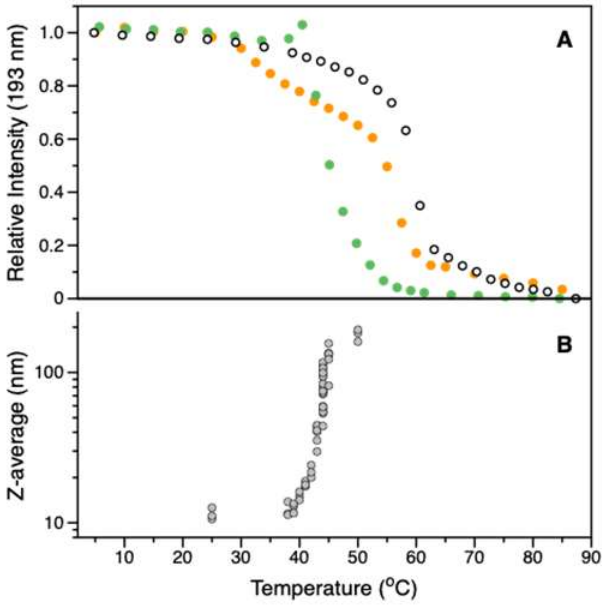

Figure 4-11: Thermal denaturation studies of Dps-WT protein (A). Data was obtained using SRCD with periscope chamber (white circles); CD with a Chirascan standard chamber (orange circles); and SRCD with standard dual peltier chamber (green circles). Thermal denaturation curves are derived from signal intensity at 193 nm versus temperature with proper baseline subtraction and normalization. Panel B shows the Z-average size calculated from raw DLS data as a function of temperature.

To corroborate this conclusion, DLS temperature dependence data was obtained for the same protein/buffer system. ***Figure 4-11B*** revealed that the Z-average size abruptly increases at approximately 40 °C, supporting the aggregation behaviour inferred from the multiple SRCD and CD data. SRCD results obtained with the periscope chamber setup also showed that no loss (or little loss) of α-helical structure occurred upon aggregation and that the thermal denaturation was better described by a Tm temperature of ~60 °C. The results described here demonstrate the importance of a horizontal mounting position such as the one used in the periscope chamber setup. More, if unaware, the possibility exists to extract misleading information about structural thermal stability from circular dichroism data. To try to explain differences in the curves obtained using the standard dual Peltier chamber and the Chirascan standard chamber, beam sizes and photon flux were measured for all used setups (***Figure 4-12***). As expected, photon flux below 185 nm was different for synchrotron beam-based setups and the Chirascan setup, but no significant differences were observed in less energetic wavelengths (up until 250 nm).

Table 4-1: Calculated secondary structures percentages obtained by analysis with Dichroweb online server.

Secondary Structure (%)	Sample			
	Dps-WT	Dps-Q14E	Dps-WT holo form	Dps-WT + pUC19
Regular Helix	55	57	58	47
Distorted Helix	23	24	24	20
Regular Strand	5	4	5	7
Distorted Strand	1	1	0	3
Turns	7	6	5	8
Unordered	9	8	8	15

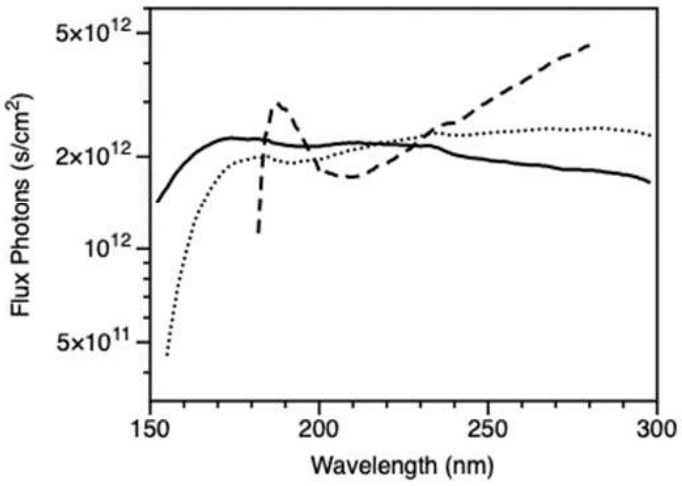

Figure 4-12: Photon flux measurements for the SRCD and CD setups used. Data was obtained using standard dual peltier chamber (solid line) periscope chamber (dotted line); and Chirascan standard chamber (dashed line).

Table 4-2: Beam size dimensions used in circular dichroism setups

Beam size	Synchrotron radiation	Benchtop	
	Standard dual Peltier chamber	Periscope chamber	Chirascan standard chamber
Width (mm)	3.5	4	8
Height (mm)	5	4.5	11
Area (mm^2)	17.5	18	88

Beam size measurements showed that the synchrotron beam-based setups have a significantly smaller area than the benchtop spectrometer used (***Table 4-2***). In fact, in the benchtop spectrometer the beam size on the sample is almost 5x that for the synchrotron radiation facility, thus explaining the different profiles observed, created by aggregation (and precipitation) displacement effects in the cell.

4.5.3.3 Protection assays

To investigate DNA and protein protection by the formation of DNA-Dps complexes, two different assays were performed.

4.5.3.3.1 Protection of Dps against protease degradation

To test if Dps was protected against proteolysis when bound to DNA, the DNA-protein complex was digested with Proteinase K (NZYTech), a serine protease with a broad substrate specificity. As expected, most of the protein was hydrolysed when *M. hydrocarbonoclasticus* Dps was incubated with Proteinase K, (***Figure 4-13*A**, lane 8 and ***Figure 4-13*B**, lane 12). The same result was obtained for BSA and Dps-Q14E protein (***Figure 4-13*B**, lanes 2 and 4). The incubation of the DNA-protein complex with Proteinase K (***Figure 4-13*A**, lanes 2 to 7) did not affect the complex, while the pre-incubation of Dps-WT protein with Proteinase K prevented the formation of protein-DNA complex (***Figure 4-13*B**).

Staining the EMSA gels for protein detection and comparison of the intensity of Dps bands revealed that the proteolysis was less extensive on the complex, than when the free protein was incubated with the protease (bottom gels in ***Figure 4-13***, lanes 2 to 7 on gel A versus lanes 6 to 11 on gel B). In the assays with Proteinase K, the Dps band presented increased mobility when compared with the free protein one, suggesting the presence of a peptide site more susceptible to proteolysis in the Dps sequence, resulting in truncated Dps molecules. This could also explain the inability of DNA binding when Dps was pre-incubated (before the DNA binding reaction) with

the protease. In fact, in the N-terminal tail (first 15 residues) there are 5 possible cleavage sites for proteinase K (in bold face in the sequence MGKN**FI**G**L**D**T** DK**T**QK).

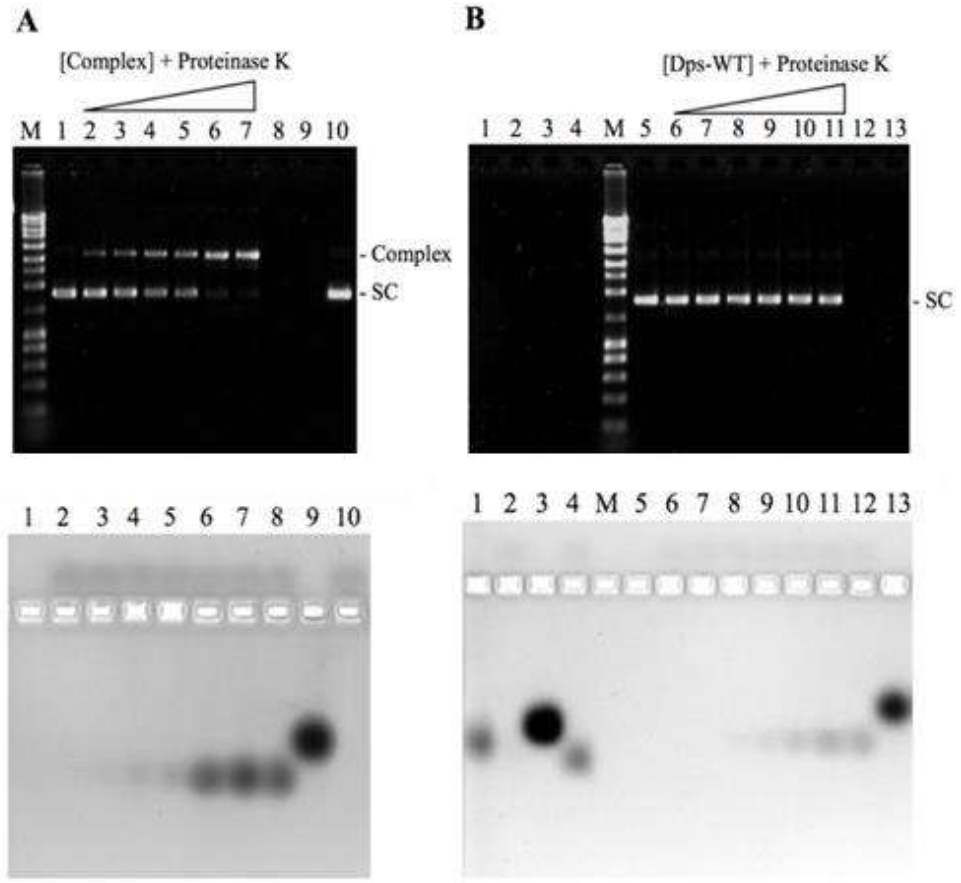

Figure 4-13: Dps protein protection assays against Proteinase K. (A) DNA-protein complex incubated with Proteinase K: 1 – Plasmid pUC19 (10 nM); 2 to 7 – Binding reaction of Dps-WT (0.50, 1.0, 3.0, 5.0, 9.0 and 13 μM) to plasmid pUC19 (10 nM) and then incubated with Proteinase K (20 mg/mL); 8 – Dps-WT (13 μM) incubated with Proteinase K; 9 – Dps-WT (13 μM); 10 – plasmid pUC19 (10 nM) incubated with Proteinase K. (B) Binding of Dps-WT pre-incubated with Proteinase K to supercoiled plasmid pUC19: 1 – BSA protein (13 μM); 2 – BSA (13 μM) incubated with Proteinase K; 3 – Dps-Q14E protein (13 μM); 4 – Dps-Q14E (13 μM) incubated with Proteinase K; 5 – Plasmid pUC19 (10 nM); 6 to 11 – Binding reaction of pre-incubated Dps-WT (0.50, 1.0, 3.0, 5.0, 9.0 and 13 μM) with Proteinase K to plasmid pUC19 (10 nM); 12 – Dps-WT (13 μM) incubated with Proteinase K; 13 – Dps- WT protein (13 μM). M – NZYLadder III. The free supercoiled plasmid band (SC) is indicated on the right. Upper gels stained with SYBR Safe for detection of DNA, and lower gels stained with Coomassie Brilliant Blue to detect protein.

4.5.3.3.2 Protection of DNA against hydroxyl radical degradation

All Dps proteins have in common the capacity to protect DNA from damage by hydroxyl radicals generated by the Fenton reaction, since they catalyse the uptake of Fe^{2+} ions and H_2O_2 from solution (ferroxidase activity). Therefore, supercoiled pUC19 was first incubated with *M. hydrocarbonoclasticus* Dps proteins (Dps-WT, Dps-Q14E and iron-loaded wild-type Dps

(containing 96 Fe atoms per protein) and then exposed to an excess of Fe^{2+} ions and H_2O_2. The results are shown in ***Figure 4-14***. Supercoiled DNA was completely hydrolysed in the presence of hydroxyl radicals (lane 2), whereas the pre-incubation with Dps proteins resulted in DNA protection (lanes 3 to 6). However, in the presence of Proteinase K (lane 4), DNA suffered more degradation when compared with the reaction without the protease (lane 3). Protection of plasmid pUC19 was still observed when pre-incubated with Dps-Q14E variant or iron-loaded Dps protein (lanes 5 and 6). Since, as demonstrated above, these two protein forms do not bind DNA (***Figure 4-8*** and ***Figure 4-9***), the results indicate that protection of DNA was due to the ferroxidase activity (ferrous irons uptake in the presence of H_2O_2) of the Dps enzyme rather than to a physical shielding effect by protein binding.

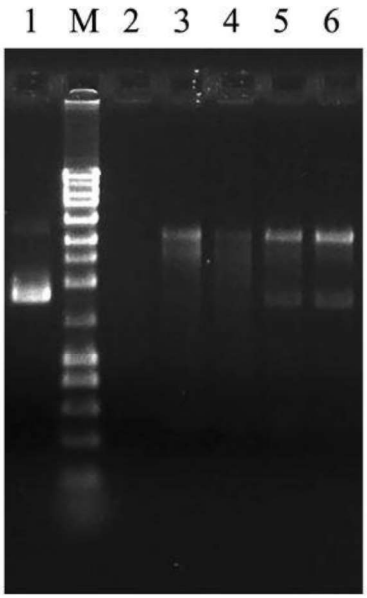

Figure 4-14: Analysis of DNA protection against hydroxyl induced DNA damage. 1—plasmid pUC19 (10 nM); 2—plasmid pUC19 (10 nM) exposed to an excess of Fe^{2+} and H_2O_2; 3—Dps-WT–pUC19 complex reacted with Fe^{2+} and H_2O_2; 4—Dps-WT–pUC19 complex pre-incubated with Proteinase K and reacted with Fe^{2+} and H_2O_2; 5— Dps-Q14E variant incubated with pUC19 and then reacted with Fe^{2+} and H_2O_2; 6—iron-loaded Dps-WT (96 Fe^{2+}/protein) incubated with pUC19 and then reacted with Fe^{2+} and H_2O_2. M—NZYLadder III

4.5.4 Conclusion

The principal function of proteins from the Dps family is DNA protection, mainly through their ferroxidase activity. Besides, some Dps also have the ability to non-specifically bind DNA protecting it from oxidative damage. When present, the flexible positively charged N- or C-terminal extensions have been indicated as responsible for the interaction of Dps with DNA. *M. hydrocarbonoclasticus* Dps contains this structural element at the N-terminal extension. In this work EMSAs were performed using wild-type Dps (apo and holo forms) and a variant. A point

mutation replacing glutamine-14 residue by a glutamate was introduced to infer about the importance of the N-terminal extension pI value. We do believe that this parameter is crucial for DNA binding. If the pI of the flexible N- or C-terminal extensions is not above the pKa of positive residues (lysines and arginines) DNA binding will not occur, even when these amino acid residues are present. Results confirmed our hypothesis and have shown the ability of *M. hydrocarbonoclasticus* Dps to bind to supercoiled plasmid DNA with a KD of 6 ± 1.0 μM. Dps variant Q14E was not able to interact with DNA.

Contrarily to what was described in the literature [41][187], iron-loaded *M. hydrocarbonoclasticus* Dps showed no affinity to bind DNA. It is likely that the presence of Fe3+ ions, may affect the structure of the "ferritin-like" pores in a mechanism similar to the one proposed for mammalian fast ferritins [186].

Formation of the protein-DNA complex proved to be important for both protein and DNA protection. Assays performed with Proteinase K demonstrated that the formation of the Dps-DNA complexes could physically shield the accessibility of the protease to the Dps molecules bound to DNA. On the other hand, protection assays against hydroxyl radicals indicated that DNA was protected due to the ferroxidase activity of the protein, since both holo-Dps-WT (containing 96 Fe per molecule) and the Dps-Q14E variant protein were still able to protect it against oxidative damage, by rapidly oxidizing Fe^{2+} ions using H_2O_2 as oxidant.

SRCD spectroscopy was used to probe the structural integrity of apo Dps-WT and holo protein, variants and protein-DNA complex. It is clear from the collected data that all protein forms retain the secondary structure characteristics expected for the Dps nanocage structure. The major difference was observed for the Dps-WT protein incubated with DNA where, while maintaining similar overall secondary structural features, a loss of α-helices components can be perceived as a perturbation upon protein-DNA complex formation. The thermal stability studies show that protein aggregation (and possible precipitation) arises above 40 ºC, a behaviour confirmed by DLS. This impairs the use of thermal denaturation studies to further probe structural differences between the studied proteins and protein-DNA complexes. It is however noteworthy to mention that the use of a periscope chamber can overcome such problem by preventing displacement effects. The smart design of a periscope chamber that places the quartz cell in a horizontal mounting position maintains the possibility of acquiring representative spectra at an extended energy range and high photon flux for protein solutions where aggregation might occur. Using this experimental setup, several SRCD experiments (in parallel with other spectroscopic and biochemical techniques) are currently being devised to understand the aggregation behaviour of *M. hydrocarbonoclasticus* Dps in different experimental conditions (ionic strength, pH,

crowding agents) and characterize the structural dynamics of such an interesting bionanocage system.

4.6.5 Supplementary Information

4.6.5.1 Protein expression and purification

Wild-type Dps protein was produced and purified as described by Penas et al. [33]. The DpsQ14E variant (in which glutamine 14 was replaced by a glutamic acid residue) was produced by site-directed mutagenesis using the pET21c-Dps as DNA template, applying the NZYMutagenesis kit from NZYTech. The oligonucleotide mutagenic primers used to produce the point mutation were designed and analysed for optimization with the software Primer Premier 6.24 from (Premier Biosoft International, Palo Alto, CA) and synthesized by NZYTech. Nucleotide sequence of mutagenic primers for point mutations (in bold) were 5'-GAC ACA GAC AAA ACC GAG AAG CTG GCA GAC G-3' (forward) and 5'-C GTC TGC CAG CTT CTC GGT TTT GTC TGT GTC-3' (reverse).

Briefly, each PCR reaction contained 35 ng of plasmid pET21c-Dps, 125 ng of each primer (forward and reverse), dNTP mix and 2.5 U of NZYProof DNA polymerase in the appropriate buffer, in a total volume of 50 μL. The PCR was performed in a thermocycler (Biometra, Invitrogen) following the manufacturer instructions using an annealing temperature of 60 °C. After analysis of the efficiency of the amplification on a 1% agarose gel, the remaining PCR mixture was incubated with 5 μL of DpnI for 1 h at 37 °C for elimination of the parental methylated and hemi-methylated DNA template. DpnI hydrolysis products (7 μL) were used to transform ultracompetent E. coli cells, NZYStar (NZYTech), plated on LB-Agar medium plates containing 100 μg/mL ampicillin and incubated overnight at 37 °C. Selected clones were analyzed by DNA sequencing of plasmid DNA (sequenced by Eurofins Genomics). Dps-Q14E variant protein was produced in *E. coli* BL21(DE3) cells (NZYTech), in LB medium, at 37 °C for 3 h, following the Dps-WT protocol.

After pelleting, cells (2.6 to 3.0 g/L of culture) were resuspended in 10 mM Tris-HCl buffer pH 7.6. The soluble extract collected after ultracentrifugation was dialyzed overnight at 4 °C against the same buffer prior to purification. The purification procedure of DpsQ14E variant was similar to the one described for the Dps-WT [33] using two chromatographic steps with ionic exchange columns (DEAE-Sepharose Fast Flow column followed by a Q-Resource column, both from GE Healthcare Life Sciences. The efficiency of all experimental steps was assessed by SDS-PAGE (12.5% acrylamide). Pure protein fractions were dialyzed overnight at 4 °C against 50 mM MOPS buffer pH 7.0, containing 200 mM NaCl and stored at −80 °C. The iron content of pure protein preparations was determined by the TPTZ method [188] and by inductively coupled plasma atomic emission spectroscopy (ICP-AES) and accounted for less than 1 Fe/protein.

4.6.5.2 Electrophoretic mobility shift assays (EMSAs)

All electrophoretic mobility shift assays (EMSAs) were performed using supercoiled plasmid pUC19, in 50 mM MOPS buffer pH 7.0, containing 50 mM NaCl (designated as EMSA buffer). Dps proteins (at variable concentration), wild type and variant, apo and holo forms, were incubated with 10 nM or 15 nM of pUC19 in EMSA buffer in a total volume of 10 µL, for 10 min, at room temperature. EMSAs were analysed in 1 % agarose gels in TAE buffer (40 mM Tris-acetate buffer pH 8.0, 0.1 mM EDTA), run at 80 V for about 1 h 30 min at room temperature. Gels were stained with SYBR Safe solution (Invitrogen) for 30 min and imaged with Safe Imager™ (Invitrogen) in a Gel Logic 100™ Imaging System (Kodak). Gel electrophoresis images were processed and analysed (free DNA and DNA-protein complex bands) using ImageJ [179]. For control assays, Dps protein was replaced by bovine serum albumin (BSA) (3 µM).

The DNA binding ability of Dps-WT loaded with iron (containing 12, 24, 48, 96, 192 and 384 iron atoms per protein) at a concentration of 13 µM was also evaluated by EMSA. Iron-loaded proteins were prepared using either H_2O_2 or O_2 as co-substrate and a ferrous sulphate solution freshly prepared in milliQ water at pH 3.0. H_2O_2 solution was prepared in 50 mM MOPS buffer pH 7.0, containing 50 mM NaCl. Dps-WT protein was first loaded with 12 Fe^{2+} per protein molecule, incubated for 5 min at room temperature and reacted with a 15-fold molar excess of H_2O_2 to the concentration of ferrous iron, for 5 more minutes. Stepwise addition of iron and hydrogen peroxide were performed as described for the first 12 Fe^{2+}/protein to achieve Fe^{2+}/protein molar ratios of 12, 24, 48, 96, 192 and 384. After the last addition, the protein was incubated for 30 min at room temperature. When oxygen was used as co-substrate, the reaction mixture was saturated with oxygen gas following iron addition and every 12 h for 5 days at 4 °C.

The relative DNA–Dps complex formation was plotted as a function of Dps protein concentration and fitted to the Hill equation, $f = f_{\max}[Dps]^n/(K_D + [Dps]^n)$, where f is the fractional complex, $fmax$ corresponds to 100% complex formation, [Dps] is the concentration of the binding protein, n is the Hill coefficient and K_D represents the macroscopic apparent dissociation constant and is a measure of the affinity of the protein to DNA.

4.6.5.3 Protease protection assays

The proteolytic resistance of the DNA-Dps-WT complex against Proteinase K was studied incubating Dps-WT (0.5 to 13 µM) with supercoiled plasmid pUC19 (10 nM), for 10 min at room temperature, in 50 mM MOPS buffer pH 7.0, 50 mM NaCl, followed by incubation with Proteinase K (20 mg/mL) for 30 min at 37 °C. A control was prepared by first incubating Dps-WT with Proteinase K for 30 min at 37 °C, and then with pUC19 for 10 min at room temperature. Additionally, Dps-WT was substituted with BSA protein (13 µM). The ability of Dps-Q14E protein (13 µM) to protect DNA was also assessed.

4.6.5.3 DNA protection assays against hydroxyl radicals

DNA protection assays against hydroxyl radicals were performed using supercoiled plasmid pUC19 and Dps-WT, iron loaded Dps-WT (96 Fe/protein) and Dps-Q14E proteins, in 50 mM MOPS buffer pH 7.0, 50 mM NaCl. First, a DNA binding reaction was prepared with pUC19 (10 nM) and protein (13 μM). Afterwards, 2.04 mM $FeSO_4.7H_2O$ and 30.7 mM H_2O_2 were added to promote Fenton reactions during 8 min at room temperature. The reaction was stopped by addition of 10% SDS and incubation at 85 °C for 2 min. In parallel, and after the DNA binding reaction, Proteinase K (20 mg/mL) was added to the DNA−Dps-WT mixture (for 30 min at 37 °C), following reaction with ferrous sulphate and H_2O_2. As before, all reactions were analysed by electrophoresis in agarose gel.

4.6.5.4 Synchrotron Radiation Circular Dichroism (SRCD)

The synchrotron radiation SRCD spectra of Dps-WT protein were collected at the SRCD facility at the AU-CD beamline on the synchrotron radiation facility ASTRID2, at ISA, Department of Physics and Astronomy, Aarhus University, Denmark. The light passed from the synchrotron through a CaF_2 window into a CD instrument purged with nitrogen. In the instrument, a Rochon polarizer initially ensures linear polarization which is then converted into alternating circularly polarized light with a photo-elastic modulator (PEM, model I/ CF50, Hinds, USA) operating at ~50 kHz. The signal was detected with a photo-multiplier tube (PMT, 9402B, Electron Tubes, UK). Camphor-sulfonic acid served as a calibration material for the instrument. The spectra were recorded with 1 nm steps and a dwell time of 2.1 s per step in triplicate, using a nominally 0.01 cm quartz cells (SUPRASIL, Hellma GmbH, Germany), for the wavelength range of 170–280 nm. The actual pathlength of the cell was determined via an interference technique [180] to be 0.01008 cm. Protein samples were in 50 mM MOPS pH 7.0, 50 mM NaF at a final protein concentration of 1 mg/mL. Buffer spectra were used as baseline and subtracted from the protein spectra. Temperature scans were acquired from 5 to 85 °C using the following equilibrium times in between each increment: 5 minutes from 5 to 40 °C, 2.5 minutes from 40 °C to 65 °C, 5 minutes from 65 to 90 °C. Triplicate spectra at 25 °C were obtained before and after the temperature scan.

Two types of observation chambers were used: i) a standard, dual Peltier chamber for which samples are mounted vertically; and ii) a periscope chamber that features a horizontal mounting position. The use of the latter eliminates or reduces the effect of displacement due to aggregation/precipitation of protein samples or buffer density changes. In order to calculate the α-helix and β-sheet content in the analysed secondary protein structure, suitable spectra were fitted using the Dichroweb online server [98][99] The fitting was conducted using the CDSSTR method and SP175 reference data set [99] CD data was acquired at the BioLab, Biological and

Chemical Analysis Facility, UCIBIO-LAQV/Requimte, FCT NOVA, Portugal, using an Applied Photophysics Chirascan™ qCD, and a UXL-150S Xenon short arc lamp, Hamamatsu. Spectra were recorded in triplicate with 1 nm steps and a dwell time of 2.7 s per step using the same pathlength calibrated 0.01008 cm quartz cells as in the SRCD measurements.

4.6.5.4 The periscope chamber for CD measurements

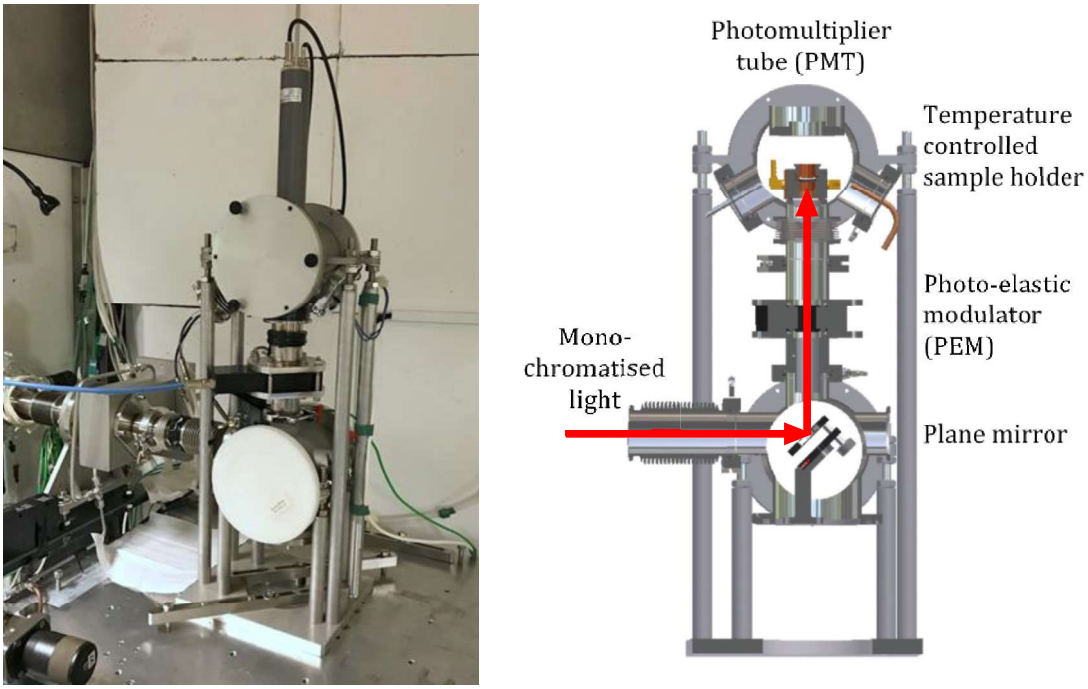

Figure 4-15: Photograph showing the AU-CD periscope chamber in place on the beam line (left) and CAD drawing of the periscope chamber (right).

Typically, samples are placed in a CD spectrometer with the light path horizontally interacting with the windows of the cell containing the sample mounted vertically ***Figure 4-15***. Some of the measurements carried out described herein utilised a periscope chamber, which features a horizontal sample mounting position, reducing the effect of possible precipitation of aggregated samples in liquids. Monochromatised light from the AU-CD beam line passes through a Rochon polarizer, before being reflected upward by a plane MgF_2 protected aluminium mirror. The light then passes a Photo Elastic Modulator (PEM) transforming the horizontally polarized light into alternating left and right handed circular polarized light, before entering the sample chamber. Transmitted light is detected using a UV enhanced photomultiplier tube (PMT) from Electron Tubes Ltd. type 9406B. The sample is placed in a copper holder surrounded by a copper block, which can be heated or cooled with two Peltier elements. The sample temperature may be set between 5 and 85°C.

4.6.5.5 Dynamic light scattering (DLS)

DLS measurements were performed using a HORIBA SZ-100 equipped with a 10 mW 532 nm laser at a detector angle of 90º. Dps-WT samples were prepared at 1 mg/mL in 200 mM MOPS pH 7, 200 mM NaCl and centrifuged for 15 min at 14000 g before each measurement. Temperature scans from 25 to 60 °C were performed by measuring the Z-average at least three times at each temperature or repeatedly until three stable readings were obtained. Raw data results were analysed using the equipment built-in software assuming a standard monodisperse form of distribution, using a particle refractive index of 1.6 (organic sample) and water settings as dispersion medium, with refractive index of 1.333 and viscosity given by the following equation:

$$\eta = 2.633\text{x}10^{-8} \cdot T^4 - 3.610\text{x}10^{-5} \cdot T^3 + 1.863\text{x}10^{-2} \cdot T^2 - 4.293 \cdot T + 3.736\text{x}10^{2}$$

4.6.5.6 Results

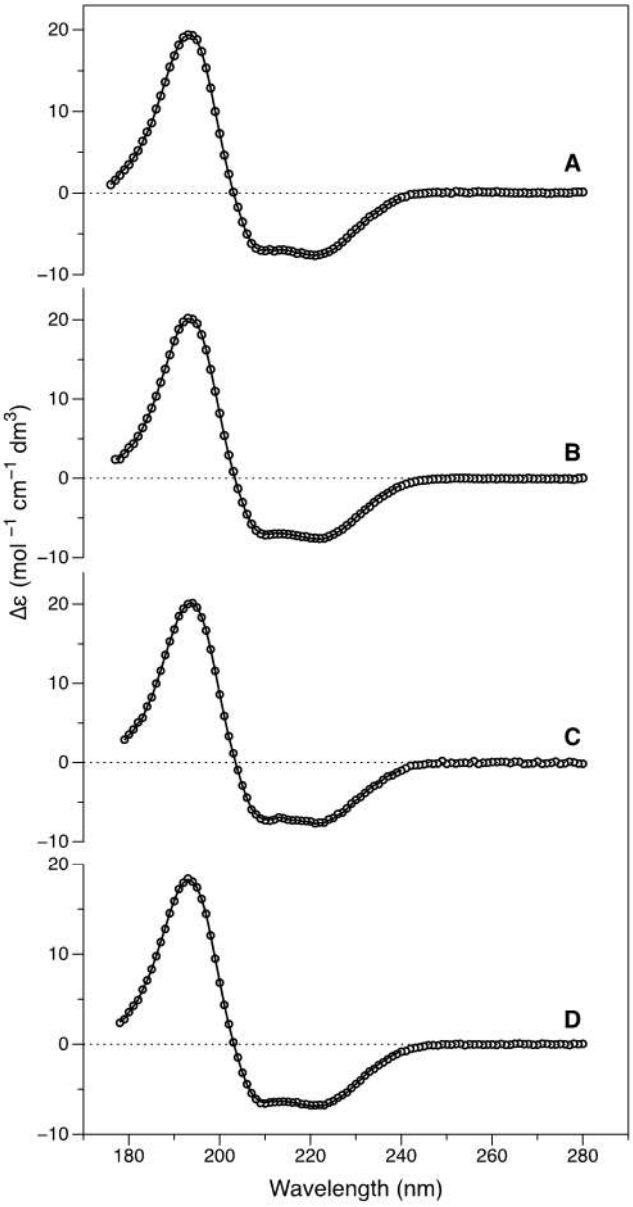

Figure 4-16: SRCD spectra obtained at 25 °C for Dps-WT (A); Dps-Q14E variant (B); Holo form with 96 Fe/Dps-WT protein (C) and Dps-WT (4 µM) incubated with supercoiled pUC19 (125 nM) (D). Solid lines represent spectral reconstructed data from Dichroweb analysis

Figure 4-16 shows the SRCD data obtained for the apo form of Dps-WT (panel A), Dps-Q14E variant (panel B) , holo form of Dps-WT (panel C) and for the apo form of Dps-WT incubated with pUC19 (panel D). Solid lines on top of experimental data were obtained by spectral reconstruction using results from Dichroweb analysis and are in good agreement with experimental results. For secondary structures percentages and results interpretations see main document (***Table 4-1***)

5

Conclusions and Future work

The work presented in this Thesis focused the characterization of Dps protein from *Marinobacter hydrocarbonoclasticus*, a mini-ferritin, in order to make use of the system for biotechnological applications.

To achieve this goal, the oligomeric state of Dps WT was firstly studied in different ionic strengths. The role played by the N-terminal tail of the protein, was also investigated as described in Chapter 2. By comparison with the WT Dps data obtained for the deletion mutant Δ15 have unequivocally shown that the assembly into dodecameric structure is dependent of the N-terminal tail, and also that other oligomeric forms can be tuned by the ionic strength of the buffer. It is therefore important to consider this dynamic equilibrium of the protein when one plans to use the system for biotechnological applications, such as encapsulation of compounds in the interior of the protein nanocage.

In Chapter 3 two different possible ways to use Dps in biotechnological applications are shown. An easy methodology for the green synthesis of nanoparticles inside the protein is presented. It was shown that different types of metal sulphides nanoparticles can be obtained. This methodology makes use of the physiological role of Dps iron and mineralization mechanism to achieve different materials without hindering the secondary structure of the protein. The possibility to employ Dps as a building block for supramolecular polymer synthesis with enhanced iron oxidation properties was also achieved. A point mutant of Dps wherein a threonine residue was replaced by a cysteine in position 10 enabled the possibility to successfully apply a click chemistry approach to functionalize the protein, which resulted in protein chains composed of Dps with over 1 μm in length. These presented enhanced physiological activity of iron oxidation. This is a proof of concept that Dps can be used as a building block for the synthesis of new material whilst maintaining its enzymatic activity.

The DNA binding activity of Dps was examined in Chapter 4, wherein DNA origami triangles were used to study the binding activity of Dps WT to either ssDNA or dsDNA strands. Even though it was shown that pH values and ionic strength influences the binding activity of Dps to DNA origami, it was not possible to show that the binding of Dps WT occurred to the strands of interest. However, it was demonstrated that a single alteration, a point mutation replacing glutamine-14 residue by a glutamate in the N-terminal tail is of importance to the N-terminal extension pI and therefore to the DNA binding process, which in this case abolished the DNA binding activity of the protein. It was also shown that the holo form of the protein is not able to bind DNA, suggesting subtle alterations of the N-terminal tail upon iron oxidation and mineralization.

As future work several approaches can be taken. The oligomeric state of Dps WT should be investigated in a wider variety of buffers to expand the realm of possibilities of the system for

encapsulation of compounds. This would increase the versability and degree of control that one has over of the system. The range of techniques that were employed in this work to access the oligomeric state of Dps in the different ionic strength conditions can be used as a template for future work.

It is also necessary to further study the mechanism used by the protein to synthesise nanoparticles in its interior. It is still not clear how the system is able to mediate the nanomaterial growth inside its nanocage. The chemical composition and structural characterization of the mineral that was synthesized inside the protein need further investigation. This could be an important step to make use of this protein not only for the nanotechnology field but also for the development of future bioremediation methodologies. The number of different materials that were able to be attained inside the protein through the simple synthesis procedure demonstrated in this thesis, allows for the possibility to employ Dps with nanoparticles in a applied context.

The supramolecular Dps polymer obtained is only the first step in the development of these protein-based new materials. The new protein fibres should be further characterised to infer about their potential use as hydrogels. Further work could be conducted with different linkers in order to access if different structures could be synthesized beside nano/microfibers. Another possible approach is the use of this new protein polymer for nanoparticle synthesis, or to take advantage of the inherent Dps DNA binding activity and studying it in this new polymer state.

Regarding the DNA-binding activity of Dps, additional experiments should be conducted with DNA origamis. Other DNA origami conformations could be employed such as squares or tubes to simplify the experiments. The binding of the Dps to the origami could be analysed by SEC or SAXS. Now that the N-terminal tail of Dps protein from *M. hydrocarbonoclasticus* is known to influence the DNA binding activity, different variations of this fragment could also be studied to better understand the mechanism upon which the DNA binding between DNA structures and Dps occurs.

Bibliography

[1] P. Arosio, L. Elia, and M. Poli, "Ferritin, cellular iron storage and regulation" *IUBMB Life*, 69, 414–422, 2017.

[2] W. R. Hagen, P.-L. Hagedoorn, and K. Honarmand Ebrahimi, "The workings of ferritin: a crossroad of opinions" *Metallomics*, 9, 595–605, 2017.

[3] S. C. Andrews, "The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor" *Biochim. Biophys. Acta-General Subj.*, 1800, 691–705, 2010.

[4] R. R. Crichton and J. Declercq, "Biochimica et Biophysica Acta X-ray structures of ferritins and related proteins" *BBA - Gen. Subj.*, 1800, 706–718, 2010.

[5] D. Penas, A. S. Pereira, and P. Tavares, "Direct evidence for ferrous ion oxidation and incorporation in the absence of oxidants by Dps from *Marinobacter hydrocarbonoclasticus*" *Angew. Chemie Int. Ed.*, 58, 1013–1018, 2019.

[6] E. F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, "UCSF Chimera - A visualization system for exploratory research and analysis" *J. Comput. Chem.*, 25, 1605–1612, 2004.

[7] H. J. H. Fenton, "Oxidation of Tartaric Acid in presence of Iron" *J. Chem. Soc. Trans.*, 65, 899–910, 1884.

[8] F. Haber and J. Weiss, "The catalytic decomposition of hydrogen peroxide by iron salts" *Proc. R. Soc. Lond. A*, 147, 332–351, 1932.

[9] J. Imlay, S. Chin, and S. Linn, "Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro" *Science*, 240, 640–642, 1988.

[10] U. S. Srinivas, B. W. Q. Tan, B. A. Vellayappan, and A. D. Jeyasekharan, "ROS and the DNA damage response in cancer" *Redox Biol.*, 25, 101084, 2019.

[11] M. Almiron, A. J. Link, D. Furlong, and R. Kolter, "A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*" *Genes Dev.*, 2646–2654, 1992.

[12] T. Haikarainen and A. C. Papageorgiou, "Dps-like proteins: Structural and functional insights into a versatile protein family" *Cell. Mol. Life Sci.*, 67, 341–351, 2010.

[13] J. M. H. R.A. Grant, D.J. Filman, S.E. Finkel, R. Kolter, "The crystal structure of Dps, a ferritin homolog that binds and protects DNA" *Nat. Struct. Biol.*, 5, 294–303, 1998.

[14] A. Ilari, S. Stefanini, E. Chiancone, and D. Tsernoglou, "The dodecameric ferritin from *Listeria innocua* contains a novel intersubunit iron-binding site" *Nat. Struct. Biol.*, 7, 38–43, 2000.

[15] G. Zhao, F. Bou-abdallah, P. Arosio, S. Levi, C. Janus-chandler, and N. D. Chasteen, "Multiple pathways for mineral core formation in mammalian apoferritin . The role of hydrogen peroxide" *Biochemistry*, 42, 3142–3150, 2003.

[16] E. Chiancone and P. Ceci, "The multifaceted capacity of Dps proteins to combat bacterial stress conditions: Detoxification of iron and hydrogen peroxide and DNA binding" *Biochim. Biophys. Acta - Gen. Subj.*, 1800, 798–805, 2010.

[17] S. Roy, R. Saraswathi, S. Gupta, K. Sekar, D. Chatterji, and M. Vijayan, "Role of N and C-terminal tails in DNA binding and assembly in Dps: Structural studies of *Mycobacterium smegmatis* Dps deletion mutants" *J. Mol. Biol.*, 370, 752–767, 2007.

[18] B. Ren, G. Tibbelin, T. Kajino, O. Asami, and R. Ladenstein, "The Multi-layered structure

of Dps with a novel di-nuclear ferroxidase center" *J. Mol. Biol.*, 329, 467–477, 2003.

[19] G. Bellapadrona, S. Stefanini, C. Zamparelli, E. C. Theil, and E. Chiancone, "Iron translocation into and out of *Listeria innocua* Dps and size distribution of the protein-enclosed nanomineral are modulated by the electrostatic gradient at the 3-fold 'Ferritin-like' pores" *J. Biol. Chem.*, 284, 19101–19109, 2009.

[20] Y. Zhang, J. Fu, S. Y. Chee, E. X. W. Ang, and B. P. Orner, "Rational disruption of the oligomerization of the mini-ferritin *E. coli* DPS through protein-protein interface mutation" *Protein Sci.*, 20, 1907–1917, 2011.

[21] G. Bhattacharyya and A. Grove, "The N-terminal extensions of *Deinococcus radiodurans* Dps-1 mediate DNA major groove interactions as well as assembly of the dodecamer" *J. Biol. Chem.*, 282, 11921–11930, 2007.

[22] A. Grove and S. P. Wilkinson, "Differential DNA binding and protection by dimeric and dodecameric forms of the ferritin homolog Dps from *Deinococcus radiodurans*" *J. Mol. Biol.*, 347, 495–508, 2005.

[23] S. Antipov S. Turishchev, Y. Purtov, U. Shvyreva, A. Sinelnikov, Y. Semov, E. Preobrazhenskaya, A. Berezhnoy, N. Shusharina, N. Novolokina, V. Vakhtel, V. Artyukhov, O. Ozoline, "The oligomeric form of the *Escherichia coli* dps protein depends on the availability of iron ions" *Molecules*, 22, 1–13, 2017.

[24] P. Ceci, A. Ilari, E. Falvo, L. Giangiacomo, and E. Chiancone, "Reassessment of Protein Stability, DNA Binding, and Protection of *Mycobacterium smegmatis* Dps" *J. Biol. Chem.*, 280, 34776–34785, 2005.

[25] X. Yang E. Chiancone, S. Stefanini, A. Ilari, N.D. Chasteen, "Iron oxidation and hydrolysis reactions of a novel ferritin from *Listeria innocua*" *Biochem. J.*, 786, 783–786, 2000.

[26] G. Zhao, P. Ceci, A. Ilari, L. Giangiacomo, T.M. Laue, E. Chiancone, N. Dennis Chasteen, "Iron and hydrogen peroxide detoxification properties of DNA-binding protein from starved cells" *J. Biol. Chem.*, 277, 27689–27696, 2002.

[27] G. Zanotti, E. Papinutto, W.G. Dundon, R. Battistutta, M. Seveso, G. Del Giudice, R. Rappuoli, C. Montecucco, "Structure of the neutrophil-activating protein from *Helicobacter pylori*" *J. Mol. Biol.*, 323, 125–130, 2002.

[28] S.-G. Kim, G. Bhattacharyya, A. Grove, and Y.-H. Lee, "Crystal structure of Dps-1, a functionally distinct Dps protein from *Deinococcus radiodurans*" *J. Mol. Biol.*, 361, 105–114, 2006.

[29] P. Ceci, A. Ilari, E. Falvo, and E. Chiancone, "The Dps protein of *Agrobacterium tumefaciens* does not bind to DNA but protects it toward oxidative cleavage" *J. Biol. Chem.*, 278, 20319–20326, 2003.

[30] S. Franceschini, P. Ceci, F. Alaleona, E. Chiancone, and A. Ilari, "Antioxidant Dps protein from the thermophilic cyanobacterium *Thermosynechococcus elongatus*" *FEBS J.*, 273, 4913–4928, 2006.

[31] K. Zeth, "Dps biomineralizing proteins: multifunctional architects of nature" *Biochem. J.*, 445, 297–311, 2012.

[32] Y. Yamamoto, L. B. Poole, R. R. Hantgan, and Y. Kamio, "An iron-binding protein, Dpr, from *Streptococcus mutans* prevents iron-dependent hydroxyl radical formation *in vitro*" *J. Bacteriol.*, 184, 2931–2939, 2002.

[33] D. Penas, A. S. Pereira, and P. Tavares, "Direct evidence for ferrous ion oxidation and

incorporation in the absence of oxidants by Dps from *Marinobacter hydrocarbonoclasticus*" *Angew. Chemie - Int. Ed.*, 58, 1013–1018, 2019.

[34] P. Chen, E. De Meulenaere, D. D. Deheyn, and P. R. Bandaru, "Iron redox pathway revealed in ferritin via electron transfer analysis" *Sci. Rep.*, 10, 1–10, 2020.

[35] G. Bellapadrona, S. Stefanini, C. Zamparelli, E. C. Theil, and E. Chiancone, "Iron translocation into and out of *Listeria innocua* Dps and size distribution of the protein-enclosed nanomineral are modulated by the electrostatic gradient at the 3-fold 'Ferritin-like' pores" *J. Biol. Chem.*, 284, 19101–19109, 2009.

[36] A. Ilari, P. Ceci, D. Ferrari, G. L. Rossi, and E. Chiancone, "Iron incorporation into *Escherichia coli* Dps gives rise to a Ferritin-like microcrystalline core" *J. Biol. Chem.*, 277, 37619–37623, 2002.

[37] T. J. Stillman M. Upadhyay, V. A. Norte, S. E. Sedelnikova, M. Carradus, S. Tzokov, P. A. Bullough, C. A. Shearman, M. J. Gasson, C. H. Williams, P. J. Artymiuk, J. Green, "The crystal structures of *Lactococcus lactis* MG1363 Dps proteins reveal the presence of an N-terminal helix that is required for DNA binding" *Mol. Microbiol.*, 57, 1101–1112, 2005.

[38] M. D. Hitchings , P. Towsend, E. Pohl, P. D. Facey, D. H. Jones, P. J. Dyson, R. del Sol, "A tale of tails: deciphering the contribution of terminal tails to the biochemical properties of two Dps proteins from *Streptomyces coelicolor*" *Cell. Mol. Life Sci.*, 71, 4911–4926, 2014.

[39] S. S. Antipov M. N. Tutukina, E. V. Preobrazhenskaya, F. A. Kondrashov, M. V. Patrushev, S. V. Toshchakov, I. Dominova, U. S. Shvyrev, V. V. Vrublevskaya, O. S. Morenkov, N. A. Sukharicheva, V. V. Panyukov, O. N. Ozoline, "The nucleoid protein Dps binds genomic DNA of *Escherichia coli* in a non-random manner" *PLoS One*, 12, 1–27, 2017.

[40] R. A. Grant, D. J. Filman, S. E. Finkel, R. Kolter, and J. M. Hogle, "The crystal structure of Dps, a ferritin homolog that binds and protects DNA" *Nat. Struct. Biol.*, 5, 294–303, 1998.

[41] P. Ceci, L. Mangiarotti, C. Rivetti, and E. Chiancone, "The neutrophil-activating Dps protein of *Helicobacter pylori*, HP-NAP, adopts a mechanism different from *Escherichia coli* Dps to bind and condense DNA" *Nucleic Acids Res.*, 35, 2247–2256, 2007.

[42] P. Ceci, S. Cellai, E. Falvo, C. Rivetti, G. L. Rossi, and E. Chiancone, "DNA condensation and self-aggregation of *Escherichia coli* Dps are coupled phenomena related to the properties of the N-terminus" *Nucleic Acids Res.*, 32, 5935–5944, 2004.

[43] D. He and J. Marles-Wright, "Ferritin family proteins and their use in bionanotechnology" *N. Biotechnol.*, 32, 651–657, 2015.

[44] S.E. Kim, K.Y. Ahn, J.S. Park, K.R. Kim, K.E. Lee, S.S. Han, J. Lee,, "Fluorescent ferritin nanoparticles and application to the aptamer sensor" *Anal. Chem.*, 83, 5834–5843, 2011.

[45] H. Yoshimura, "Protein-assisted nanoparticle synthesis" *Colloids Surfaces A Physicochem. Eng. Asp.*, 282–283, 464–470, 2006.

[46] G. Jutz, P. Van Rijn, B. Santos Miranda, and A. Böker, "Ferritin: A versatile building block for bionanotechnology" *Chem. Rev.*, 115, 1653–1701, 2015.

[47] A. M. Sevcenco, M. Paravidino, J. S. Vrouwenvelder, H. T. Wolterbeek, M. C. M. van Loosdrecht, and W. R. Hagen, "Phosphate and arsenate removal efficiency by thermostable ferritin enzyme from *Pyrococcus furiosus* using radioisotopes" *Water Res.*, 76, 181–186, 2015.

[48] L. Chen G. Bai, S. Yang, R. Yang, G. Zhao, C. Xu, W. Leung, "Encapsulation of curcumin in recombinant human H-chain ferritin increases its water-solubility and stability" *Food Res. Int.*, 62, 1147–1153, 2014.

[49] J. Zang, H. Chen, G. Zhao, F. Wang, and F. Ren, "Ferritin cage for encapsulation and delivery of bioactive nutrients: From structure, property to applications" *Crit. Rev. Food Sci. Nutr.*, 57, 3673–3683, 2017.

[50] M. Okuda, Y. Suzumoto, K. Iwahori, S. Kang, M. Uchida, T. Douglas, I. Yamashita, "Bio-templated CdSe nanoparticle synthesis in a cage shaped protein, *Listeria*-Dps, and their two dimensional ordered array self-assembly" *Chem. Commun.*, 46, 8797–8799, 2010.

[51] M. Kim Y. Rho, K. S. Jin, B. Ahn, S. Jung, H. Kim, M. Ree, "pH-dependent structures of ferritin and apoferritin in solution: Disassembly and reassembly" *Biomacromolecules*, 12, 1629–1640, 2011.

[52] B. Chiou and J. R. Connor, "Emerging and dynamic biomedical uses of ferritin" *Pharmaceuticals*, 11, e124-141, 2018.

[53] S. Mazzucchelli, M. Truffi, F. Baccarini, M. Beretta, L. Sorrentino, M. Bellini, M.A. Rizzuto, R. Ottria, A. Ravelli, P. Ciuffreda, D. Prosperi, F. Corsi, "H-Ferritin-nanocaged olaparib: a promising choice for both BRCA-mutated and sporadic triple negative breast cancer" *Sci. Rep.*, 7, 7505, 2017.

[54] M. Liang, K. Fan, M. Zhou, D. Duan, J. Zheng, D. Yang, J. Feng, X. Yan, "H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection" *Proc. Natl. Acad. Sci.*, 111, 14900–14905, 2014.

[55] X. Li, L. Qiu, P. Zhu, X. Tao, T. Imanaka, J. Zhao, Y. Huang, Y. Tu, X. Cao, "Epidermal Growth Factor-Ferritin H-Chain Protein Nanoparticles for Tumor Active Targeting" *Small*, 8, 2505–2514, 2012.

[56] M. Kanekiyo, C.J. Wei, H.M. Yassine, P.M. McTamney, J.C. Boyington, J.R.R. Whittle, S.S. Rao, W.P. Kong, L. Wang, G.J. Nabel, "Self-assembling influenza nanoparticle vaccines elicit broadly neutralizing H1N1 antibodies" *Nature*, 499, 102–106, 2013.

[57] A. Prastaro, P. Ceci, E. Chiancone, A. Boffi, R. Cirilli, M. Colone, G. Fabrizi, A. Stringaro, S. Cacchi, "Suzuki-Miyaura cross-coupling catalyzed by protein-stabilized palladium nanoparticles under aerobic conditions in water: application to a one-pot chemoenzymatic enantioselective synthesis of chiral biaryl alcohols" *Green Chem.*, 11, 1929, 2009.

[58] I. Yamashita, "Biosupramolecules for nano-devices: biomineralization of nanoparticles and their applications" *J. Mater. Chem.*, 18, 3813, 2008.

[59] G. Jutz and A. Böker, "Bionanoparticles as functional macromolecular building blocks - A new class of nanomaterials" *Polymer*., 52, 211–232, 2011.

[60] A. de la Escosura, M. Verwegen, F.D. Sikkema, M. Comellas-Aragonès, A. Kirilyuk, T. Rasing, R.J.M. Nolte, J.J.L.M. Cornelissen, "Viral capsids as templates for the production of monodisperse Prussian blue nanoparticles" *Chem. Commun.*, 13, 1542–1544, 2008.

[61] T. Ueno, M. Suzuki, T. Goto, T. Matsumoto, K. Nagayama, and Y. Watanabe, "Size-selective olefin hydrogenation by a Pd nanocluster provided in an apo-ferritin cage" *Angew. Chemie - Int. Ed.*, 43, 2527–2530, 2004.

[62] T. Douglas, D. P. E. Dickson, S. Betteridge, J. Charnock, C. D. Garner, and S. Mann, "Synthesis and structure of an iron(III) sulfide-ferritin bioinorganic nanocomposite" *Science.*, 269, 54–57, 1995.

[63] S.H. Choi, J.-W. Kim, S.-H. Chu, Y. Park, G.C. King, P.T. Lillehei, S.-J. Kim, J.R. Elliott, "Ferritin-templated quantum-dots for quantum logic gates" *Proc. SPIE - Int. Soc. Opt. Eng.*, 5763, 213–232, 2005.

[64] M. Okuda, K. Iwahori, I. Yamashita, and H. Yoshimura, "Fabrication of nickel and chromium nanoparticles using the protein cage of apoferritin" *Biotechnol. Bioeng.*, 84, 187–194, 2003.

[65] S. A. Shamsudin, H. Hasegawa, M. Takenaka, and S. Kenji, "Aligning CdS quantum dots in apo-ferritin protein and PS-b-P2VP organic templates" *Adv. Mater. Res.*, 832, 675–680, 2013.

[66] K. Iwahori, M. Yamane, S. Fujita, and I. Yamashita, "Synthesizing CdSe nanoparticles by using a low concentration of cadmium ions and the apoferritin protein cage of marine pennate diatoms" *Mater. Lett.*, 160, 154–157, 2015.

[67] T. Wang, H. W. Choi, W.-J. Kim, J. S. Kim, and S. J. Park, "Two-dimensional array of ZnSe-ferritin nanodots as a sensor media for gamma-aminobutyric acid" *Mol. Cryst. Liq. Cryst.*, 519, 27–35, 2010.

[68] T.D. Bradshaw, M. Junor, A. Patanè, P. Clarke, N.R. Thomas, M. Li, S. Mann, L. Turyanska, "Apoferritin-encapsulated PbS quantum dots significantly inhibit growth of colorectal carcinoma cells" *J. Mater. Chem. B*, 1, 6254–6260, 2013.

[69] K. Yoshizawa, K. Iwahori, K. Sugimoto, and I. Yamashita, "Fabrication of gold sulfide nanoparticles using the protein cage of apoferritin" *Chem. Lett.*, 35, 1192–1193, 2006.

[70] K.R. Hansen, J.R. Peterson, A. Perego, M. Shelley, C.R. Olsen, L.D. Perez, H.L. Hogg, R.K. Watt, J.S. Colton, "Lead sulfide quantum dots inside ferritin: synthesis and application to photovoltaics" *Appl. Nanosci.*, 8, 1687–1699, 2018.

[71] K. Iwahori, T. Enomoto, H. Furusho, A. Miura, K. Nishio, Y. Mishima, I. Yamashita, "Cadmium sulfide nanoparticle synthesis in Dps protein from *Listeria innocua*" *Chem. Mater.*, 19, 3105–3111, 2007.

[72] M. Grzelczak, J. Vermant, E. M. Furst, and L. M. Liz-Marzán, "Directed self-assembly of nanoparticles" *ACS Nano*, 4, 3591–3605, 2010.

[73] C. M. Soto and B. R. Ratna, "Virus hybrids as nanomaterials for biotechnology" *Curr. Opin. Biotechnol.*, 21, 426–438, 2010.

[74] R. Yang, L. Chen, T. Zhang, S. Yang, X. Leng, and G. Zhao, "Self-assembly of ferritin nanocages into linear chains induced by poly(α, l-lysine)" *Chem. Commun.*, 50, 481–483, 2014.

[75] Q. Luo, Z. Dong, C. Hou, and J. Liu, "Protein-based supramolecular polymers: progress and prospect" *Chem. Commun.*, 50, 9997, 2014.

[76] H. C. Kolb, M. G. Finn, and K. B. Sharpless, "Click chemistry: Diverse chemical function from a few good reactions" *Angew. Chemie - Int. Ed.*, 40, 2004–2021, 2001.

[77] B. Le Droumaguet and K. Velonia, "Click chemistry: A powerful tool to create polymer-based macromolecular chimeras" *Macromol. Rapid Commun.*, 29, 1073–1089, 2008.

[78] B. Liu, M. Ianosi-Irimie, and S. Thayumanavan, "Reversible Click Chemistry for ultrafast and quantitative formation of protein-polymer nanoassembly and intracellular protein delivery" *ACS Nano*, 13, 9408–9420, 2019.

[79] P. W. K. Rothemund, "Folding DNA to create nanoscale shapes and patterns" *Nature*, 440, 297–302, 2006.

[80] I. Bald and A. Keller, "Molecular processes studied at a single-molecule level using DNA origami nanostructures and atomic force microscopy" *Molecules*, 19, 13803–13823, 2014.

[81] "OpenStax College, Nucleic Acids." Available: http://cnx.org/content/m44403/latest/Figure_03_05_01.jpg. [Accessed: 02-Feb-2020].

[82] J. Prinz, B. Schreiber, L. Olejko, J. Oertel, J. Rackwitz, A. Keller, I. Bald, "DNA origami substrates for highly sensitive surface-enhanced raman scattering" *J. Phys. Chem. Lett.*, 4, 4140–4145, 2013.

[83] N. Wu, X. Zhou, D.M. Czajkowsky, M. Ye, D. Zeng, Y. Fu, C. Fan, J. Hu, B. Li, "In situ monitoring of single molecule binding reactions with time-lapse atomic force microscopy on functionalized DNA origami" *Nanoscale*, 3, 2481, 2011.

[84] I. H. Stein, C. Steinhauer, and P. Tinnefeld, "Single-molecule four-color FRET visualizes energy-transfer paths on DNA origami" *J. Am. Chem. Soc.*, 133, 4193–4195, 2011.

[85] S. H. Ko, G. M. Gallatin, and J. A. Liddle, "Nanomanufacturing with DNA origami: Factors affecting the kinetics and yield of quantum dot binding" *Adv. Funct. Mater.*, 22, 1015–1023, 2012.

[86] X. -c. Bai, T. G. Martin, S. H. W. Scheres, and H. Dietz, "Cryo-EM structure of a 3D DNA-origami object" *Proc. Natl. Acad. Sci.*, 109, 20012–20017, 2012.

[87] D. P. Allison, N. P. Mortensen, C. J. Sullivan, and M. J. Doktycz, "Atomic force microscopy of biological samples" *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology*, 2, 618–634, 2010.

[88] R. Jungmann, C. Steinhauer, M. Scheible, A. Kuzyk, P. Tinnefeld, and F. C. Simmel, "Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami" *Nano Lett.*, 10, 4756–4761, 2010.

[89] V. V Thacker, L.O. Herrmann, D.O. Sigle, T. Zhang, T. Liedl, J.J. Baumberg, U.F. Keyser, "DNA origami based assembly of gold nanoparticle dimers for surface-enhanced Raman scattering" *Nat. Commun.*, 5, 3448, 2014.

[90] A. Moeinian, F.N. Gür, J. Gonzalez-Torres, L. Zhou, V.D. Murugesan, A.D. Dashtestani, H. Guo, T.L. Schmidt, S. Strehle, "Highly localized SERS measurements using single silicon nanowires decorated with DNA origami-based SERS probe" *Nano Lett.*, 19, 1061–1066, 2019.

[91] A. Shaw, E. Benson, and B. Högberg, "Purification of functionalized DNA origami nanostructures" *ACS Nano*, 9, 4968–4975, 2015.

[92] S. Ramakrishnan, S. Subramaniam, A. F. Stewart, G. Grundmeier, and A. Keller, "Regular nanoscale protein patterns via directed adsorption through self-assembled DNA origami masks" *ACS Appl. Mater. Interfaces*, 8, 31239–31247, 2016.

[93] M. C. M. Guilherme, "Estudos mecanísticos e estruturais da oxidação e armazenamento de ferro por ferritinas rápidas", PhD Thesis, Universidade Nova de Lisboa, 2009.

[94] D. Penas, "Structural and Mechanistic Studies of Prokaryotic Ferritins", PhD Thesis, Universidade Nova de Lisboa, 2018.

[95] N. J. Greenfield, "Using circular dichroism spectra to estimate protein secondary structure" *Nat. Protoc.*, 1, 2876–2890, 2006.

[96] B. A. Wallace and R. W. Janes, "Synchrotron radiation circular dichroism (SRCD) spectroscopy: an enhanced method for examining protein conformations and protein interactions" *Biochem. Soc. Trans.*, 38, 861–73, 2010.

[97] L. Whitmore and B. A. Wallace, “Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases” *Biopolymers*, 89, 392–400, 2008.

[98] L. Whitmore and B. A. Wallace, “DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data” *Nucleic Acids Res.*, 32, W668-673, 2004.

[99] J. G. Lees, A. J. Miles, F. Wien, and B. A. Wallace, “A reference database for circular dichroism spectroscopy covering fold and secondary structure space” *Bioinformatics*, 22, 1955–1962, 2006.

[100] C.E. Blanchet, A. Spilotros, F. Schwemmer, M.A. Graewert, A. Kikhney, C.M. Jeffries, D. Franke, D. Mark, R. Zengerle, F. Cipriani, S. Fiedler, M. Roessle, D.I. Svergun, “Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY)” *J. Appl. Crystallogr.*, 48, 431–443, 2015.

[101] D. Franke, M. V. Petoukhov, P. V. Konarev, A. Panjkovich, A. Tuukkanen, H.D.T. Mertens, A.G. Kikhney, N.R. Hajizadeh, J.M. Franklin, C.M. Jeffries, D.I. Svergun, “ATSAS 2.8 : a comprehensive data analysis suite for small-angle scattering from macromolecular solutions” *J. Appl. Crystallogr.*, 50, 1212–1225, 2017.

[102] R. Chiaraluce, V. Consalvi, S. Cavallo, A. Ilari, S. Stefanini, and E. Chiancone, “The unusual dodecameric ferritin from *Listeria innocua* dissociates below pH 2.0” *Eur. J. Biochem.*, 267, 5733–5741, 2000.

[103] B.R. Lee, H.K. Ko, J.H. Ryu, K.Y. Ahn, Y.H. Lee, S.J. Oh, J.H. Na, T.W. Kim, Y. Byun, I.C. Kwon, K. Kim, J. Lee, “Engineered human ferritin nanoparticles for direct delivery of tumor antigens to lymph node and cancer immunotherapy” *Sci. Rep.*, 6, 1–12, 2016.

[104] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, “NIH Image to ImageJ: 25 years of image analysis” *Nat. Methods*, 9, 671–675, 2012.

[105] A. Kauko, S. Haataja, A. T. Pulliainen, J. Finne, and A. C. Papageorgiou, “Crystal structure of *Streptococcus suis* dps-like peroxide resistance protein Dpr: Implications for iron incorporation” *J. Mol. Biol.*, 338, 547–558, 2004.

[106] J. Pršek, D. Ozdín, and J. Sejkora, “Eclarite and associated Bi sulfosalts from the Brezno-Hviezda occurrence (Nízke Tatry Mts, Slovak Republic)” *Neues Jahrb. fur Mineral. Abhandlungen*, 185, 117–130, 2008.

[107] J. Donohue, “Introduction to crystal geometry (Buerger, Martin J.)” *J. Chem. Educ.*, 49, A192, 1972.

[108] S. A. T. Redfern, "Manual of mineralogy (After James D. Dana)", 21st Ed. by C. Klein and C. S. Hurlbut Jr, Wiley, New York, 30, ISBN 0471535605, 1995.

[109] S. Das and M. J. Hendry, “Application of Raman spectroscopy to identify iron minerals commonly found in mine wastes,” *Chem. Geol.*, 290, 101–108, 2011.

[110] B. Lafuente, R. T. Downs, H. Yang, and N. Stone, “The power of databases: the RRUFF project. In "Highlights in mineralogical crystallography", Edited by T. Armbruster and R.M. Danisi, Berlin, München Boston: De Gruyter, 30, 2015.

[111] J. Wang, D. L. Rousseau, H. M. Abu-Soud, and D. J. Stuehr, “Heme coordination of NO in NO synthase” *Proc. Natl. Acad. Sci. U. S. A.*, 91, 10512—10516, 1994.

[112] L. Ashton, V. L. Brewster, E. Correa, and R. Goodacre, “Detection of glycosylation and iron-binding protein modifications using Raman spectroscopy” *Analyst*, 142, 808–814,

2017.

[113] N. Howell and E. Li-Chan, “Elucidation of interactions of lysozyme with whey proteins by Raman spectroscopy” *Int. J. Food Sci. \& Technol.*, 31, 439–451, 1996.

[114] P. E. Prevelige, D. Thomas, K. L. Aubrey, S. A. Towse, and G. J. Thomas, “Subunit conformational changes accompanying bacteriophage P22 capsid maturation” *Biochemistry*, 32, 537–543, 1993.

[115] S. Vohník, C. Hanson, R. Tuma, J. A. Fuchs, C. Woodward, and G. J. J. Thomas, “Conformation, stability, and active-site cysteine titrations of *Escherichia coli* D26A thioredoxin probed by Raman spectroscopy” *Protein Sci.*, 7, 193–200, 1998.

[116] D. Diaz, A. Care, and A. Sunna, “Bioengineering strategies for protein-based nanoparticles” *Genes (Basel).*, 9, e370-399, 2018.

[117] M. Rother, M. G. Nussbaumer, K. Renggli, and N. Bruns, “Protein cages and synthetic polymers: A fruitful symbiosis for drug delivery applications, bionanotechnology and materials science” *Chem. Soc. Rev.*, 45, 6213–6249, 2016.

[118] H. Ren, S. Zhu, and G. Zheng, “Nanoreactor design based on self-assembling protein nanocages” *Int. J. Mol. Sci.*, 20, 592, 2019.

[119] S. Tetter and D. Hilvert, “Enzyme encapsulation by a ferritin cage” *Angew. Chemie - Int. Ed.*, 56, 14933–14936, 2017.

[120] M. Marguet, C. Bonduelle, and S. Lecommandoux, “Multicompartmentalized polymeric systems: Towards biomimetic cellular structure and function” *Chem. Soc. Rev.*, 42, 512–529, 2013.

[121] E. J. Lee, N. K. Lee, and I.-S. Kim, “Bioengineered protein-based nanocage for drug delivery” *Adv. Drug Deliv. Rev.*, 106, 157–171, 2016.

[122] P. Sahandi Zangabad, M. Karimi, F. Mehdizadeh, H. Malekzad, A. Ghasemi, S. Bahrami, H. Zare, M. Moghoofei, A. Hekmatmanesh, M.R. Hamblin, “Nanocaged platforms: modification, drug delivery and nanotoxicity. Opening synthetic cages to release the tiger” *Nanoscale*, 9, 1356–1392, 2017.

[123] R. R. Crichton and J.-P. Declercq, “X-ray structures of ferritins and related proteins” *Biochim. Biophys. Acta - Gen. Subj.*, 1800, 706–718, 2010.

[124] S. Bhaskar and S. Lim, “Engineering protein nanocages as carriers for biomedical applications” *NPG Asia Mater.*, 9, e371, 2017.

[125] J. G. Heddle, “Protein cages, rings and tubes: useful components of future nanodevices?” *Nanotechnol. Sci. Appl.*, 1, 67–78, 2008.

[126] B. Maity, S. Abe, and T. Ueno, “Observation of gold sub-nanocluster nucleation within a crystalline protein cage” *Nat. Commun.*, 8, 14820–14828, 2017.

[127] M.A. Kostiainen, P. Hiekkataipale, A. Laiho, V. Lemieux, J. Seitsonen, J. Ruokolainen, P. Ceci, “Electrostatic assembly of binary nanoparticle superlattices using protein cages” *Nat. Nanotechnol.*, 8, 52–56, 2013.

[128] C.C. Broomell, H. Birkedal, C.L.P. Oliveira, J.S. Pedersen, J.-A. Gertenbach, M. Young, T. Douglas, “Protein cage nanoparticles as secondary building units for the synthesis of 3-dimensional coordination polymers” *Soft Matter*, 6, 3167–3171, 2010.

[129] K. Zhou, H. Chen, S. Zhang, Y. Wang, and G. Zhao, “Disulfide-mediated reversible two-dimensional self-assembly of protein nanocages” *Chem. Commun.*, 55, 7510–7513, 2019.

[130] C. Gu, H. Chen, Y. Wang, T. Zhang, H. Wang, and G. Zhao, "Structural insight into binary protein metal–organic frameworks with ferritin nanocages as linkers and nickel clusters as nodes" *Chem. – A Eur. J.*, 26, 3016–3021, 2020.

[131] K. Iwahori, R. Takagi, N. Kishimoto, and I. Yamashita, "A size controlled synthesis of CuS nano-particles in the protein cage, apoferritin" *Mater. Lett.*, 65, 3245–3247, 2011.

[132] R. K. Watt, O. D. Petrucci, and T. Smith, "Ferritin as a model for developing 3rd generation nano architecture organic/inorganic hybrid photo catalysts for energy conversion" *Catal. Sci. Technol.*, 3, 3103–3110, 2013.

[133] F. Beuerle and B. Gole, "Covalent organic frameworks and cage compounds: design and applications of polymeric and discrete organic scaffolds" *Angew. Chemie - Int. Ed.*, 57, 4850–4878, 2018.

[134] N. Kobayashi, K. Inano, K. Sasahara, T. Sato, K. Miyazawa, T. Fukuma, M.H. Hecht, C. Song, K. Murata, R. Arai, "Self-assembling supramolecular nanostructures constructed from de novo extender protein nanobuilding blocks" *ACS Synth. Biol.*, 7, 1381–1394, 2018.

[135] Y. Wang and D. H.-C. Chou, "A thiol-ene Coupling approach to native peptide stapling and macrocyclization" *Angew. Chemie - Int. Ed.*, 54, 10931–10934, 2015.

[136] A. Massi and D. Nanni, "Thiol-yne coupling: Revisiting old concepts as a breakthrough for up-to-date applications" *Org. Biomol. Chem.*, 10, 3791–3807, 2012.

[137] E. M. Valkevich, R. G. Guenette, N. A. Sanchez, Y. Chen, Y. Ge, and E. R. Strieter, "Forging isopeptide bonds using thiol–ene chemistry: site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases" *J. Am. Chem. Soc.*, 134, 6916–6919, 2012.

[138] Z. Z. Brown, M. M. Müller, S. U. Jain, C. D. Allis, P. W. Lewis, and T. W. Muir, "Strategy for 'detoxification' of a cancer-derived histone mutant based on mapping its interaction with the methyltransferase PRC2" *J. Am. Chem. Soc.*, 136, 13498–13501, 2014.

[139] M. Bertoni, F. Kiefer, M. Biasini, L. Bordoli, and T. Schwede, "Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology" *Sci. Rep.*, 7, 10480–10494, 2017.

[140] P. Benkert, M. Biasini, and T. Schwede, "Toward the estimation of the absolute quality of individual protein structure models" *Bioinformatics*, 27, 343–350, 2011.

[141] N. Guex, M. C. Peitsch, and T. Schwede, "Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective" *Electrophoresis*, 30, S162–S173, 2009.

[142] S. Bienert, A. Waterhouse, T.A.P. de Beer, G. Tauriello, G. Studer, L. Bordoli, T. Schwede, "The SWISS-MODEL Repository—new features and functionality" *Nucleic Acids Res.*, 45, D313–D319, 2017.

[143] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, "SWISS-MODEL: homology modelling of protein structures and complexes" *Nucleic Acids Res.*, 46, W296–W303, 2018.

[144] R. Hussain, T. Jávorfi, C. S. Hughes, and G. Siligardi, "Circular dichroism and synchrotron radiation circular dichroism applications to biomaterials" in "Radiation in Bioanalysis", Edited by A.S. Pereira, P. Tavares, and P. Limão-Vieira, Spinger, 147–172, 2019,

[145] L. Whitmore and B. A. Wallace, "Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases" *Biopolymers*, 89, 392–400, 2008.

[146] N. Sreerama and R. W. Woody, "Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set" *Anal. Biochem.*, 287, 252–260, 2000.

[147] B. Woollett, L. Whitmore, R. W. Janes, and B. A. Wallace, "ValiDichro: a website for validating and quality control of protein circular dichroism spectra" *Nucleic Acids Res.*, 41, W417-W421, 2013.

[148] W. Kabsch and C. Sander, "Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features" *Biopolymers*, 22, 2577–2637, 1983.

[149] O. Miyawaki, M. Dozen, and K. Hirota, "Cooperative hydration effect causes thermal unfolding of proteins and water activity plays a key role in protein stability in solutions" *J. Biosci. Bioeng.*, 122, 203–207, 2016.

[150] R. A. Grant, D. J. Filman, S. E. Finkel, R. Kolter, and J. M. Hogle, "The crystal structure of Dps, a ferritin homolog that binds and protects DNA" *Nat. Struct. Biol.*, 5, 294–303, 1998.

[151] K. Sugimoto, S. Kanamaru, K. Iwasaki, F. Arisaka, and I. Yamashita, "Construction of a ball-and-spike protein supramolecule" *Angew. Chemie - Int. Ed.*, 45, 2725–2728, 2006.

[152] R. Wang, J. Li, X. Li, J. Guo, J. Liu, and H. Li, "Engineering protein polymers of ultrahigh molecular weight via supramolecular polymerization: towards mimicking the giant muscle protein titin" *Chem. Sci.*, 10, 9277–9284, 2019.

[153] C. Hou, J. Li, L. Zhao, W. Zhang, Q. Luo, Z. Dong, J. Xu, J. Liu, "Construction of protein nanowires through cucurbit[8]uril-based highly specific host-guest interactions: an approach to the assembly of functional proteins" *Angew. Chemie Int. Ed.*, 52, 5590–5593, 2013.

[154] C. G. Timóteo, M. Guilherme, D. Penas, F. Folgosa, P. Tavares, and A. S. Pereira, "*Desulfovibrio vulgaris* bacterioferritin uses H_2O_2 as a co-substrate for iron oxidation and reveals DPS-like DNA protection and binding activities" *Biochem. J.*, 446, 125–133, 2012.

[155] A.S. Pereira, P. Tavares, S.G. Lloyd, D. Danger, D.E. Edmondson, E.C. Theil, B.H. Huynh, "Rapid and parallel formation of Fe^{3+} multimers, including a trimer, during H-type subunit ferritin mineralization" *Biochemistry*, 36, 7917–7927, 1997.

[156] B. D. Fairbanks, M. P. Schwartz, C. N. Bowman, and K. S. Anseth, "Photoinitiated polymerization of PEG-diacrylate with lithium phenyl-2,4,6-trimethylbenzoylphosphinate: polymerization rate and cytocompatibility" *Biomaterials*, 30, 6702–6707, 2009.

[157] A. A. Abramov, A. V Anisimov, and A. A. Bobyleva, "Synthesis of oxathiacrown compounds by the reactions of sulfur dichloride with unsaturated compounds and their extraction characteristics. (Review)" *Chem. Heterocycl. Compd.*, 38, 261–273, 2002.

[158] C. K. Riener, G. Kada, and H. J. Gruber, "Quick measurement of protein sulfhydryls with Ellman's reagent and with 4,4′-dithiodipyridine" *Anal. Bioanal. Chem.*, 373, 266–276, 2002.

[159] U. K. Laemmli, "Cleavage of structural proteins during the assembly of the head of bacteriophage T4" *Nature*, 227, 680–685, 1970.

[160] E. K. Koepf, H. M. Petrassi, M. Sudol, and J. W. Kelly, "WW: An isolated three-stranded antiparallel β-sheet domain that unfolds and refolds reversibly; evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state" *Protein Sci.*, 8, 841–853, 1999.

[161] D. Nečas and P. Klapetek, "Gwyddion: an open-source software for SPM data analysis" *Open Phys.*, 10, 181–188, 2012.

[162] H. Kim, S. P. Surwade, A. Powell, C. O'Donnell, and H. Liu, "Stability of DNA origami nanostructure under diverse chemical environments" *Chem. Mater.*, 26, 5265–5273, 2014.

[163] E. Chiancone and P. Ceci, "Role of Dps (DNA-binding proteins from starved cells) aggregation on DNA" *Front. Biosci.*, 15, 122, 2010.

[164] S. Kumagai, S. Yoshii, K. Yamada, N. Matsukawa, I. Fujiwara, K. Iwahori, I. Yamashita, "Electrostatic placement of single ferritin molecules" *Appl. Phys. Lett.*, 88, 153103, 2006.

[165] J. G. Heddle, S. Chakraborti, and K. Iwasaki, "Natural and artificial protein cages: design, structure and therapeutic applications" *Curr. Opin. Struct. Biol.*, 43, 148–155, 2017.

[166] W. M. Aumiller, M. Uchida, and T. Douglas, "Protein cage assembly across multiple length scales" *Chem. Soc. Rev.*, 47, 3433–3469, 2018.

[167] S. Nair and S. E. Finkel, "Dps protects cells against multiple stresses during stationary phase" *J. Bacteriol.*, 186, 4192–4198, 2004.

[168] V. O. Karas, I. Westerlaken, and A. S. Meyer, "The DNA-binding protein from starved cells (dps) utilizes dual functions to defend cells against multiple stresses" *J. Bacteriol.*, 197, 3206–3215, 2015.

[169] S. G. Wolf, D. Frenkiel, T. Arad, S. E. Finkel, R. Kolter, and A. Minsky, "DNA protection by stress-induced biocrystallization" *Nature*, 400, 83–85, 1999.

[170] Y. Zhang and B. P. Orner, "Self-assembly in the ferritin nano-cage protein superfamily" *Int. J. Mol. Sci.*, 12, 5406–5421, 2011.

[171] E. C. Theil, T. Tosha, and R. K. Behera, "Solving biology's iron chemistry problem with ferritin protein nanocages" *Acc. Chem. Res.*, 49, 784–791, 2016.

[172] L. N. Calhoun and Y. M. Kwon, "Structure, function and regulation of the DNA-binding protein Dps and its role in acid and oxidative stress resistance in *Escherichia coli*: a review" *J. Appl. Microbiol.*, 110, 375–386, 2011.

[173] E. Paszko and M. O. Senge, "Immunoliposomes" *Curr. Med. Chem.*, 19, 5239–5277, 2012.

[174] J. Pesek, R. Büchler, R. Albrecht, W. Boland, and K. Zeth, "Structure and mechanism of iron translocation by a dps protein from *Microbacterium arborescens*" *J. Biol. Chem.*, 286, 34872–34882, 2011.

[175] M.D. Hitchings, P. Townsend, E. Pohl, P.D. Facey, D.H. Jones, P.J. Dyson, R. Del Sol, "A tale of tails: deciphering the contribution of terminal tails to the biochemical properties of two Dps proteins from *Streptomyces coelicolor*" *Cell. Mol. Life Sci.*, 71, 4911–4926, 2014.

[176] R. Crichton, "Intracellular iron storage and biomineralization" in *Inorganic Biochemistry of Iron Metabolism*, John Wiley & Sons, Ltd, pp. 133–165, 2002.

[177] S. C. Andrews, "The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor" *Biochim. Biophys. Acta - Gen. Subj.*, 1800, 691–705, 2010.

[178] R. Janissen, M.M.A. Arens, N.N. Vtyurina, Z. Rivai, N.D. Sunday, B. Eslami-Mossallam, A.A. Gritsenko, L. Laan, D. de Ridder, I. Artsimovitch, N.H. Dekker, E.A. Abbondanzieri, A.S. Meyer, "Global DNA compaction in stationary-phase bacteria does not affect transcription" *Cell*, 174, 1188-1199.e14, 2018.

[179] M. D. Abràmoff, P. J. Magalhães, and S. J. Ram, "Image processing with imageJ" *Biophotonics Int.*, 11, 36–41, 2004.

[180] S. V. Hoffmann, M. Fano, and M. van de Weert, "Circular dichroism spectroscopy for structural characterization of proteins" in "Analytical Techniques in the Pharmaceutical Sciences. Advances in Delivery Science and Technology". Edited by A. Müllertz, Y. Perrie, and T. Rades, New York, NY: Springer, pp. 223–251, 2016.

[181] J. G. Lees, A. J. Miles, F. Wien, and B. A. Wallace, "A reference database for circular dichroism spectroscopy covering fold and secondary structure space" *Bioinformatics*, 22, 1955–1962, 2006.

[182] K. H. Nguyen, L. T. Smith, L. Xiao, G. Bhattacharyya, and A. Grove, "On the stoichiometry of *Deinococcus radiodurans* Dps-1 binding to duplex DNA" *Proteins Struct. Funct. Bioinforma.*, 80, 713–721, 2012.

[183] L. N. Calhoun and Y. M. Kwon, "Structure, function and regulation of the DNA-binding protein Dps and its role in acid and oxidative stress resistance in *Escherichia coli*: A review" *J. Appl. Microbiol.*, 110, 375–386, 2011.

[184] V. K. Moparthi, S. B. Moparthi, C. Howe, P. Raleiras, J. Wenger, and K. Stensjö, "Structural diffusion properties of two atypical Dps from the cyanobacterium *Nostoc punctiforme* disclose interactions with ferredoxins and DNA" *Biochim. Biophys. Acta - Bioenerg.*, 1860, 148063, 2019.

[185] N. R. de Alcântara, F. M. de Oliveira, W. Garcia, O. A. L. dos Santos, A. P. Junqueira-Kipnis, and A. Kipnis, "Dps protein is related to resistance of *Mycobacterium abscessus* subsp. *massiliense* against stressful conditions" *Appl. Microbiol. Biotechnol.*, 104, 5065–5080, 2020.

[186] T. Tosha, R. K. Behera, H. L. Ng, O. Bhattasali, T. Alber, and E. C. Theil, "Ferritin protein nanocage ion channels: Gating by N-terminal extensions" *J. Biol. Chem.*, 287, 13016–13025, 2012.

[187] L. F. Huergo, H. Rahman, A. Ibrahimovic, C. J. Day, and V. Korolik, "*Campylobacter jejuni* Dps protein binds DNA in the presence of iron or hydrogen peroxide" *J. Bacteriol.*, 195, 1970–1978, 2013.

[188] P. F. Collins, H. Diehl, and G. F. Smith, "2,4,6-Tripyridyl- s -triazine as reagent for iron. determination of iron in limestone, silicates, and refractories" *Anal. Chem.*, 31, 1862–1867, 1959.

[189] M. M. Juban, M. M. Javadpour, and M. D. Barkley, "Circular Dichroism studies of secondary structure of peptides" *Antibact. Pept. Protoc.*, 78, pp. 73–78, 1997.

[190] S. M. M. M. Kelly and N. C. C. C. Price, "The use of circular dichroism in the investigation of protein structure and function" *Curr. Protein Pept. Sci.*, 1, 349–84, 2000.

[191] A. J. Miles and B. A. Wallace, "Synchrotron radiation circular dichroism spectroscopy of proteins and applications in structural and functional genomics" *Chem. Soc. Rev.*, 35, 39–51, 2006.

[192] S. M. Kelly, T. J. Jess, and N. C. Price, "How to study proteins by circular dichroism" *Biochim. Biophys. Acta - Proteins Proteomics*, 1751, 119–139, 2005.

[193] B. A. Wallace, “Protein characterisation by synchrotron radiation circular dichroism spectroscopy” *Q. Rev. Biophys.*, 42, 317–370, 2009.

[194] M. A. Andrade, P. Chacón, J. J. Merelo, and F. Morán, “Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network” *Protein Eng. Des. Sel.*, 6, 383–390, 1993.

[195] B. A. Wallace and C. L. Teeters, “Differential absorption flattening optical effects are significant in the circular dichroism spectra of large membrane fragments” *Biochemistry*, 26, 65–70, 1987.

[196] M. Bloemendal, A. Toumadje, and W. C. Johnson, “Bovine lens crystallins do contain helical structure: a circular dichroism study” *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.*, 1432, 234–238, 1999.

[197] S. W. Provencher and J. Glöckner, “Estimation of globular protein secondary structure from circular dichroism” *Biochemistry*, 20, 33–37, 1981.

Appendix

A. Circular dichroism and synchrotron radiation circular dichroism

A1. Fundamentals

Circular dichroism (CD) is the differential absorption of left and right circularly polarized light of molecules that are either chiral or placed in chiral environments. L and D amino acids are chiral molecules since they do not superimpose their mirror images, with the exception of glycine. In proteins there are a number of chromophores that exhibit CD signals. Two spectral regions are of relevance when acquiring CD spectra of proteins, the far UV region (240-180 nm), which is relevant to the peptide bond absorption. In this region it is possible to acquire information of the regular secondary structural features of proteins such as α-helix and β-sheet [189]. When one analyses the near UV region (320 -260 nm) the information is relevant to the aromatic amino acid side chains environment and the tertiary structure of the protein. As such, CD is nowadays extensively used for protein structure characterisation and also ligand binding, being an extremely versatile technique in structural biology [190]. The development of synchrotron radiation circular dichroism (SRCD), which makes use of the intense light of a synchrotron beam, has further expanded the CD technique, due to the huge increase (over 1000 fold) of the intensity in the far UV region when compared with conventional CD spectrometers. SRCD provides a mean to acquire data at lower wavelength, allowing to obtain more structural information and also requiring smaller sample concentration, ***Figure A-1*** [191].

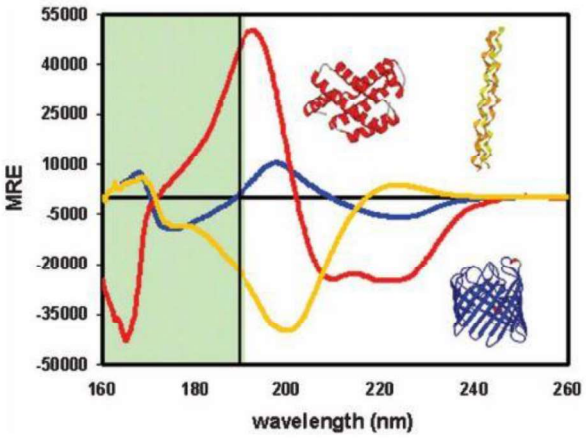

Figure A-1: SRCD of a most helical protein in red, a protein predominantly composed of β sheets in blue and a polyproline helix in yellow. The green rectangle marks the information only acquired by SRCD. Mean Residue Ellipticity (MRE) [163].

As previously stated, the radiation used in CD spectroscopy is circularly polarised, that is a plane-polarised radiation is split in two circularly polarised components when passing through a modulator wherein it is subjected to an alternating electric field of 50 kHz. This modulator is composed of a piezoelectric quartz crystal and a thin plate of an isotropic material coupled to the crystal. The alternating electric field induces structural changes in the piezoelectric, which will

make the isotropic material transmit circularly polarized light. When the radiation passes through the sample and either the left or right polarized components of light are absorbed a combination of the components is regenerated in the original plane and will now be elliptically polarised. The CD instrument detects the two radiation components separately and displays dichroism at a given radiation wavelength. This is expressed either by the difference of absorbance, equation A.1 of the two components or as the ellipticity in degrees equation A.2, where *b* and *a* are the minor and major axes of the resultant ellipse

$$\Delta A = A_L - A_R \qquad \text{(A.1).}$$

$$\theta = tan^{-1}(b/a) \qquad \text{(A.2)}$$

Conventional CD instruments use a xenon arc lamp as a light source, which covers the range of 178 to 1000 nm, that can be employed to almost all protein studies. However, it is first necessary to remove O_2 from the lamp housing in order to prevent ozone formation to minimize damage to the optical system, and to allow measurements below 180 nm. In order to be able to acquire data below 180 nm it is necessary to use synchrotron radiation sources.

When collecting a CD spectrum it is important to calibrate the instrument with a reliable chiral standard such as 1S(+)-10-camphrsulphonic acid. The experimental instrument parameters that can be adjusted are the time constant, the scan rate, the number of scans, and the bandwidth. As a rule of thumb, the product of the time constant and the scan speed should be smaller than 0.5 nm to avoid distortion of the spectrum [190].

One of the main requirements of CD spectroscopy that is common to other spectroscopic techniques is sample homogeneity. This should be accounted with either centrifugation or filtration of the sample, prior to measurement. In biological samples it is of upmost importance to pay attention to the buffer and solvents used. Phosphate, borate and low molarity (20 mM) Tris have low absorbances above 190 nm in cells with a pathlength of 0.1 cm or less. Buffers that usually cover low pH values, from 4.0 to 6.0, usually have carboxylate groups and tend to absorb intensively below 200 nm. As a golden rule, the sample buffer should always be measured as a "blank" in order to account for undesirable CD absorption. Regarding electrolytes that are commonly found in buffers it is important to keep in mind that chloride ions absorb strongly below 195 nm, therefore these should not be present in the sample and should be replaced, when possible, by fluoride or sulphate salts, to maintain the ionic strength.

The secondary structure of the protein under study will also only be able to be correctly estimated form a CD spectrum when the protein concentration is accurately known (the later can be determined by UV/visible spectroscopy or by amino acid analysis) [192]. In the spectral range of far UV CD there is the contribution of the peptide, therefore one has to know the molar

concentration of these bonds, which can be calculated by dividing the concentration in mass terms by the mean residue weight MRW. The MRW is obtained by dividing the molecular mass by (N – 1), where N is the number of amino acids in the polypeptide chain.

The units utilized in terms of absorbance are expressed in equation A.3 and units are $cm^{-1}M^{-1}$. In ellipticity the mean residue ellipticity at a specific wavelength is expressed in $deg.cm^2.dmol^{-1}$ and given by equation A.4, where θ is the observed ellipticity (degrees), d is the pathlength (cm) and c is the concentration of the protein (g/mL). At any wavelength the numerical relationship between values in the two sets of units is given by equation A.5

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R \qquad \text{(A.3)}$$

$$[\theta]_{mrw}\lambda = \frac{MRW\theta}{10dc} \qquad \text{(A.4)}$$

$$[\theta]_{mrw} = 3298\Delta\varepsilon \qquad \text{(A.5)}$$

Typical cell pathlengths for CD spectroscopy are in the range of 0.01 to 0.05 cm and protein concentrations are in the range of 0.2 to 1 mg/mL. The sample volume is dependent on the design of the cell and can go from about 1 mL to as little as 50 μL. It is then possible to acquire a CD spectrum with about10 μg of sample, however in order to explore optimal conditions for data acquisition between 100 and 500 μg should be considered [193].

The use of CD spectroscopy as a structural biology tool has a set of advantages in comparison with other techniques such as X-ray crystallography and NMR spectroscopy. One of the principal advantages is the fastness of the measurements. In under 30 minutes it is possible to acquire a good CD spectrum in the near UV spectral region. The fact that CD spectroscopy is a non-destructive technique, being possible to recover the sample used in the measurement, that can then be analysed by other techniques. The versatility of cells with different pathlengths also gives flexibility to the user to study a wide range of protein concentrations. The main limitation of this particular spectroscopy is that even though it provides a reliable estimate of the secondary structural content of a protein it does not indicate which regions of the protein are of a determined structural type. Given this it is almost always relevant to use CD spectroscopy as a complementary method to study a particular system [96].

A2. Data analysis

Concerning CD data analysis an effort has been made to determine the protein secondary structure based on empirical analysis of CD spectra. These include multilinear regression least squares, singular value position (SVD), ridge regression, the self-consistent method and neural networks. All these methods have in common the use of the spectral data from proteins with

known crystallographic structures to create reference databases, from which they calculate the structure of the measured protein [191]

Generally, multiple scans of both the protein and the baseline must be obtained. It is important to compare the individual scans and not just the averaged spectrum, this enables the user to determine systematic changes during acquisition, which could arise from denaturation or precipitation during the measurement session. The following step consist on the subtraction of the average baselines to the averaged sample spectrum.

The methods used to calculate the secondary structure of the sample assume that the crystal structure and structures in solution of the measured protein are the same, and contributions to the CD spectrum from individual secondary structures are additive.

Least square methods derive a spectrum, one for each secondary structural type, from the spectra of the individual proteins and then minimizes the deviation between the experimental and the calculated protein spectra. In constrained methods the sum of fractional weights is constrained to be equal to 1.0 [194]. Unconstrained, normalized least squares analyses do not constrain the fractions to sum to 1.0, but divide the fractions by the sum of the fractions [195]. This has the advantage that can be employed if the concentration of the protein is not accurately known. SVD [196] is an eigenvector method that extracts the principal component basis curves from a set of protein spectra. Each of this basis curve is then associated with a known mixture of secondary structures and used to analyse the spectrum of the unknown protein.

Other methods utilize selection procedures from a reference set with close CD spectra to the protein measured, which are then combined to produce the final spectrum. The CONTIN program [197] utilizes this type of approach. Another selection method that also includes a number of constrains is the self-consistent method SELCON3 [146] and the CDSSTR algorithm based on SVD [196]. Neural networks are also being employed wherein a network [194] is trained on a reference dataset of proteins to identify characteristic patterns associated with specific structural types. K2D is a type of neural network that employs this approach.

In general, the results given by all these methods depend on the proteins used in the basis dataset, and therefore it is expected that given the initial characteristics of the reference set of spectra the results obtained might differ. This problem usually does not arise if the sample protein has a relatively standard structure, all reference databases and methods will give similar results. However, one must be careful when the studied protein has unusual contributions, either by the presence of aromatic groups or disulphide bridges, which might impair the ability of the various methods to replicate the result.

Presently the best approach is to test different combinations of algorithms and different databases to determine the best result. DICHROWEB is a user-friendly web-based calculation server where this can be made. This server also calculate the normalized root mean square deviation (NRMSD) which allows the comparison of results between different methods and algorithms employed in a simple and easy manner [98].

B. DNA origami staples

B1.1 DNA origami staples map

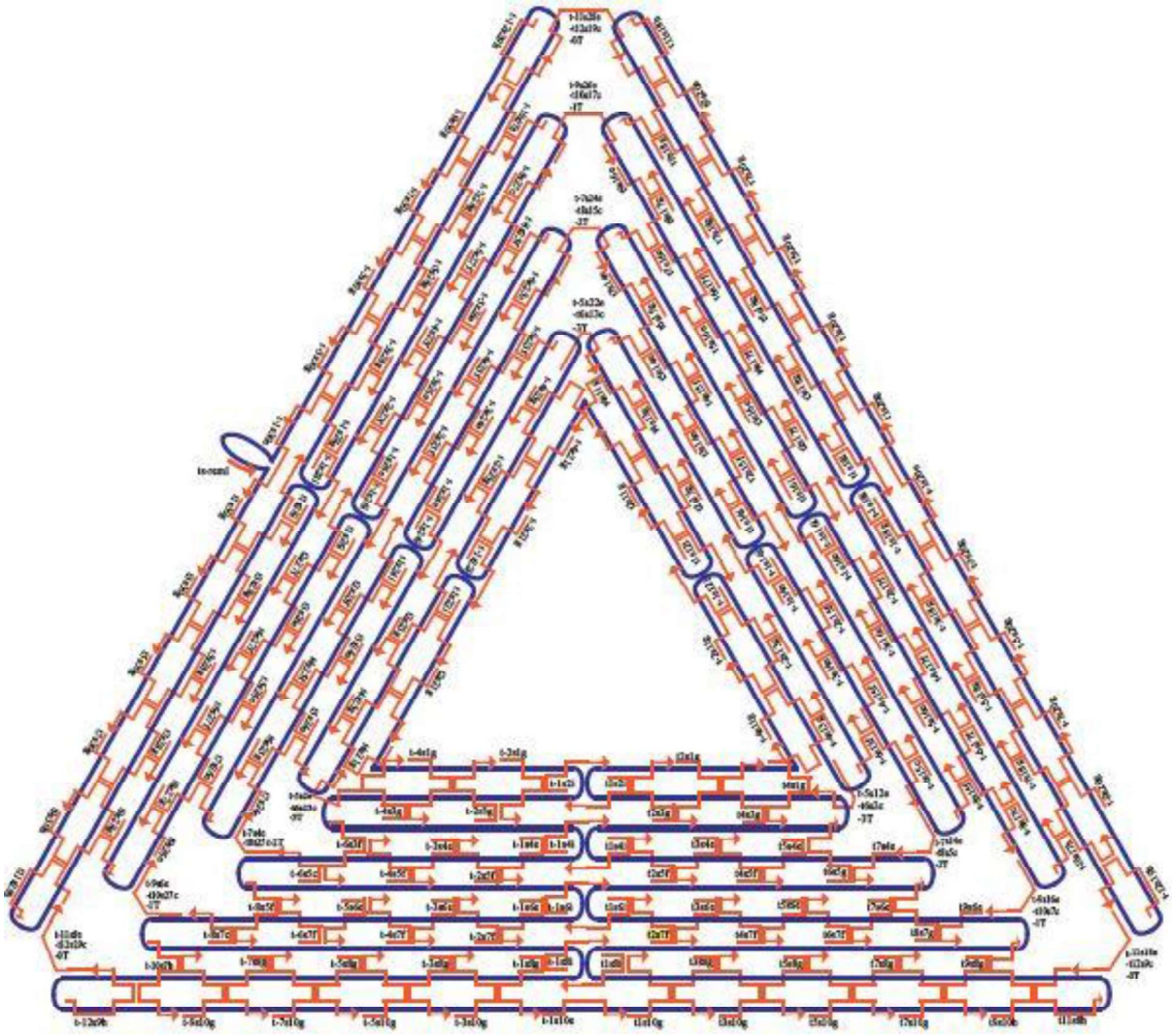

Figure B-1: DNA origami staple map

B1.2 DNA origami triangle staple sequence

t11s18h	AAT ACT GCG GAA TCG TAG GGG GTA ATA GTA AAA TGT TTA GAC T
t11s28h	TCT TTG ATT AGT AAT AGT CTG TCC ATC ACG CAA ATT AAC CGT T
t11s8h	CAG AAG GAA ACC GAG GTT TTT AAG AAA AGT AAG CAG ATA GCC G
t1s10g	GAC GGG AGA ATT AAC TCG GAA TAA GTT TAT TTC CAG CGC C
t1s12i	TCA TAT GTG TAA TCG TAA AAC TAG TCA TTT TC
t1s14i	GTG AGA AAA TGT GTA GGT AAA GAT ACA ACT TT
t1s16i	GGC ATC AAA TTT GGG GCG CGA GCT AGT TAA AG
t1s18i	TTC GAG CTA AGA CTT CAA ATA TCG GGA ACG AG

t1s20g	GAA TAC CAC ATT CAA CTT AAG AGG AAG CCC GAT CAA AGC G
t1s22i	TCG GGA GAT ATA CAG TAA CAG TAC AAA TAA TT
t1s24i	CCT GAT TAA AGG AGC GGA ATT ATC TCG GCC TC
t1s26i	GCA AAT CAC CTC AAT CAA TAT CTG CAG GTC GA
t1s28i	CGA CCA GTA CAT TGG CAG ATT CAC CTG ATT GC
t1s2i	CGG GGT TTC CTC AAG AGA AGG ATT TTG AAT TA
t1s30g	TTG ACG AGC ACG TAT ACT GAA ATG GAT TAT TTA ATA AAA G
t1s4i	AGC GTC ATG TCT CTG AAT TTA CCG ACT ACC TT
t1s6i	TTC ATA ATC CCC TTA TTA GCG TTT TTC TTA CC
t1s8i	ATG GTT TAT GTC ACA ATC AAT AGA TAT TAA AC
t2s11g	AGA AAA GCC CCA AAA AGA GTC TGG AGC AAA CAA TCA CCA T
t2s13g	ACA GTC AAA GAG AAT CGA TGA ACG ACC CCG GTT GAT AAT C
t2s15f	ATA GTA GTA TGC AAT GCC TGA GTA GGC CGG AG
t2s17f	AAC CAG ACG TTT AGC TAT ATT TTC TTC TAC TA
t2s1g	GAT AAG TGC CGT CGA GCT GAA ACA TGA AAG TAT ACA GGA G
t2s21g	CCT GAT TGC TTT GAA TTG CGT AGA TTT TCA GGC ATC AAT A
t2s23g	TGG CAA TTT TTA ACG TCA GAT GAA AAC AAT AAC GGA TTC G
t2s25f	AAG GAA TTA CAA AGA AAC CAC CAG TCA GAT GA
t2s27f	GGA CAT TCA CCT CAA ATA TCA AAC ACA GTT GA
t2s3g	TTT GAT GAT TAA GAG GCT GAG ACT TGC TCA GTA CCA GGC G
t2s5f	CCG GAA CCC AGA ATG GAA AGC GCA ACA TGG CT
t2s7f	AAA GAC AAC ATT TTC GGT CAT AGC CAA AAT CA
t3s10g	GTC AGA GGG TAA TTG ATG GCA ACA TAT AAA AGC GAT TGA G
t3s14e	CAA TAT GAC CCT CAT ATA TTT TAA AGC ATT AA
t3s16e	CAT CCA ATA AAT GGT CAA TAA CCT CGG AAG CA
t3s18g	AAC TCC AAG ATT GCA TCA AAA AGA TAA TGC AGA TAC ATA A
t3s20g	CGC CAA AAG GAA TTA CAG TCA GAA GCA AAG CGC AGG TCA G
t3s24e	TAA TCC TGA TTA TCA TTT TGC GGA GAG GAA GG
t3s26e	TTA TCT AAA GCA TCA CCT TGC TGA TGG CCA AC
t3s28g	AGA GAT AGT TTG ACG CTC AAT CGT ACG TGC TTT CCT CGT T
t3s30g	AGA ATC AGA GCG GGA GAT GGA AAT ACC TAC ATA ACC CTT C
t3s4e	TGT ACT GGA AAT CCT CAT TAA AGC AGA GCC AC
t3s6e	CAC CGG AAA GCG CGT TTT CAT CGG AAG GGC GA
t3s8g	CAT TCA ACA AAC GCA AAG ACA CCA GAA CAC CCT GAA CAA A
t4s11g	GCA AAT ATT TAA ATT GAG ATC TAC AAA GGC TAC TGA TAA A
t4s13g	CGT TCT AGT CAG GTC ATT GCC TGA CAG GAA GAT TGT ATA A
t4s15f	CAG GCA AGA TAA AAA TTT TTA GAA TAT TCA AC
t4s17f	GAT TAG AGA TTA GAT ACA TTT CGC AAA TCA TA
t4s1g	TAG CCC GGA ATA GGT GAA TGC CCC CTG CCT ATG GTC AGT G
t4s21g	GCG CAG AGG CGA ATT AAT TAT TTG CAC GTA AAT TCT GAA T
t4s23g	GAT TAT ACA CAG AAA TAA AGA AAT ACC AAG TTA CAA AAT C
t4s25f	TAG GAG CAT AAA AGT TTG AGT AAC ATT GTT TG
t4s27f	TGA CCT GAC AAA TGA AAA ATC TAA AAT ATC TT
t4s3g	TTT AAC GGT TCG GAA CCT ATT ATT AGG GTT GAT ATA AGT A
t4s5f	CTC AGA GCA TAT TCA CAA ACA AAT TAA TAA GT
t4s7f	GGA GGG AAT TTA GCG TCA GAC TGT CCG CCT CC

t5s10g	GAT AAC CCA CAA GAA TGT TAG CAA ACG TAG AAA ATT ATT C
t5s14e	TTA ATG CCT TAT TTC AAC GCA AGG GCA AAG AA
t5s16e	TTA GCA AAT AGA TTT AGT TTG ACC AGT ACC TT
t5s18g	TAA TTG CTT TAC CCT GAC TAT TAT GAG GCA TAG TAA GAG C
t5s20g	AAC ACT ATC ATA ACC CAT CAA AAA TCA GGT CTC CTT TTG A
t5s24e	AAT GGA AGC GAA CGT TAT TAA TTT CTA ACA AC
t5s26e	TAA TAG ATC GCT GAG AGC CAG CAG AAG CGT AA
t5s28g	GAA TAC GTA ACA GGA AAA ACG CTC CTA AAC AGG AGG CCG A
t5s30g	TTA AAG GGA TTT TAG ATA CCG CCA GCC ATT GCG GCA CAG A
t5s4e	CCT TGA GTC AGA CGA TTG GCC TTG CGC CAC CC
t5s6e	TCA GAA CCC AGA ATC AAG TTT GCC GGT AAA TA
t5s8g	TTG ACG GAA ATA CAT ACA TAA AGG GCG CTA ATA TCA GAG A
t6s15g	ATA AAG CCT TTG CGG GAG AAG CCT GGA GAG GGT AG
t6s17f	TAA GAG GTC AAT TCT GCG AAC GAG ATT AAG CA
t6s25g	TCA ATA GAT ATT AAA TCC TTT GCC GGT TAG AAC CT
t6s27f	CAA TAT TTG CCT GCA ACA GTG CCA TAG AGC CG
t6s5g	CAG AGC CAG GAG GTT GAG GCA GGT AAC AGT GCC CG
t6s7f	ATT AAA GGC CGT AAT CAG TAG CGA GCC ACC CT
t7s10g	ATA AGA GCA AGA AAC ATG GCA TGA TTA AGA CTC CGA CTT G
t7s14e	ATG ACC CTG TAA TAC TTC AGA GCA
t7s16e	TAA AGC TAT ATA ACA GTT GAT TCC CAT TTT TG
t7s18g	CGG ATG GCA CGA GAA TGA CCA TAA TCG TTT ACC AGA CGA C
t7s20g	GAT AAA AAC CAA AAT ATT AAA CAG TTC AGA AAT TAG AGC T
t7s24e	ACA ATT CGA CAA CTC GTA ATA CAT
t7s26e	TTG AGG ATG GTC AGT ATT AAC ACC TTG AAT GG
t7s28g	CTA TTA GTA TAT CCA GAA CAA TAT CAG GAA CGG TAC GCC A
t7s30g	GAA TCC TGA GAA GTG TAT CGG CCT TGC TGG TAC TTT AAT G
t7s4e	GCC GCC AGC ATT GAC ACC ACC CTC
t7s6e	AGA GCC GCA CCA TCG ATA GCA GCA TGA ATT AT
t7s8g	CAC CGT CAC CTT ATT ACG CAG TAT TGA GTT AAG CCC AAT A
t8s17g	TAA TTG CTT GGA AGT TTC ATT CCA AAT CGG TTG TA
t8s27g	CGC GAA CTA AAA CAG AGG TGA GGC TTA GAA GTA TT
t8s7g	AGC CAT TTA AAC GTC ACC AAT GAA CAC CAG AAC CA
t9s10h	TAT CTT ACC GAA GCC CAA ACG CAA TAA TAA CGA AAA TCA CCA G
t9s16e	ACT AAA GTA CGG TGT CGA ATA TAA
t9s18g	TGC TGT AGA TCC CCC TCA AAT GCT GCG AGA GGC TTT TGC A
t9s20h	AAA GAA GTT TTG CCA GCA TAA ATA TTC ATT GAC TCA ACA TGT T
t9s26e	ACC ACC AGC AGA AGA TGA TAG CCC
t9s28g	TAA AAC ATT AGA AGA ACT CAA ACT TTT TAT AAT CAG TGA G
t9s30h	GCC ACC GAG TAA AAG AAC ATC ACT TGC CTG AGC GCC ATT AAA A
t9s6e	CCA TTA GCA AGG CCG GGG GAA TTA
t9s8g	GAG CCA GCG AAT ACC CAA AAG AAC ATG AAA TAG CAA TAG C
t-10s17h	ACC AAC CTA AAA AAT CAA CGT AAC AAA TAA ATT GGG CTT GAG A
t-10s27h	AAC TCA CAT TAT TGA GTG TTG TTC CAG AAA CCG TCT ATC AGG G
t-10s7h	ACG ACA ATA AAT CCC GAC TTG CGG GAG ATC CTG AAT CTT ACC A
t-12s19h	CCT GAC GAG AAA CAC CAG AAC GAG TAG GCT GCT CAT TCA GTG A

t-12s29h	ACG TGG ACT CCA ACG TCA AAG GGC GAA TTT GGA ACA AGA GTC C
t-12s9h	TGC TAT TTT GCA CCC AGC TAC AAT TTT GTT TTG AAG CCT TAA A
t-1s10e	AGA GAA TAA CAT AAA AAC AGG GAA GCG CAT TA
t-1s12i	AGG GAT AGC TCA GAG CCA CCA CCC CAT GTC AA
t-1s14e	ATT TTC TGT CAG CGG AGT GAG AAT ACC GAT AT
t-1s14i	CAA CAG TTT ATG GGA TTT TGC TAA TCA AAA GG
t-1s16e	ATT CGG TCT GCG GGA TCG TCA CCC GAA ATC CG
t-1s16i	GCC GCT TTG CTG AGG CTT GCA GGG GAA AAG GT
t-1s18g	CGA CCT GCG GTC AAT CAT AAG GGA ACG GAA CAA CAT TAT T
t-1s18i	GCG CAG ACT CCA TGT TAC TTA GCC CGT TTT AA
t-1s20e	ACA GGT AGA AAG ATT CAT CAG TTG AGA TTT AG
t-1s22i	CGC GTC TGA TAG GAA CGC CAT CAA CTT TTA CA
t-1s24e	CAG TTT GAC GCA CTC CAG CCA GCT AAA CGA CG
t-1s24i	AGG AAG ATG GGG ACG ACG ACA GTA ATC ATA TT
t-1s26e	GCC AGT GCG ATC CCC GGG TAC CGA GTT TTT CT
t-1s26i	CTC TAG AGC AAG CTT GCA TGC CTG GTC AGT TG
t-1s28g	TTT CAC CAG CCT GGC CCT GAG AGA AAG CCG GCG AAC GTG G
t-1s28i	CCT TCA CCG TGA GAC GGG CAA CAG CAG TCA CA
t-1s2i	CCT TTT TTC ATT TAA CAA TTT CAT AGG ATT AG
t-1s30e	CGA GAA AGG AAG GGA AGC GTA CTA TGG TTG CT
t-1s4e	TTA TCA AAC CGG CTT AGG TTG GGT AAG CCT GT
t-1s4i	TTT AAC CTA TCA TAG GTC TGA GAG TTC CAG TA
t-1s6e	TTA GTA TCG CCA ACG CTC AAC AGT CGG CTG TC
t-1s6i	AGT ATA AAA TAT GCG TTA TAC AAA GCC ATC TT
t-1s8g	TTT CCT TAG CAC TCA TCG AGA ACA ATA GCA GCC TTT ACA G
t-1s8i	CAA GTA CCT CAT TCC AAG AAC GGG AAA TTC AT
t-2s11g	CCT CAG AAC CGC CAC CCA AGC CCA ATA GGA ACG TAA ATG A
t-2s13g	AGA CGT TAC CAT GTA CCG TAA CAC CCC TCA GAA CCG CCA C
t-2s15f	CAC GCA TAA GAA AGG AAC AAC TAA GTC TTT CC
t-2s17f	ATT GTG TCT CAG CAG CGA AAG ACA CCA TCG CC
t-2s1g	AAA ACA AAA TTA ATT AAA TGG AAA CAG TAC ATT AGT GAA T
t-2s21g	GCT CAT TTT TTA ACC AGC CTT CCT GTA GCC AGG CAT CTG C
t-2s23g	GTA ACC GTC TTT CAT CAA CAT TAA AAT TTT TGT TAA ATC A
t-2s25f	ACG TTG TAT TCC GGC ACC GCT TCT GGC GCA TC
t-2s27f	CCA GGG TGG CTC GAA TTC GTA ATC CAG TCA CG
t-2s3g	AGA GTC AAA AAT CAA TAT ATG TGA TGA AAC AAA CAT CAA G
t-2s5f	ACT AGA AAT ATA TAA CTA TAT GTA CGC TGA GA
t-2s7f	TCA ATA ATA GGG CTT AAT TGA GAA TCA TAA TT
t-3s10g	AAC GTC AAA AAT GAA AAG CAA GCC GTT TTT ATG AAA CCA A
t-3s14e	GTT TTG TCA GGA ATT GCG AAT AAT CCG ACA AT
t-3s16e	GAC AAC AAG CAT CGG AAC GAG GGT GAG ATT TG
t-3s18g	TAT CAT CGT TGA AAG AGG ACA GAT GGA AGA AAA ATC TAC G
t-3s20g	TTA ATA AAA CGA ACT AAC CGA ACT GAC CAA CTC CTG ATA A
t-3s24e	TGT AGA TGG GTG CCG GAA ACC AGG AAC GCC AG
t-3s26e	GGT TTT CCA TGG TCA TAG CTG TTT GAG AGG CG
t-3s28g	GTT TGC GTC ACG CTG GTT TGC CCC AAG GGA GCC CCC GAT T

t-3s30g	TAG AGC TTG ACG GGG AGT TGC AGC AAG CGG TCA TTG GGC G
t-3s4e	GAT TAA GAA ATG CTG ATG CAA ATC AGA ATA AA
t-3s6e	CAC CGG AAT CGC CAT ATT TAA CAA AAT TTA CG
t-3s8g	AGC ATG TAT TTC ATC GTA GGA ATC AAA CGA TTT TTT GTT T
t-4s11g	AGG TTT AGT ACC GCC ATG AGT TTC GTC ACC AGG ATC TAA A
t-4s13g	AGC GTA ACT ACA AAC TAC AAC GCC TAT CAC CGT ACT CAG G
t-4s15f	TAG TTG CGA ATT TTT TCA CGT TGA TCA TAG TT
t-4s17f	GTA CAA CGA GCA ACG GCT ACA GAG GAT ACC GA
t-4s1g	GAG CAA AAG AAG ATG AGT GAA TAA CCT TGC TTA TAG CTT A
t-4s21g	GTT AAA ATT CGC ATT AAT GTG AGC GAG TAA CAC ACG TTG G
t-4s23g	GGA TAG GTA CCC GTC GGA TTC TCC TAA ACG TTA ATA TTT T
t-4s25f	AGT TGG GTC AAA GCG CCA TTC GCC CCG TAA TG
t-4s27f	CGC GCG GGC CTG TGT GAA ATT GTT GGC GAT TA
t-4s3g	ACA TAG CGC TGT AAA TCG TCG CTA TTC ATT TCA ATT ACC T
t-4s5f	GTT AAA TAC AAT CGC AAG ACA AAG CCT TGA AA
t-4s7f	CCC ATC CTC GCC AAC ATG TAA TTT AAT AAG GC
t-5s10g	TCC CAA TCC AAA TAA GAT TAC CGC GCC CAA TAA ATA ATA T
t-5s16e	AAC AGC TTG CTT TGA GGA CTA AAG CGA TTA TA
t-5s18g	CCA AGC GCA GGC GCA TAG GCT GGC AGA ACT GGC TCA TTA T
t-5s20g	ACC AGT CAG GAC GTT GGA ACG GTG TAC AGA CCG AAA CAA A
t-5s26e	TGC TGC AAA TCC GCT CAC AAT TCC CAG CTG CA
t-5s28g	TTA ATG AAG TTT GAT GGT GGT TCC GAG GTG CCG TAA AGC A
t-5s30g	CTA AAT CGG AAC CCT AAG CAG GCG AAA ATC CTT CGG CCA A
t-5s6e	GTG TGA TAA GGC AGA GGC ATT TTC AGT CCT GA
t-5s8g	ACA AGA AAG CAA GCA AAT CAG ATA ACA GCC ATA TTA TTT A
t-6s13f	ACA GAC AGC CCA AAT CTC CAA AAA AAA ATT TCT TA
t-6s15c	CGA GGT GAG GCT CCA AAA GGA GCC
t-6s17f	ACC CCC AGA CTT TTT CAT GAG GAA CTT GCT TT
t-6s23f	CGG CGG ATT GAA TTC AGG CTG CGC AAC GGG GGA TG
t-6s25c	TGG CGA AAT GTT GGG AAG GGC GAT
t-6s27f	TGT CGT GCA CAC AAC ATA CGA GCC ACG CCA GC
t-6s3f	TCC CTT AGA ATA ACG CGA GAA AAC TTT TAC CGA CC
t-6s5c	GTT TGA AAT TCA AAT ATA TTT TAG
t-6s7f	AAT AGA TAG AGC CAG TAA TAA GAG ATT TAA TG
t-7s10g	GCC AGT TAC AAA ATA ATA GAA GGC TTA TCC GGT TAT CAA C
t-7s18g	AAA ACA CTT AAT CTT GAC AAG AAC TTA ATC ATT GTG AAT T
t-7s20g	ACC TTA TGC GAT TTT ATG ACC TTC ATC AAG AGC ATC TTT G
t-7s28g	TTC CAG TCC TTA TAA ATC AAA AGA GAA CCA TCA CCC AAA T
t-7s30g	CAA GTT TTT TGG GGT CGA AAT CGG CAA AAT CCG GGA AAC C
t-7s8g	GCG CCT GTT ATT CTA AGA ACG CGA TTC CAG AGC CTA ATT T
t-8s15f	CGG TTT ATC AGG TTT CCA TTA AAC GGG AAT ACA CT
t-8s17c	GGC AAA AGT AAA ATA CGT AAT GCC
t-8s25f	TCT TCG CTA TTG GAA GCA TAA AGT GTA TGC CCG CT
t-8s27c	GCG CTC ACA AGC CTG GGG TGC CTA
t-8s5f	TTC TGA CCT AAA ATA TAA AGT ACC GAC TGC AGA AC
t-8s7c	TCA GCT AAA AAA GGT AAA GTA ATT

t-9s10g	ACG CTA ACG AGC GTC TGG CGT TTT AGC GAA CCC AAC ATG T
t-9s20g	TGG TTT AAT TTC AAC TCG GAT ATT CAT TAC CCA CGA AAG A
t-9s30g	CGA TGG CCC ACT ACG TAT AGC CCG AGA TAG GGA TTG CGT T
ts-rem1	GCG CTT AAT GCG CCG CTA CAG GGC
t-5s2e-t6s23c	TTA ATT AAT TTT TTA CCA TAT CAA A
t-7s4e-t8s25c	TTA ATT TCA TCT TAG ACT TTA CAA
t-9s6e-t10s27c	CTG TCC AGA CGT ATA CCG AAC GA
t-11s8e-t12s29c	TCA AGA TTA GTG TAG CAA TAC T
t-5s12e-t6s3c	TGT AGC ATT CCT TTT ATA AAC AGT T
t-7s14e-t8s5c	TTT AAT TGT ATT TCC ACC AGA GCC
t-9s16e-t10s7c	ACT ACG AAG GCT TAG CAC CAT TA
t-11s18e-t12s9c	ATA AGG CTT GCA ACA AAG TTA C
t-5s22e-t6s13c	GTG GGA ACA AAT TTC TAT TTT TGA G
t-7s24e-t8s15c	CGG TGC GGG CCT TCC AAA AAC ATT
t-9s26e-t10s17c	ATG AGT GAG CTT TTA AAT ATG CA
t-11s28e-t12s19c	ACT ATT AAA GAG GAT AGC GTC C

B1.3 Modified DNA origami triangle strands

Sequence name	Sequence	5' Modification
t2s7f(ACGTU)(ACGTT)(GC)$_5$	AAA GAC AAC ATT TTC GGT CAT AGC CAA AAT CAA CGT UAC GTT GCG CGC GCG C	
t2s17f(ACGTU)(ACGTT)(GC)$_5$	AAC CAG ACG TTT AGC TAT ATT TTC TTC TAC TAA CGT UAC GTT GCG CGC GCG C	
t2s27f(ACGTU)(ACGTT)(GC)$_5$	GGA CAT TCA CCT CAA ATA TCA AAC ACA GTT GAA CGT UAC GTT GCG CGC GCG C	
(GC)5(AACGT)$_2$	G CGC GCG CGC AAC GTA ACG T	Biotin
t2s7f_78	AAA GAC AAC ATT TTC GGT CAT AGC CAA AAT CAA TGT CCG GCC CTA ATC TTC GCA CAA CTA TAC TGC ATA CGG TCC TGA GAT CTA CGC GAC AAC GCA ACC	
t2s17f_78	AAC CAG ACG TTT AGC TAT ATT TTC TTC TAC TAA TGT CCG GCC CTA ATC TTC GCA CAA CTA TAC TGC ATA CGG TCC TGA GAT CTA CGC GAC AAC GCA ACC CGA ACG TAC GTC GTA CAT TAG	
(TC)5(GACAT)_t2s7f_78	GGT TGC GTT GTC GCG TAG ATC TCA GGA CCG TAT GCA GTA TAG TTG TGC GAA GAT TAG GGC CGG ACA T	Biotin
(TC)5(GACAT)_t2s17f_78	CTA ATG TAC GAC GTA CGT TCG GGT TGC GTT GTC GCG TAG ATC TCA GGA CCG TAT GCA GTA TAG TTG TGC GAA GAT TAG GGC CGG ACA T	Biotin

In red, protruding nucleotides from DNA origami triangle; in blue complementary strand.

C. Plasmid vector maps

C1.1 pUC19 map

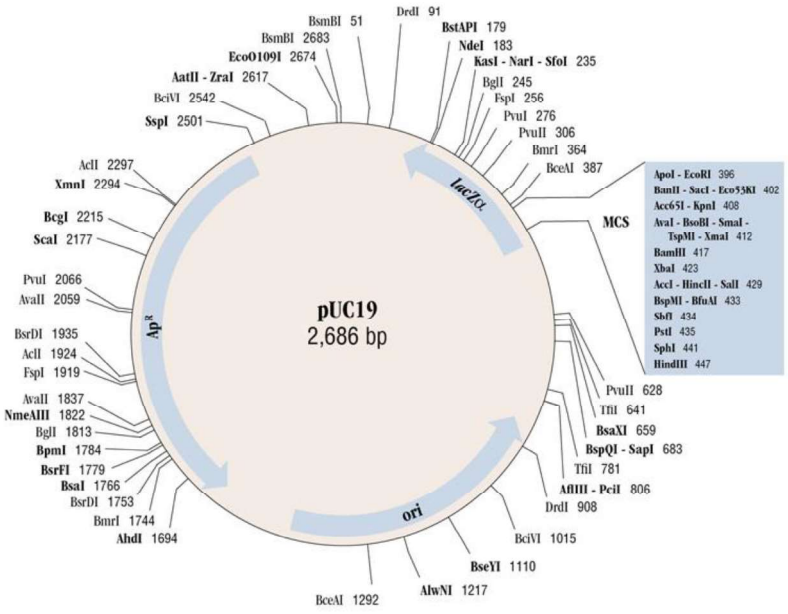

Figure C-1: Plasmid pUC19 vector map

C1.2 pET-21c(+) vector (Novagen)

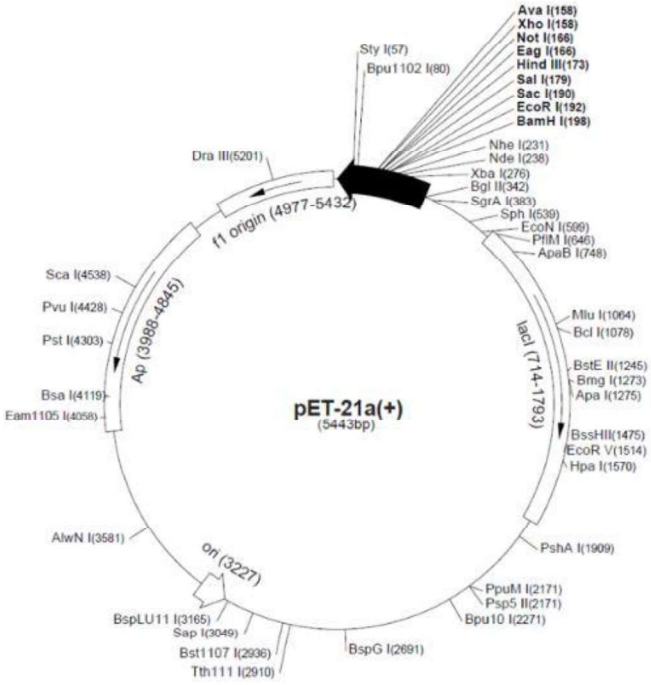

Figure C-2: Plasmid pET-21c(+) vector map

C1.3 pET-21c – Dps and variants expression vector

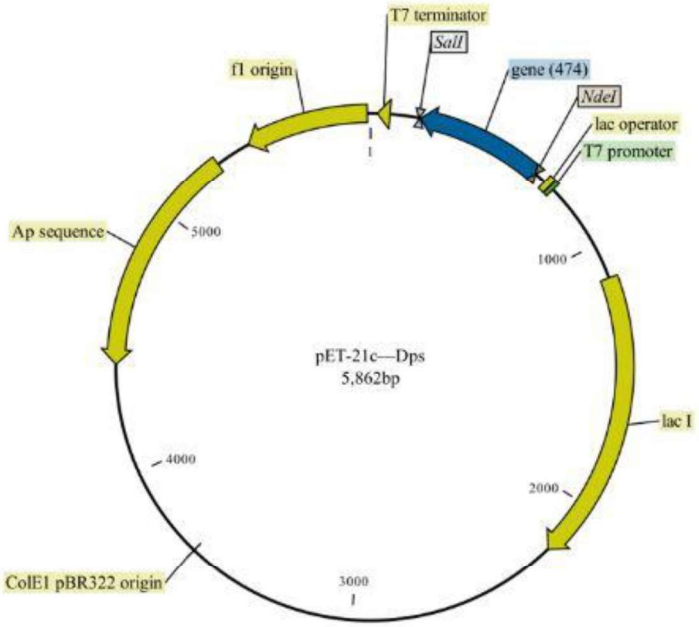

Figure C-3: Plasmid pET-21c–Dps map. Gene that code for Dps-WT and variants. The size of the WT and Q14E coding genes are indicated in the map. Dps T10C coding gene contains 474 bp, while Dps-Δ15 protein variant is codified by a gene of 432 bp.

D. Markers

D1. DNA markers

D1.1 Molecular weight NZYLadder II (NZYTech)

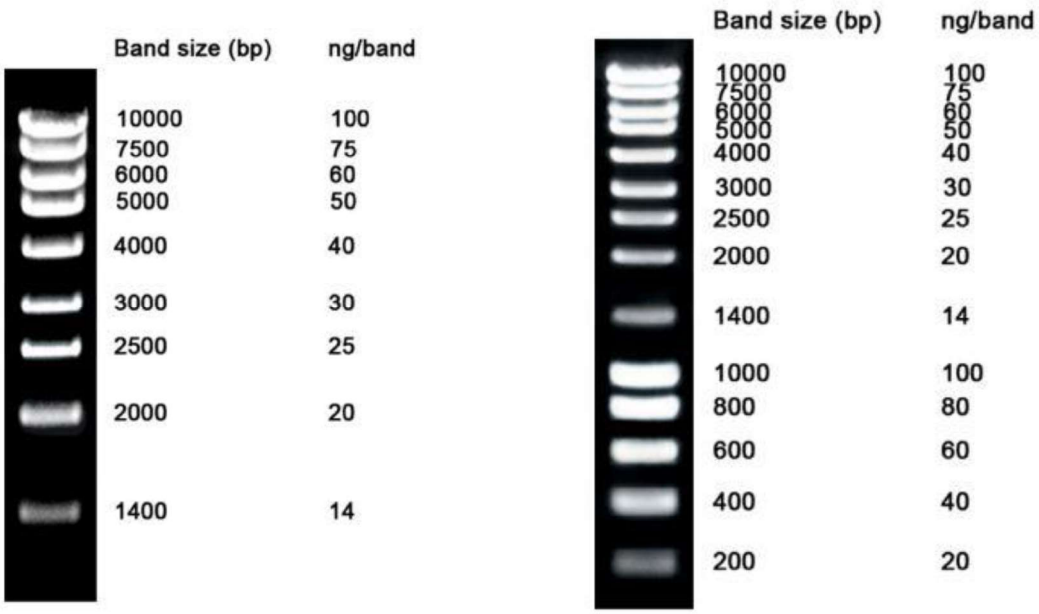

Figure D-1: Molecular weight markers NZYLadder II (left) and NZYLadder III (right) (NZYTech). 5 μL in a 1% agarose gel 1xTAE buffer (NZYTech)

D2. Protein markers

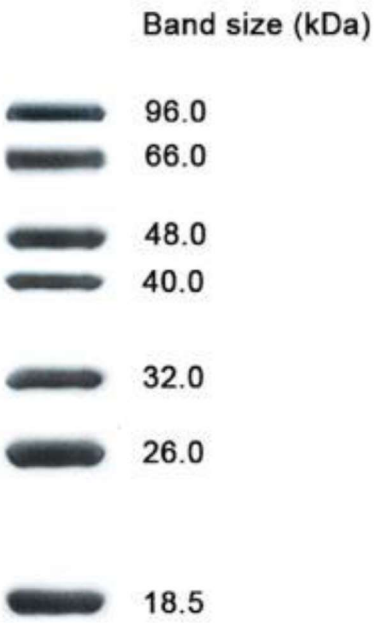

Figure D-2: Molecular weight LMW. 15 μL in a 14% Tris-glycine SDS-PAGE (NZYTech)