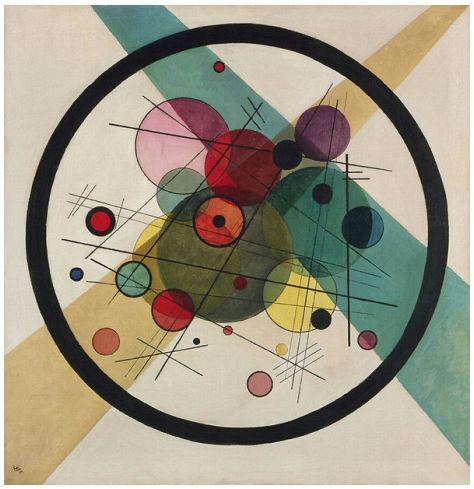
SAPIENZA
UNIVERSITÀ DI ROMA

PhD Course in Biochemistry

Cycle XXX (2014-2017)

Innovative ferritin nanocages
for drug-delivery and biotechnological applications



Matilde Cardoso Trabuco

Tutor:
Prof. Alessandra Bonamore

Coordinator:
Prof. Francesco Malatesta

Co-tutor:
Prof. Alberto Boffi

Everything starts with a dot.

Wassily Kandinsky

Acknowledgements

To my supervisors and mentors Alessandra Bonamore, Alberto Boffi and Paola Baiocco (honorary supervisor) for the tutoring, the guidance and for your kind patience. I could not have done it without you! I am extremely grateful to Alberto Boffi for accepting me in his lab and for always having a positive take on any situation.

A big thank you to my lab colleagues Irene Benni and Lorenzo Calisti, and all the other PhD students of Biochemistry, for all the help and most of all for the friendship. It has been fun working with you! Thank you Cécile Exertier, for the friendship, the companionship in all the travels, and for the hard work in the tiresome but fun beamtimes.

A sincere thank you to Francesco Malatesta, Beatrice Vallone, Matteo Levantino and Marco Cammarata, for all the long hours of exciting scientific discussion and specially for their patience.

I am grateful to to my family, whom I miss deeply and dearly, who has provided me nothing else than endless love and support, despite the distance. A special thank you to my big brother João, to whom I look up to and the most likely responsible for my passion for science.

To all the X-probe community, for being part of such an impressive team due to shape the future of the field.

Finally, I would like to thank the European Union's Horizon 2020 Research and Innovation Program for the funding, under the Marie Skłodowska-Curie grant agreement number 637295.

Table of Contents

Acknowledgements....3
1 Introduction....5
1.1 Protein cages....6
1.2 The ferritin scaffold....10
1.2.1 Probe encapsulation....14
1.2.2 Uptake and delivery mechanisms....20
1.2.3 Other applications....25
1.3 Assembly of protein cages....28
1.3.1 The ferritin self-assembly....32
1.4 Future prospects....34
2 Aim of the work....36
3 Materials and methods....39
3.1 Archaeal ferritin's permeation properties....40
3.2 Humanized archaeal ferritin....44
3.3 Lanthanide binding ferritin....52
4 Results and discussion....58
4.1 Archaeal ferritin's permeation properties....59
4.2 Humanization of an archaeal ferritin....67
4.2.1 "Humanized"*A. fulgidus* ferritin design....67
4.2.2 Structural characterization....71
4.2.3 Cellular uptake of Humanized ferritin....77
4.2.4 Humanized ferritin self-assembly....83
4.3 Lanthanide binding ferritin....94
4.3.1 Binding of Terbium to MuHF-LBT....96
4.3.2 Structural Analysis of MuHF-LBT with and without Tb(III)....99
4.3.3 Uptake of MuHF-LBT by tumor cell lines....102
5 Conclusions....105
6 Abbreviations....111
7 References....115
Appendix I....131
Appendix II....139
Appendix III....149

1 Introduction

1.1 Protein cages

Nature has since long served as a great source of inspiration in biomedice and biotechnology and the use of naturally occurring structures has shown great potential as nanoscaled building blocks and biotemplates for diverse nanotechnology applications. The use of proteins as multivalent multi-functional carriers for biotechnology applications is experiencing rapid development. Protein cages are container-like protein architectures formed from protein subunits that self-assemble into nearly monodisperse size distribution and whose functionality can be tailored by introduction of chemical modifications or by genetic engineering. Naturally available protein cages, such as viral capsids, small heatshock proteins and ferritins provide a privileged template for engineering modifications. These nanodevices can be loaded with chemotherapeutics and/or contrast agents, and thus develop theranostic devices of great clinical value. They come in a wide variety of sizes, stability and functionalities and work by providing a protected environment for their cargoes (Uchida et al., 2007). Protein cages serve as multipurpose containers for the transport of viral payloads (Cann, 2012), biomineralization (Theil et al., 2013), protein folding (Saibil, 2008) and degradation (Kish-Trier et al., 2013), enzyme encapsulation (Sutter et al., 2008) and can even work as catalysis microenvironment (Maity et al., 2016; Milne, 2002), but find their most prominent application in the biomedical field as drug-delivery systems (DDS) (Molino & Wang, 2014). Typically, protein cages self assemble in highly symmetrical arrangements based on

helical, icosahedral, cubic or tetrahedral symmetries with three topologically distinct regions: the interior, the exterior or the interface between subunits. These interfaces provide a platform for insertion of modifications allowing for customizable applications. There are few other nanoplatforms that offer simultaneously biocompatibility, control over size and shape, and the ability to modify and combine functionalities. Their application in the biotechnological and biomedical field are only limited by one's imagination (Molino & Wang, 2014; Sasaki et al., 2017; Uchida et al., 2007).

Some noteworthy examples of the use of protein cages are described in the literature:

I. Kramer et al. have reported the use of an engineered ferritin nanocage for the synthesis of silver nanoparticles to overcome the issue of size polydispersity and low crystal ordering of physical nanoparticle synthesis.(Kramer et al., 2004);

II. Falvo et al. have described the use of an innovative antibody-drug conjugate for the treatment of melanoma using a ferritin cage encapsulating the anticancer drug cisplatin conjugated with monoclonal antibodies (mAb) to the human melanoma-specific antigen Chondroitin sulfate proteoglycan 4 (CSPG4).(Falvo et al., 2013);

III. Flenniken et al. demonstrate that Heatshock protein (Hsp) cage from *Methanococcus jannaschii* can be used to confer cell-specific targeting of mammalian cells through combination of genetic and

chemical modifications. Hsp was conjugated on its outer surface with a tumor vascular targeting peptide or anti-cluster of differentiation (CD) 4 antibodies and on its inner surface with fluorescent molecules, thus conferring the ability to simultaneously incorporate cell targeting and imaging agents within a single protein cage (Flenniken et al., 2006);

IV. Of note is the work carried by Millan et al., where cowpea chloroticmoottle virus (CCMV) capsid was used to encapsulate smaller paramagnetic micelles containing a compound complexed with Gd(III) or Zn(II) phtalacyanine for their use as contrast agents for magnetic resonance imaging (MRI) (Millán et al., 2014).

Nevertheless, protein cages find their most striking application as DDS. The uniform size, multifunctionality, biocompatibility and biodegradability push them to the frontiers of drug-delivery (MaHam et al., 2009). All kinds of DDS, both synthetic (silica, metallic, polymers, gels) and natural (proteins, lipids, sugars), have been developed to address the issues of conventional drugs and improve their pharmacological properties (Safari & Zarnegar, 2014). As the medical world is undergoing a transition from traditional medicine to personalized medicine, more and highly versatile scaffolds for DDS are needed, as well as deeper understanding of their properties. The potential of conventional drugs in clinic is sometimes hampered due to poor pharmacodynamics – what the drug does to the body, including its mechanism of action – and pharmacokinetics – what the body does to the drug, commonly divided in 4 phases: absorption, distribution, metabolism

and excretion (Levy & Bauer, 1986). The development of appropriate DDS can overcome barriers at any of those levels and bring drugs to therapeutic areas otherwise unaccessible. A striking example where the development of DDS is the of the utmost necessity is in cancer therapy. Upon administration, anticancer drugs are distributed throughout the body usually with significant off-site effects that are responsible for the grave side-effects of cancer therapy. This reduces the effectiveness of the treatment as it limits it to a tolerable dose with sub-optimal amount of on-target drug (Cho et al., 2008). Another common limitation to anticancer drugs is their short "time of action", i.e., the time available to be in contact with their intended target and perform their action. This can result from a rapid clearing from the bloodstream by the kidneys/immune system or from their metabolization into inactive forms. Furthermore, drugs can have poor water solubility and relatively short shelf-stability. (Cho et al., 2008; MaHam et al., 2009). All these issues can be addressed with the development of advanced, tailored DDS.

When compared to their non-proteic counterparts, protein-based DDS perform better at various levels: proteins raise less biocompatibility issues and can be exploited to bypass the defenses of the immune system; they usually have a well characterized atomic structure; when assembled into cages are characterized by narrow size-distribution because of rigid geometrical constraints; they can be easily submitted to conjugation at any of the conceptual interfaces by adding chemical compounds, antibodies, oligosaccharides or lipids, or any conceivable combination of these,

depending on the intended application; they can be modified by genetic engineering, either to add specific conjugation sites or to alter intrinsic characteristics of the protein scaffold itself; and they allow for large scale production using relatively simple recombinant expression in bacterial systems (Molino & Wang, 2014; Todd et al, 2013).

Over 30 therapeutically relevant compounds have been investigated with protein-based DDS, such as the classic anticancer model drugs doxorubicin (Simsek & Akif Kilic, 2005), rapamycin and colchicine, others compounds such as progesterone, glucocorticosteroids, or even metal ions as cobalt, manganese, cadmium. Currently, a wide range of protein-based platforms are being investigated for the treatment of various cancers, infectious diseases, chronic diseases and autoimmune diseases (Kuruppu et al., 2015; Lee et al., 2015; Vannucci et al., 2012; Z. Yang et al., 2007). A particularly successful case of a protein-based conjugate that reached clinics is the albumin-paclitaxel nanoparticle (Abraxane ®) that has been recognized as one of the first approved protein-based nanoparticle for human use by the United States Food and Drug Administration (US FDA) (Todd et al., 2013). This success inspires researchers to pursue the development of other innovative and more efficient protein-based nanocarriers.

1.2 The ferritin scaffold

Ferritin was first described by the German pharmacologist Oswals Schmiedeberg in 1894, who described it as an iron-rich component in horse

livers. Later, in 1937, it was purified from horse spleen and characterized by the Czech biologist Vilém Laufberger who proposed ferritin to act as an iron depot in the organism. Early studies of ferritin were possible due to its distinct biochemical characteristics: stability at high temperatures, relative insolubility in ammonium sulfate solutions and its crystallization with cadmium salts (Alkhateeb et al., 2013).

Alongside with the research, the term *ferritin* as evolved and presently ferritin describes a family of proteins, ubiquitous in all life forms with the notable exception of yeast, whose physiological role as the cellular repository of excess iron is unique. Ferritin's role can be seen as that of a tiny organelle for its ability to store iron and to release it in a controlled manner, avoiding toxicity by excess iron. Ferritin can be found in both extracellular and intracellular compartment: although mainly cytosolic, it can be found in mitochondria, nucleus or free in the serum (Andrews, 2010). The cytosolic ferritins' key contribution for iron homeostasis is reflected by its ubiquitousness, but the physiological function of secreted ferritin remains unclear. Elevated levels of serum ferritin have been linked to inflammation, angiogenesis and tumors, and thus considered a marker for those conditions (Meyron-Holtz et al., 2011; Wang et al., 2010). Members of the ferritin family of proteins share characteristic structural features: the presence of a conserved fold made of a four-helix bundle domain in each subunit and the assemble of twenty-four monomeric units in a roughly spherical macro-structure resembling a cage. This translates in the appearance of an hollow interior cavity that provides an environment sheltered from the external

conditions. In nature the cavity is used to sequester iron in a non-reactive, biomineralized form and to serve as a cellular stock of iron for biological reactions, such as integration of cofactors for iron proteins (heme, FeS, mono and diiron). The iron biomineralization reaction also provides an antioxidant effect, acting as a depot for iron and oxidants arising from damaged proteins. (Theil et al., 2013; Uchida et al., 2010)

Other than storing functions, ferritins have catalytic functions. The active site is called the ferroxidase center (FOC) and it catalyzes the oxidation of Fe(II) into Fe(III), reducing either dioxygen or hydrogen peroxide to synthesize $Fe_2O_3 \bullet H_2O$, forming an iron-core within the protein shell (Theil et al., 2013).

The ferritin family of proteins is subdivided in three sub-classes, with different morphological and functional characteristics: DNA-binding protein from starved cells (Dps), bacterioferritins (Bfr) and the classical ferritins (Andrews, 2010).

Dps protein is the most dissimilar within the family. It has the ability to bind DNA non-specifically and functions mostly in ferrous iron detoxification. Rather than storage, Dps acts as oxidative stress protecting factor, most likely to avoid DNA oxidative damage. As such, Dps protects the cell from reactive oxygen species (ROS) generated by the Fenton reaction: $Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + \bullet OH$ (Zhao et al., 2002). In these proteins, the FOC sits between two subunits at the two-fold axis interface instead of being located within the four helix bundle motif (Ilari et al., 2000). Another striking feature of Dps is its smaller size that accounts for it being referred as mini-ferritin, as to contrast with the other members of the family – the macro-ferritins. Dps structure is

composed of twelve subunits instead of the canonical twenty-four, forming a cage with approximately 9.5 nm outer diameter and a significantly lower iron-storage capacity (Haikarainen & Papageorgiou, 2010). The crystal structure also revealed a characteristic 23 point symmetry that contrasts with the classic 432 symmetry found in other ferritins members, the absence of the short final E helix, and the presence of a short helix within the BC loop (Grant et al., 1998).

Bacterioferritin resembles classic ferritin in almost all aspects. It performs the same biological role as canonical ferrritins but its distribution among species is restricted to the bacterial and archaeal domains of life (Andrews et al., 1991). It retains the 24-meric quaternary structure and the typical four helix domain enclosing the FOC in each subunit as well as the final short E helix (Frolow et al., 1994). Its most notable difference is the presence of 12 heme groups within the protein shell, located at the dimer two-fold interface. The role of the heme group is unclear. but is thought to be involved in the release of iron from the iron-core (Frolow et al., 1993; Weeratunga et al., 2009).

Classical ferritins are the archetype of the ferritin family of proteins and can be found in all domains of life. Their cage structure is composed of 24 subunits assembled in a rhombic dodecahedron and, similarly to Bfr, they have the capacity to store-iron in a compact iron core within the protein shell. Each monomer folds in a four helix bundle motif with a long loop connecting helices B and C plus a fifth short helix at the C-terminal pointing inwards the cavity (Andrews, 2010; Theil et al., 2006). The FOC is located within the four helix bundle with most residues involved in the ferroxidase activity

conserved between classical and Bfrs. The subunits assemble in a 24-meric cage of around 450 kDa with an outer diameter of approximately 12 nm and an inner diameter of around 8 nm.(Arosio et al., 2017; Harrison & Arosio, 1996) The assembled ferritin can be formed by distinct but homologous subunits, like the case of human ferritin, where different contents of L and H subunits (L for light and H for heavy) co-assemble to form heteropolymers optimized for iron incorporation and mineralization. Different ratios of L to H subunits are are found throughout the various human tissues (Levi et al., 1987). The topological distribution of monomers in the 24-mer protein shell results in a 432 point symmetry, with eight hydrophilic pores at the three-fold symmetry axis, and six hydrophobic channels at the four-fold symmetry axis. The channels at the three-fold axis have been proposed to serve as entry channels for Fe(II) atoms (Alkhateeb & Connor, 2013; Arosio et al., 2017).

1.2.1 Probe encapsulation

Understanding the keys aspects of probe loading, entrapment and release is of major importance for the use of ferritin-based DDS. Molino and Wang (2014) have reviewed and summarized the strategies used for drug encapsulation and release for protein cages (Figure 1.1): drug-loading strategies include assembly/disassembly, covalent modification, pore entry and physical interactions. Drug-releasing strategies include pH triggered release, reducing environment, biodegradation and diffusion (Molino & Wang, 2014). Most of these strategies have been attempted using a variety of ferritin scaffolds.

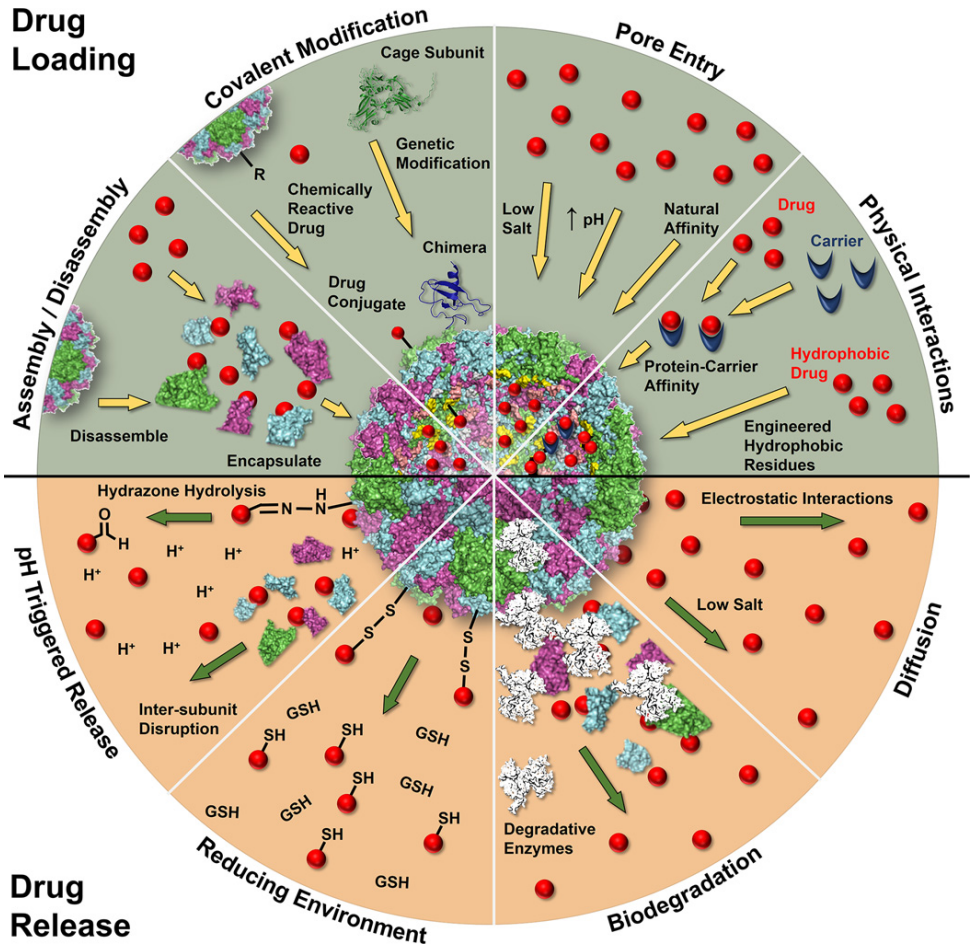


Figure 1.1 – Summary of strategies for drug loading and release using protein cages. Image adapted from Molino et al. (Molino & Wang, 2014)

Assembly/Disassembly

The loading of payloads within the ferritin's cavity can be achieved by exploiting the association-dissociation properties of the multimeric shell, thus inducing the disassembly into subunits followed by reassembly in the presence of the wanted cargo under selected experimental conditions. Payload entrapment is then achieved as a physical confinement into the cavity under reassembly conditions. Nevertheless, given the high stability of the 24-mer species, the dissociation process requires harsh conditions, such as acidic pH values, with suboptimal loading and low protein recovery. The association-dissociation process is in fact dependent upon an extensive network of interactions at the subunit interfaces, comprising hydrophobic interactions, salt bridges and hydrogen bonds that can be eventually manipulated by genetic engineering in order to affect the strength of intersubunit contacts (Huard et al., 2013). In this framework, the incomplete understanding of the physical aspects governing the assembly/disassembly of ferritins poses a challenge for successful application of this strategy. Human ferritin has been the most frequently used to entrap a number of compounds. Its disassembly is induced by lowering the pH to around 2.0 and re-assembly achieved by restoring it to neutral pH. Under this procedure, only pH resistant compounds can be loaded, and the re-assembly process has been reported to be incomplete, leading to a statistical distribution of "subunit defects" within the 24-meric structure (Kim et al., 2011). Nonetheless, the pH jump method is still the most popular procedure in order to achieve successful encapsulation.

Covalent Modifications

The covalent conjugation of drugs to proteins is a common method of drug-loading in protein carriers, and has shown its potential with the use of antibody-drug conjugates (ADC) in clinics. The added value of using a protein cage is the protected environment it offers, useful for particularly sensitive drugs, and the possibility to combine multiple strategies. The bioconjugation of small molecules to aminoacid side-chains, native or recombinantly engineered, such as cysteines, tyrosines, or the amine or carboxyl groups, can be attained using a wide range of chemical techniques. The approach for bioconjugation has to be carefully designed as it may result in poor yields and protein instability. (Molino & Wang, 2014) In a very recent report, a successful conjugation of aldoxorubicin (doxorrubin with and extra acid sensitive linker) with human ferritin external cysteine has been obtained. However, further covalent modification with polyethylene glycol molecules was necessary (Huang et al., 2017).

Pore Entry

Even though the protein shell offers a protected environment for its cargo, the 24-meric scaffold does show diverse possible pathways for small molecules entry/exit. The eight hydrophilic channels located at the 3-fold symmetry axis display a relatively broad selectivity and allow for the exchange of cations during mineralization/demineralization, with a preference for divalent cations (Haikarainen et al., 2011). They have also

been proposed as the main route of entry for a number of other substances, such as hydrophilic drugs and labeling fluorophores (He & Marles-Wright, 2015; Wu et al., 2008). The remaining hydrophobic channels, located at the four-fold axis can serve as entry for hydrophobic drugs and are thought to work as proton outlets (Douglas & Ripoll, 1998). Besides the small four-fold or three-fold channels, the overall intersubunit interfaces might offer room for exchange of small molecules, as proposed very recently by Zhang et al (2017). in a remarkable set of high resolution microscopy observations, they have demonstrated that the ferritin shell is not entirely rigid, but does possess a degree of flexibility, possibly due to a collective "breathing" motion that involves dynamic readjustments of the subunits within the 24-mer. During these motions, large organic compounds like curcumin and doxorubicin can be incorporated within the cavity simply driven by the outward/inward concentration gradient. Thus, under appropriately high concentration of the payload, a sufficient incorporation can be achieved simply by mixing the small molecule with the protein even at neutral pH values (Zhang et al., 2017).

Physical Interactions

Although secondary, tertiary and quaternary structures of ferritins are well conserved, their DNA sequences can have low homology. (Harrison & Arosio, 1996) Nevertheless, apo-ferritin, i.e., when there is no formed iron-core, ferritins' cavity is most likely filled with the solvent from the surrounding solution. (Theil et al., 2006; Yang & Chasteen, 1996) Its inner

surface can be redesigned to improve the biding of cargo molecules by genetically engineering residues to turn the cavity more hydrophobic/hydrophilic or by addition of specific binding sequences (Falvo et al., 2016; Kitagawa et al., 2012).

Drug-release strategies

There are as many different ways to encapsulate probes as there are to release them. The selection of the release mechanism is dependent on the desired application and on the stability and chemical properties of the probe-cage system. Engineering of this properties can be done in a similar fashion as for probe-encapsulation. Using non-specific interactions for encapsulation results in slow release of the probe over time. Controlled release can be achieved by interfering with release kinetics, as for example: amide or thioether bonds are disrupted during biodegradation (Kang et al., 2012). Disulfide bonds allow probe release in reducing environments and hydrazone bonds enable release in slightly acidic conditions. These mimic conditions reported to occur during endocytosis (Molino & Wang, 2014). In particular, it is expected that most payloads can be delivered once the protein is confined within a low pH environment in lysosomes. However, given the complex intracellular pathway of ferritins, these expectations are not always fulfilled.

1.2.2 Uptake and delivery mechanisms

Ferritin is responsible for iron storage within cells, however it also circulates and has been demonstrated to bind specifically to diverse cell types. In the 1960's several research groups reported the selected uptake of human ferritin by tumor cells (Easty et al., 1964; Ryser et al., 1962), and later, in 1988, Fargion et al. Demonstrated human H-chain ferritin's (HuHF) ability to bind a 100 kDa protein (Fargion et al., 1988), however it was not until 2010 that Li et al. identified human transferrin receptor-1 (TfR1) as mediator of the binding of HuHF, but not human L-chain ferritin (HuLF), to a variety of cell types. (Li et al., 2010). Later on, Fan et al. reported that HuHF targets TfR1 on tumor cells and tissues obtained from clinical samples, including nine of the most common solid tumors – liver, lung, colon, cervical, ovarian, prostate, breast and thymus cancers (Fan et al., 2012). This dual capacity of the receptor to interact with both transferrin (Tf) and ferritin raised the possibility that TfR1 acts a potential regulation center for the processing and use of cellular iron storages.

Tfr1 is a homodimeric type II transmembrane receptor located on the cell surface that has been identified as the receptor for Tf, essential for iron delivery from Tf to cells. Each polypeptide subunit has approximately 760 aminoacids and is composed of a short N-terminal cytoplasmic domain, a hydrophobic transmembrane domain and a large, globular extracellular C-terminal domain with the binding site for transferrin (Lawrence et al., 1999). TfR1 is thus established as a gatekeeper for the regulation of Tf-TfR1 iron uptake pathway, the main route for cellular iron uptake. Additionally, TfR1

has been shown to take part in the regulation of cell growth (Aisen, 2004; Neckers & Trepel, 1986) and multiple studies report an upregulation of TfR1 expression in proliferating cells, such as cancer cells (Calzolari et al., 2007; Kondo et al., 1990; Prutki et al., 2006; Ryschich et al., 2004). Overexpression might be a consequence of the high rate metabolism of cancer cells resulting in increased iron requirements since iron acts as an important co-factor for enzymes involved in DNA synthesis (Daniels et al., 2006). Although TfR1 is also present in normal physiological conditions, the overexpression of TfR1 in cancer cells and its efficient internalization highlight its potential as a pharmacological target for cancer diagnosis and treatment.(Fan et al., 2013)

The cellular uptake of iron via the TfR1 occurs via clathrin-mediated endocytosis. Upon Tf binding to TfR1, two distinct pathways are described. Either the complex is recycled back to the cell surface or it proceeds to lysosomal degradation (Figure 1.2).

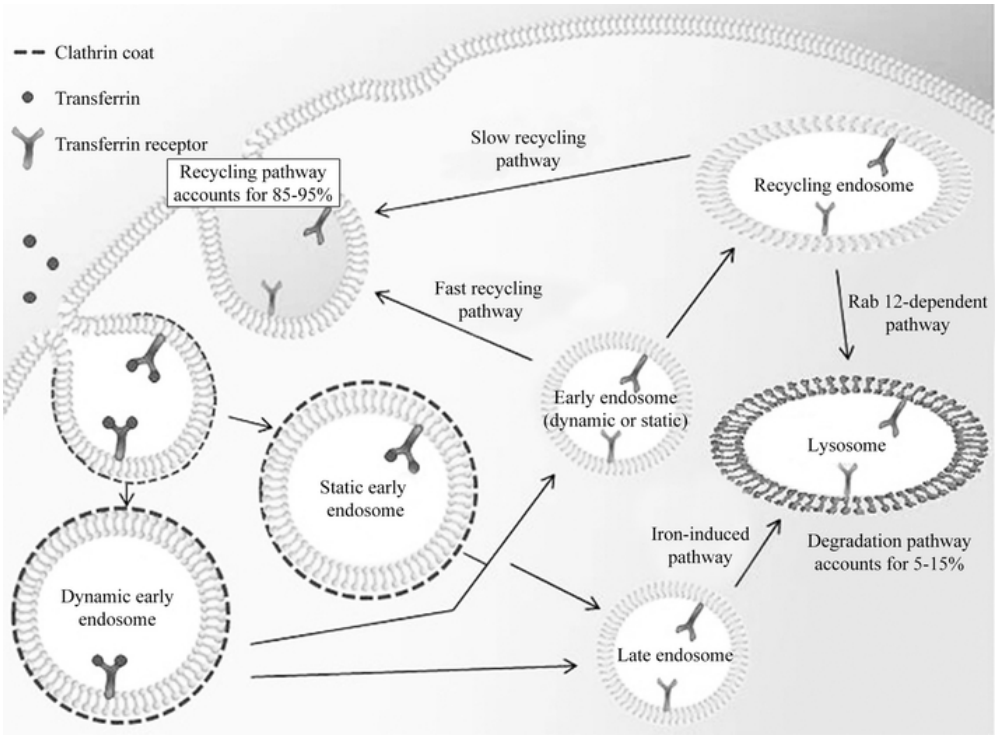


Figure 1.2 – Transferrin receptor-mediated endocytosis. Upon binding, endocytosis begins with the formation of clathrin-coated pits, and migration of the complex to the cytoplasm. Transferrin then enters the endosomes where dissociation of iron from transferrin occurs. At this point, one of two routes can be taken: either the complex is recycled back to the surface or undergo lysosomal degradation. The majority (85-95%) of the receptors will be recycled and a small fraction (5-15%) will be follow the degradation pathway for the purpose of receptor turnover. Image adapted from Tortorella et al. (Tortorella et al., 2014)

A considerable amount of effort has been dedicated to the development of strategies targeting TfR1 for cancer diagnosis and treatment. One of the most extensively studied systems concerns the doxorubicin-transferrin conjugate (Munns et al., 1998; Tortorella et al., 2014). Doxorubicin is widely used in

the treatment of a large variety of cancers but despite its large spectrum applicability to treat malignancies, its use is dose-limited as it accumulates in cardiac tissue, causing significant cardiotoxicity (Carvalho et al., 2009). By directing doxorubicin to cancer cells, one can reduce its accumulation in heart and improve the concentration of drug in the tumor tissue.

Another noteworthy strategy using TfR1 targeting is the use of transferrin /anti-transferrin receptor antibodies for delivery of therapeutics across the blood brain barrier. Crossing the blood brain barrier is a major limitation in the treatment of many brain conditions as many free drugs do not have the necessary characteristic to cross it. Transferrin/anti-transferrin receptor antibodies complexes have demonstrated to cross the blood brain barrier and modified versions of it might be developed as carriers for brain therapeutics. (Burdo & Connor, 2003; Dufes et al., 2013).

The demonstration that HuHF binds to cells through the TfR1, puts ferritin on the spotlight for the development of next generation cancer therapeutics. The ferritin's natural tropism to TfR1 can be exploited to design molecular trojan horses against cancer. Upon binding, HuHF follows a similar pathway to that of Tf. It undergoes endocytosis and is directed to endossomes and lysosomes, where it is able to release its iron content in sufficient quantities to alter the gene expression dependent of iron. Further studies defining the fate of HuHF after it reaches the lysosome will help to develop strategies for controlled cargo delivery (Li et al., 2010). Nevertheless, it is now clear that when ferritin is uptaken and internalized by TfR1, the complex does not necessarily follows the typical pathway sketched in Figure 1.2 and further

branching of the intracellular pathway have been proposed, including targeting to the nucleus (Alkhateeb & Connor, 2010; Surguladze et al., 2005; Thompson et al., 2002).

Other than TfR1, other receptors have been fount to bind ferritin. Li et al found that mouse scavenger receptor class A member 5 (Scara5) was responsible for binding and endocytosis of serum ferritin, in mice. Scara5 binds preferentially to murine L-ferritin, and it plays an important role in kidney organogenesis (Li et al., 2009). Han et al describe the specific and saturable binding of ferritin to T cell immunoglobuline and mucin domain containing receptor 2 (TIM-2), on mice (Han et al., 2011). TIM-2 has been found to be overexpressed in oligodendrocytes. These cells have increased iron needs for the production of mielin and authors hypothesized that TIM-2 functions as an iron uptake mechanism (Todorich et al., 2008). However both Scara5 and TIM-2 receptors are not expressed in humans, and lack of evidence that either murine or human ferritin bind to murine TfR1 suggests that the underlying recognition mechanisms have diverged between humans and mice (Han et al., 2011; Li et al., 2010).

The recent advances in understanding of the ferritin uptake and its fate in many cellular models strongly reinforce the potential of using ferritin-based DDS. In addition to the specific and selective binding to cancer cells, ferritin nanoparticles dimensions are ideal to tumor targeting through enhanced permeability and retention (EPR) effect, where the abnormal vascular growth in tumors increases the likelihood of accumulation of nano-sized particles (Lee et al., 2016).

Collectively, broad knowledge in cancer biology and promising results obtained in conjugation/encapsulation of cancer drugs with ferritin provide rationale to pursue the development of ferritin-based therapies.

1.2.3 Other applications

Although ferritin has showed enormous potential as *in vivo* DDS for cancer. However, the ferritin scaffold can be used for many other applications. In this section a few selected examples of ferritin used in biotechnology are described.

Industrial phosphate and arsenate removal from water

The use of *Pyrococcus furiosus* ferritin for the removal of phosphates from water is one of the very few industrial-scale uses of ferritin (Jacobset al., 2010). It has been developed as an industrial scale solution to prevent the eutrophication of water as a consequence of excessive growth of algae and aquatic plants whose limiting nutrient is inorganic phosphate. The current available methods for phosphate removal rely on the precipitation of phosphate using calcium and aluminum sulfates or biological methods where bacteria are used to sequester calcium. Later on, Sevcenco et al. expanded this approach to the removal of arsenate from water, a problem taking considerable proportions in India and Bangladesh (Sevcenco et al., 2015). Ferritin has the the ability to sequester phosphate and arsenate through the iron oxides and hydroxides within its core. The adsorption of phosphates and

arsenate can occur either concurrently with the iron core formation or after its completion. *P. furiosus* ferritin seems specially adequate for this purpose since it can be easily and abundantly expressed recombinantly, making it suitable for large-scale production, and since it is endowed of remarkable thermostability due to the fact that it is isolated from a thermophilic organism. This approach shows great potential to reduce costs, nevertheless its wide spread applicability still remains to be assessed (Jacobs et al., 2010).

Ferritin for medical imaging

Ferritin has shown to be and ideal tool for cellular imaging, as labeled heavy atoms can be readily sequestered within its core. Furthermore, iron loaded ferritin has been used as contrast agent in both electron microscopy and MRI. Ferritin fulfills the requirement of a good cellular label for electron cryotomography (cryo-ET) since its iron core provides significantly higher contrast than the cellular background. Moreover, inclusion of a fluorescent moiety in the ferritin allows to combine fluorescence microscopy and cryo-transmission electron microscopy (cryo-TEM) imaging techniques (Wang et al., 2011). For MRI, a powerful tool in both clinical and research context, the superparamagnetic properties of the iron oxides and hydroxides found in the ferritin core provide the necessary dark contrast in MRI images. Endothelial tumor cells were visualized by MRI using recombinant gadolinium-loaded ferritin bearing a peptide targeting a protein whose expression is limited to cancer cells (Geninatti Crich et al., 2015). This strategy can be extended for imaging of a many other targets simply by changing the moiety responsible

for the targeting. Ferritin-based imaging probes recall the properties of liposomes and other vesicle-like system, with the advantage of having a smaller size which improves its accessibility to targets outside the bloodstream.

Ferritin as biomineralization template

The use of biomolecules for the synthesis of inorganic nanomaterials is an emerging trend. Biological systems are unique in their versatility and nanoparticle size, shape, morphology, can be controlled by modifying interaction between inorganic materials and biomolecules. Moreover, biotemplate-directed synthesis can be more "green" than traditional chemical reactions as they usually require less harsh conditions of pH and temperature and use less often organic solvents.

By taking advantage of the physiological role of ferritin to store iron in the form of a ferrihydride core, many have employed ferritin as size-constrained reaction environment for the synthesis of different types of nanoparticles – metals, oxides, hydroxides, carbonates, semiconductors. For example, Fan et al. have described a new two-step method for the synthesis of gold nanoparticles in horse spleen ferritin (HoSP), using the natural negatively charged character of ferritin's cavity to form clusters than nucleate nanoparticle growth within the protein shell (Fan et al., 2010).

Ferritin in vaccine design

The structural resemblance of ferritins to viral capsids makes them an attractive target for vaccine development. An elegant example is the work carried by Kanekiyo et al., where the protein heamagglutinin (HA) from influenza virus was fused to the surface of *Helicobacter pylori* ferritin. These decorated ferritins were able to elicit the production of broadly neutralizing antibodies with enhanced potency when compared to a commercially available vaccine. The use of synthetic nanoparticle vaccines, entirely recombinantly produced, could eliminate the need to work with live viruses in culture, reducing both risks and costs (Kanekiyo et al., 2013).

1.3 Assembly of protein cages

Proteins are often referred as building blocks, puzzle pieces or LEGO bricks. These analogies refer to the modular nature of proteins, whose functions are tightly related to their structure. Nature, similarly to engineering, tends to combine smaller parts, or modules, to construct larger, more complex objects: nature reuses instead of reinventing. Biological systems use modularity at many levels, ranging from the molecular level up to the formation of complete organisms (Daviset al., 2002). A striking example is provided at the levels of organization of protein structure: at the primary level, by combining the same basic twenty-one aminoacids in different ways, one obtains defined secondary structures such as α-helices and β-sheets, which in turn organize in specific tertiary level structural domains such as zinc-finger domains, four-

helix bundles or globin folds. To sum up, from the same basic blocks it is possible to make virtually all the proteins necessary for function and thriving of life. Moreover, in multidomain proteins, smaller domains with varying sizes and functions are joined in different combination to achieve different functions. Development of “modules” simplifies and accelerates the evolution of biological entities (Moore et al., 2008). Despite the usefulness of the LEGO analogy in explaining a complicated molecular process, it is important to state that it is an over simplification of the complexity of proteins three-dimensional structure, overlooking fundamental properties such as the plasticity, softness and dynamics that are essential for protein function (Levy, 2017).

Protein cages can be seen as a combination of smaller modules having a self-assembly behavior which permits the encapsulation of a variety of cargoes. These modules can be chemically and genetically modified giving them new, unique properties (Maassen et al., 2016). Protein assembly comprises the formation of secondary, tertiary, quaternary and supramolecular structure. Nevertheless, when considering protein-cages, self-assembly usually refers to formation of quaternary structure and the understanding of the fundamental biochemical and biophysical processes governing protein-cage self-assemble is essential for the design of tailored cages for a wide range of purposes.

Self-assembly of proteins in complexes has a crucial role in biology. Some complexes self-assemble by design, such as naturally occurring protein cages. Others assemble as an unwanted consequence, as in the case of the amyloid β-protein aggregation found in Alzheimer’s disease (Hardy &

Selkoe, 2002), or the single-point mutation of hemoglobin causing its polymerization and underlying the pathogenesis of sickle cell anemia (Noguchi & Schechter, 1981).

In recent decades, with the arising of better and faster methods for determination of protein atomic structure, much has been improved in the understanding of the principles of quaternary structure assembly. Such that our conception of protein complexes has shifted from static representations to representations where dynamic aspects of quaternary structure are accounted for, including conformational changes upon binding, multistep ordered assembly pathways, and structural fluctuations occurring within fully assembled complexes (Marsh & Teichmann, 2015).

For any protein complex with more than two subunits, assembly becomes more and more complex as it is no longer a simple binary interaction. An analogy can be drawn from Levinthal's folding paradox (Zwanzig et al., 1992) – if a protein were to attain its folded state by sequentially sampling all the possible conformations, it would require an enormously long time, yet most proteins fold in the micro to millisecond time scale – whereby complex formation is likely to occur via an energetically favorable pathway, or, similarly to what has been proposed for protein folding, via multiple pathways exploring a funnel-like energy landscape (Dill & Chan, 1997).

Protein-cages can be described as particular case of homomeric (for most cases, assembly units can be simplified to a homomer) protein complexes, arranged with a high degree of symmetry in a spherical macro-structure. These can generally be classified as belonging to tetrahedral, octahedral, or

icosahedral symmetry group. Tetrahedral homomers have twelve subunits with two-fold and three-fold symmetry axes. Octahedral homomers have twenty-four subunits with two-fold, three-fold, and four-fold symmetry axes. Icosahedral complexes have sixty subunits with two-fold, three-fold, and five-fold symmetry axes (Marsh & Teichmann, 2015).

Much has been done in the study of self-assembly of protein cages. Specifically viral capsids, since its self-assembly can be triggered upon mixing monomers with nucleic acid, which makes them ideal model systems. Experiments and simulations have shown that capsid assembly can be divided in two phases: nucleation and growth. In nucleation a critical nucleus – smallest intermediate that has more than 50% probability of growing into a complete capsid before disassembling – is formed, followed by a growth phase where remaining subunits are added sequentially (Prevelige et al., 1993). For both processes the loss of translational entropy during self-assembly is balanced by gains from specific hydrophobic, electrostatic, van der Waals, and hydrogen-bonding interactions (Perlmutter & Hagan, 2015). Smaller intermediates are unstable and transient formation of critical nucleus is a rare event, while in the growth phase intermediates are relatively stable and addition of subunits is independent, thus the timescale to achieve full assembly has a low order dependence on the free subunit concentration (Hagan & Elrad, 2010). Due to the geometry of the protein shell, the first few intermediates have relatively few subunit-subunit contacts and are relatively unstable. The critical nucleus geometry maximizes the number of interactions

and often corresponds to a small polygon and subunit conformation changes may provide additional stabilization upon polygon formation (Hagan, 2014).

Although many experiments have provided vast information on the kinetics of capsid assembly, detailed characterization of pathways remains challenging since intermediates a unstable and short lived (Zlotnick et al., 1999). Some inventive approaches have been investigated, such as the works carried by Zhou et al. where restive pulse sensing was used to probe the passage of Hepatitis B virus (HBV) capsids through conical nanopores in a membrane and successfully distinguished between two different final assembled states (Zhou et al., 2011).

Upon assembly, capsid can undergo maturation, which involves conformational rearrangements or even covalent modifications that mostly increase the stability of the final assembled particle. Nevertheless, it remains a dynamic structure able to undergo conformational changes required to attain its intended biological function (Mateu, 2013).

1.3.1 The ferritin self-assembly

The self-assembly mechanism of protein supramolecular complexes is a challenging subject especially if the structures are highly symmetrical and homo-oligomeric. In general, the characterization of thermodynamically and kinetically accessible intermediates for the association and disassembly pathways can be considered as a pre-requisite in order to achieve a mechanistic insight. In practice, these studies are complex due to difficulties

with the determination of folding and assembly rate-limiting steps and coupling between the two. Inspection of the threedimensional structures of ferritins suggests that stable dimers are the most likely candidates as assembly intermediates due to the extensive intersubunit contacts that suggest higher interaction energies with respect to other oligomers. Sedimentation velocity analytical ultracentrifugation on horse heart ferritin as a function of pH not only demonstrated the stability of dimers in HoSF (Stefanini et al., 1987) but also suggested that assembly entails the formation of other discrete oligomeric species such as tetramers and octamers. Other investigations on homomeric H type ferritins confirmed that the dimer is the most highly populated species but suggested that other discrete species such as hexamers, in small amounts, and dodecamers are also formed during the assembly pathway (Sato et al., 2015). These results led to a refined model where the 24-meric cage assembles from dimer via tetramers and hexamers. Recently, a complete quantitative data analysis of SAXS measurements indicated that apoferritin can undergo stepwise disassembly through several structural intermediates below pH 3.40 (Kim et al., 2011). The dissociation process starts with hollow spherical structures with two holes, followed by "headphone"-shaped structures, and lastly rod-like oligomers (mainly trimers) or monomers. The structural recovery of the intermediates during the pH-induced reassembly process was proposed to be dependent on the history of the disassembly process, and that hole defect (22-mers or 23-mers) could never be recovered back to the intact 24-mers. How this data relates to proposed mechanisms of assembly has yet to be explored (Zhang & Orner, 2011).

1.4 Future prospects

The current research on ferritin assembly mechanism and its nanotechnological and biotechnological applications is currently exploding as witnessed by the increasing number of experimental and theoretical observation published recently. Major efforts are devoted to the understanding of the physical basis of the complex assembly mechanism on one side, and on the growing attempts to set up a more efficient scaffold for payload encapsulation on the other. The key advantage of ferritins as nanocarriers over other systems are well known: uniform and homogeneous size distribution of the nanoparticle, easy manufacturing procedure, unmatched overall thermodynamic stability and easy functionalization on both the internal and external surfaces. Under these premises, it is envisaged that ferritin based technologies will rapidly expand in wide nanotechnological applications in many diverse fields, spanning from biomedicine to nanoengineering, nanochemistry and smart materials. The biomedical applications, in particular, are gaining enormous popularity in view of the unique property of ferritin based nanocarriers, i.e., the targeting capability to the TfR1 receptor, a key target for tumor delivery. However, many reports have already pointed out that ferritins might also be engineered to carry alternate epitopes for the recognition of specific cell surface receptors other that TfR1. Such versatility will possibly expand the scope of therapeutic application thus transforming ferritins in universal, broad range scaffolds for cellular targeting. As a last mention, one may infer that the

advantage of having the possibility of 24 epitopes on the same molecule would greatly enhance the possibility of having stronger interactions with the desired targets even over antibodies. Research efforts in this sense are only at the early stage, but promise a revolution in drug delivery in the near future.

2 Aim of the work

The work presented in this thesis aimed to set out the basis for the rational design of innovative protein-based carriers for drug-delivery and biotechnological applications. In this context, ferritin stood out as promising protein system due to its remarkable characteristics. Ferritins are versatile biocompatible proteins scaffolds that display a cage-like structure that provides shielding of the cavity's content from harsh external conditions, and are amenable to modifications in a relatively straightforward manner.

Previously published information suggested that ferritins cages present a unexpected degree of plasticity in that they are able to undergo significant structural rearrangements in order to host small molecules within the internal cavity, a process referred to as "encapsulation". However, in order to clarify the process of molecule confinement within ferritin the internal cavity, a critical analysis of the process of ligand entry/release through the protein matrices of ferritins of diverse origin was deemed necessary.

The current project thus focuses on the design of ferritin-based nanocarriers, that exploit the convenient properties of archaeal ferritins. These proteins displays unique assembly properties and extreme thermodynamic stability, however lack the cell recognition properties of human ferritin. Thus we designed a chimeric protein in which the relevant recognition sequence of human ferritin was grafted into the corresponding sequence into archaeal ferritins surface exposed regions in order to confer specific recognition of human epitopes. The development of such innovative ferritins nanocages aimed to generate universally applicable ferritin nanocarriers for biomedical

and biotechnological purposes. The assembly properties of the novel archaeal-human chimeric ferritin nanocarrier was extensively investigated in this work with the aim of greatly improving the current methods of probe encapsulation within ferritins. These studies can thud provide key elements for the development of strategies for efficient encapsulation/release of probes in a controlled way.

On a different approach, this work aimed to exploit the versatility of ferritins for bioimaging applications, by inserting an extra functional segment for binding of luminescent lanthanides, within a mammalian ferritin's cavity. These systems could be used for advanced cell imaging applications, merging the recognition capabilities of ferritins with the notable properties of lanthanide-based fluorescence.

3 Materials and methods

3.1 Archaeal ferritin's permeation properties

Point mutations and protein expression

The genes encoding for bacterial ferritin from *A. fulgidus* ferritin (AfFt) and *P. furiosus* ferritin (PfFt) were cloned into the expression vector pET22b (Novagen). Point mutants AfFt-M54C, AfFt-M54C/K150A/R151A, PfFt-G52C and PfFt-P77C were obtained by polymerase chain reaction (PCR) using QuickChange Mutagenesis kit (Stratagene). The recombinant plasmids were transformed into *E. coli* TOP 10 cells and the resulting colonies were screened by DNA sequencing. Plasmids bearing the desired mutations were transformed into BL21(DE3) *E. coli* strain for protein expression. For each mutant, protein over-expression was obtained as it follows: 1 L Luria-Bertani (LB) broth medium was inoculated with 2 mL overnight culture of a single colony and the gene expression was induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) when the absorbance at 600 nm reached 0.6. Cells were harvested by centrifugation after overnight induction at 37 °C and the cell pellets were stored at -20 °C.

Protein purification

Harvested cells from 1 L culture over-expressing AfFt-M54C and AfFt-M54C/K150A/R151A mutants were resuspended in 20 ml buffer (20 mM HEPES pH 7.4, 300 mM NaCl) containing a cOmplete™ Mini Protease

Inhibitor Cocktail Tablet (Roche) and disrupted by sonication. The soluble fraction was thermally purified by heating at 85 °C for 10 minutes followed by removal of denatured proteins by centrifugation at 11000 rpm for 30 minutes at 4 °C. The supernatant was fractioned by ammonium sulfate precipitation. 70% ammonium sulfate pellet containing highly purified protein was resuspended in 20 mM HEPES pH 7.4, 20 mM $MgCl_2$, dialyzed versus the same buffer, sterile filtered and stored at 4 °C. Cells over-expressing PfFt-G52C and PfFt-P77C were sonicated in 20 mM HEPES buffer at pH 7.4, 0.3 M NaCl and cOmplete™ Mini Protease Inhibitor Cocktail Tablet. After sonication, the crude bacterial extract was digested with DNase for 1 hour at 37 °C, heated at 80 °C for 10 minutes followed by centrifugation to remove insoluble material. Then, the supernatant was further purified by ammonium sulfate precipitation. 70% ammonium sulfate pellet was resuspended in 20 mM HEPES pH 7.5, 150 mM NaCl, dialyzed versus the same buffer and loaded onto a HiLoad 26/600 Superdex 200 pg column (GE Healthcare). Fractions containing highly purified protein were pooled, sterile filtered and store at 4°C.

Preparation of ferritin-DTNB adducts

All mutants were reduced with 3 mM tris(2-carboxyethyl)phosphine (TCEP) in their storage buffers and then loaded onto a desalting column (GE Healthcare) to remove the reducing agent. Each mutant was reacted with 40-fold molar excess of Ellman's Reagent, 5.5'-dithio-bis-(2-nitrobenzoic acid) (DTNB) per cysteine for 3 hours at room temperature. Stock DTNB solutions

were prepared in ethanol. The excess (non-reacted) reagent was removed by ultra-filtration using 100 kDa Amicon Ultra-15 centrifugal devices (Millipore Corporate).

Stopped flow experiments

Kinetic measurements were carried out on a thermostated Applied Photophysics stopped-flow apparatus (Leatherhead, UK) by mixing 8-10 μM protein solutions, previously reduced with TCEP, with solutions containing different concentrations of DTNB (from 0.2 to 0.7 mM after mixing) in 20 mM HEPES, 20 mM $MgCl_2$ pH 7.5. In order to avoid interference of the instrument phototube from the high concentrations of DTNB and the released chromophore 5-thio-2-nitro-benzoic acid (TNB), the reaction was followed at 430 nm and the extinction coefficient calculated to be 12205 $mM^{-1}cm^{-1}$, as determined from the extinction coefficient of 14150 $mM^{-1}cm^{-1}$ at 412 nm (Riddles et al., 1979). All fitting procedures were carried out by using Matlab software (Mathworks, USA). Experimental traces were fitted by non-linear regression to either exponential or biexponential processes by using an Levenberg-Marquardt algorithm.

Self-assembly study

$MgCl_2$ mediated self-assembly of ferritin mutants was confirmed by incubating aliquots of proteins (1 mg/ml) with different salt concentrations in 20 mM HEPES buffer, pH 7.4. Molecular sizes of AfFt-M54C, AfFt-M54C/

K150A/R151A, PfFt-G52C and PfFt-P77C were determined by size exclusion chromatography (SEC) using HiPrep 16/60 Sephacryl S300 column (GE Healthcare). The column was equilibrated with 20 mM HEPES pH 7.4, 20 mM $MgCl_2$. Molecular weight of each mutant were determined by comparing their elution volumes with the elution volumes of standard proteins at the same salt concentration.

Liquid Chromatography – Mass Spectrometry

Liquid Chromatography – Mass Spectrometry (LC–MS) was performed on protein samples before and after DTNB titration, after dialysis in distilled water in the presence of 0.1 mM ethylenediaminetetraacetic acid (EDTA), using a Waters AcquityuPLC connected to Waters Acquity Single Quad Detector. A Hypersil Gold C4 column was used: 1.9 μm, 2.1 × 50 mm at 254 nm observation wavelength; mobile phase: 95:5 water (0.1% formic acid):MeCN (0.1% formic acid); gradient over 6 min (to 5:75 water (0.1% formic acid):MeCN (0.1% formic acid); flow rate: 0.4 ml min^{-1}; MS mode was set at a scan range: m/z=250–2,000 (ES+); scan time: 0.25 s. Data were obtained in continuum mode by setting the electrospray source of the MS with a capillary voltage of 3.5 kV and a cone voltage of 50 V. N_2 gas was used as nebulizer and desolvation gas at a total flow of 300 l/hours. Ion series were generated by integration of the UV signal at 254 nm chromatogram over 1.2–1.8 min range. Mass spectra were subsequently reconstructed for proteins from the ion series using the MaxEnt 1 algorithm on MassLynx software program.

3.2 Humanized archaeal ferritin

"Humanized" Archaeoglobus ferritin design

The gene encoding for a mutated ferritin from *A. fulgidus* was synthesized by GeneArt (ThermoFisher) and sub-cloned into a pET22b vector (Novagen) between the restriction sites NdeI and HindIII at 5′ and 3′ respectively. The recombinant plasmid was transformed into BL21(DE3) *E. coli* strain for protein expression.

Protein expression and purification

E. coli cells, containing the Humanized *A. fulgidus* ferritin (HumAfFt) plasmid, were grown and induced with 1 mM IPTG at OD600 = 0.6. Cells were harvested by centrifugation 18 hours post induction at 37 °C. Cells harvested from 1 L culture were resuspended in 20 mM HEPES buffer, pH 7.4, containing 300 mM NaCl, and a cOmplete™ Mini Protease Inhibitor Cocktail Tablet (Roche). Cells were disrupted by sonication and the soluble fraction was purified by heat treatment at 80 °C for 10 minutes. Denatured *E. coli* proteins were removed by centrifugation at 11000 rpm at 4 °C for 30 min. The soluble protein was further purified by ammonium sulfate precipitation. The precipitated fraction at 70% ammonium sulfate was resuspended in 20 mM HEPES, 50 mM $MgCl_2$, pH 7.4 and dialyzed versus the same buffer. As a final purification step, the protein was loaded onto a

HiLoad 26/600 Superdex 200 pg column previously equilibrated in the same buffer using an ÄKTA-Prime system (GE Healthcare). The purified protein was concentrated to obtain the final protein preparation and protein concentration were routinely calculated by measuring the UV spectrum using an extinction coefficient of 32400 M^{-1} cm^{-1}.

Self-assembly assessment in solution

SEC $MgCl_2$-mediated self-assembly was studied using a Superdex 200 26/600 GL column (GE Healthcare). The molecular size of HumAfFt was determined under different conditions by comparing the elution volume with that of standard proteins. Composition of the mobile phase was 20 mM HEPES pH 7.4 with different $MgCl_2$ concentrations accordingly to the experiment requirements.

Crystallization and crystal structure determination

The purified protein was concentrated to 20 mg mL^{-1} and initial crystallization screening was performed using a Phenix Robot. Crystals were obtained by mixing in a 2 µL hanging drop of the purified protein with a solution containing 22% (vol/vol) polyacrylic acid (PAA), 0.1 M Tris, 0.02 M $MgCl_2$, pH 7.4, at 25 °C within a week, cryo-protected by increasing the precipitant concentration and flash-frozen in liquid nitrogen for transport. Diffraction data have been collected at ID23-2 beamline at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. Data were

processed with XDS (Kabsch, 2010) and scaled with Aimless (ccp4 suite) at a final resolution of 2.87 Å. The structure was solved by Molecular Replacement with MolRep (ccp4 suite) using the wild type structure AfFt (PDB code 1S3Q) as the search model (Winn et al., 2011). Model Building and refinement were done using Coot (Emsley & Cowtan, 2004) and Refmac5 (Murshudov et al., 2011a), respectively. The final model was analyzed with PROCHECK (Laskowski e al., 1993) and Molprobity (Chen et al., 2010). The Ramachandran plot showed that 97.8% of residues are in preferred regions, 2.2% in allowed regions and no outlier is observed. The final atomic coordinates and structure factors were deposited with the PDB Data Bank (http://www.rcsb.org) with accession code: 5LS9.

Cryo-Electron microscopy

Holey-gold grids were prepared as described before (Russo & Passmore, 2014) from Quantifoil R1.2/1.3 (Quantifoil Micro Tools GmbH, Germany). 3 μL of HumAfFt was applied to the holey-gold grids after plasma cleaning with a mixture of H_2 and O_2. Grids were blotted for 4 seconds and vitrified by rapidly plunging into liquid ethane at -180 °C with a Vitrobot. Data acquisition was done using a FEI Titan Halo (FEI, Eindhoven) operating at 300 kV. Datasets were imaged with a Volta phase-plate (Danev et al.,, 2014) (FEI, Eindhoven) and were collected with the automated data collection system EPU (FEI, Eindhoven) at a nominal magnification of 59000× on an FEI Ceta camera (FEI, Eindhoven) with a camera pixel size of 14 μm, corresponding to a calibrated pixel size of 1.49 Å on the specimen scale and

with a dose of 50 e^{-} $Å^{-2}$. Images were processed with the reference-based automated particle picking procedure implemented in RELION 1.3 (Scheres, 2012) . Contrast transfer function (CTF) correction was not applied since the data were collected within 200 nm of focus and the first CTF zero crossing was well beyond the achievable resolution of the dataset. Those particles were subjected to 2D classification using RELION with k = 100 classes. Good particles were then subjected to 3D classification using RELION with the number of classes K = 8. Resulting classes were refined with the autorefine procedure in RELION. Reported resolution is based on the 'gold-standard' protocol with the Fourier shell correlation (FSC) = 0.143 criterion using soft masks with an 8 pixel soft edge, and were corrected for the effects of the mask on the FSC curve using high-resolution noise substitution (Chen et al., 2013).

Protein FITC labeling

HumAfFt, AfFt and holo-transferrin were labeled with fluorescein-isothiocyanide (FITC) according to the manufacturer's standard protocol. Briefly, 2 mg mL^{-1} of the purified protein was added with 10-fold molar excess of in protein storage buffer stirring for 4 hours at room temperature. The nonreacted dye was removed by gel filtration chromatography and the fluorescent dye to protein ratio was determined by UV-spectroscopy. LC-MS spectrometry measurements on HumAfFt-FITC confirmed that >60% of monomers are FITC labeled.

Cell cultures and ferritin internalization

HeLa cells were grown at 37 °C in Eagle's MEM supplemented with 10% (v/v) fetal bovine serum (FBS), Glutamax (Invitrogen) and penicillin–streptomycin solution (Sigma). When needed, the cells were induced with doxycycline 0.2 μg mL^{-1}. The internalization assay was performed as it follows: after seeding the cells on the relevant substrate depending on the experiment, cells were left one day to attach and then incubated with FITC-ferritin nanoparticles (AfFt-FITC, HumAfFt-FITC or Tf-FITC as specified in each experiment) at the final concentration of 30 μg ml^{-1} for the time indicated (1 h, 3 h or 20 h).

For evaluation of the TfR1 silencing effect on HumAfFt uptake by Hela cells, cells were transfected with anti-TfR1 siRNA (Sigma-Aldrich) with Lipofectamine® RNAiMAX (Life Technologies) following standard procedure. Cells were then incubated with 30μg/ml of Hum AfFt-FITC, treated and analyzed as described in the methods session. The FITC intensity was normalized for the FITC intensity of the scr samples.

Flow cytometry analysis

For flow cytometry analysis, or fluorescence activated cell sorting (FACS), HeLa cells were seeded on multiwell plates. Cells were incubated with FITC-ferritin nanoparticles as described previously, then washed two times with phosphate-buffered saline (PBS), detached with trypsin-EDTA (Euroclone),

washed with PBS and resuspended in BD-FACS flow buffer. Half of each sample was treated with Trypan Blue (TB) to quench the FITC signal from membrane-bound nanoparticles that were not internalized. The quenching was performed with 0.04% TB for 10 min, on ice. Control cells were treated in the same way but without FITC-ferritin incubation. Internalization of ferritins before and after TB treatments was measured at the BD LSFORTESSA (BD Biosciences, San Jose, CA, USA) equipped with a 488 nm laser and FACS Diva software (BD Biosciences version 6.1.3). Live cells were first gated by forward and side scatter area (FSC-A and SSC-A) plots, then detected in the green channel for FITC expression (530/30 nm filter) and side scatter parameter. The gate for the final detection was set in the control sample. Data were analyzed using FlowJo 9.3.4 software (Tree Star, Ashland, OR, USA).

Confocal microscopy of live cells

To visualize ferritin internalization by live cells using a confocal microscope, HeLa TagRFP cells were seeded on a μ-slide 8-well ibiTreat (ibidi) and induced with 0.2 μg mL^{-1} of doxycycline. Cells were then incubated with FITC-ferritin nanoparticles as previously described, for 20 hours and, before microscopy, cells were washed two times with an imaging medium (DMEM without phenol red, 10% FBS, 10 mM Hepes, Glutamax and penicillin–streptomycin solution) to eliminate the unbound FITC-ferritin nanoparticles. The confocal laser-scanning microscope used was an Olympus FV10i platform equipped with a built-in incubator. Images were acquired with a

60×/ 1.2NA water-immersion objective, LD lasers, 473 nm and 559 nm, and filter sets for FITC.. Phase-contrast images were acquired simultaneously.

Dynamic Light Scattering

Dynamic Light Scattering (DLS) experiments were carried out with a Zetasizer Nano S (Malvern Instruments, Malvern, U.K.) equipped with a 4 mW He-Ne laser (633 nm). Measurements were performed at 25 °C, at an angle of 173° with respect to the incident beam. Peak intensity analyses were used to determine the average hydrodynamic diameters (Z-average diameter) of the scattering particles. All samples were prepared at 2 mg/mL in HEPES 20 mM pH=7.4 and filtered with a 0.2 µm filter before analysis.

Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were carried out using an iTC200 microcalorimeter (MicroCal). HumAfFt was prepared in 20 mM HEPES pH 7.4 and an $MgCl_2$ stock solution was prepared by dissolving the powder into the same buffer solution to ensure buffer-matching between the protein and the titrant. 1.5-µl aliquots of 4 mM $MgCl_2$ were injected into a 15 µM protein solution at 25°C (each injection was time-spaced for 240 sec). Data were fitted using a "one-binding-site model" of the MicroCal version of ORIGIN. The heat of binding (ΔH), the stoichiometry (n), and the dissociation constant (K_d) were then calculated from plots of the heat evolved per mole of ligand injected versus the molar ratio of ligand to protein.

Small Angle X-ray Scattering

Small Angle X-ray Scattering (SAXS) experiments were carried at the ESRF at the ID02 and BM29 beamlines. X-ray wavelength was set at 0.99 Å (12.5 KeV) and temperature kept at 20 °C in all experiments. two-dimensional scattering images were obtained using Rayonix MX170-HS detector and a 1M Pilatus detector, for ID2 and BM29 respectively. Static measurements at BM29 were obtained using using a 1.7 mm capillary, while flowing the sample, with 0.5 seconds exposure time. For time-resolved experiments, sample was mixed using BioLogic SFM stopped-flow SFM400 apparatus, equipped with a quartz cell of 1.4 mm diameter.

The self-assembly reaction of HumAfFt was triggered by adding 20 mM HEPES pH 7.4 with $MgCl_2$ at various concentrations with a mixing ratio of 1:1. Time-resolved scattering data was obtained with 5 millisecond exposure and an average of thirty time-points were collected per mixing. Reaction time-course was obtained by combining different time-delays, and 10 to 15 mixings were averaged to increase signal to noise. For longer time intervals, shutter was closed between acquisitions. X-ray damage was evaluated with static measurements to guarantee its absence from time-resolved collected data.

Two-dimensional images were circularly integrated to obtain one-dimensional scattering profiles. After subtraction of the scattering profile of the buffer solution, the radius of gyration, R_g, and the forward scattering intensity, I_0, were calculated using the Guinier approximation, where the scattering vector $q=(4\pi/\lambda) \sin \theta$ (λ is the wavelength and 2θ is the scattering

angle) and I_q is the scattering intensity at a given q value. The I_0 and R_{app} values were calculated from the intercept and slope of the Guinier plot in the q range satisfying $q.R_g \leq 1.3$.

3.3 Lanthanide binding ferritin

Protein design, expression and purification

A synthetic gene encoding for mouse H chain ferritin (MuHF) fused with a lanthanide binding peptide (LBT) was designed, synthesized, and optimized for *E. coli* codon usage by Geneart (Geneart AG). LBT sequence YIDTNNDGWIEGDELLA(Martin et al., 2007) was added to the C-terminal of MuHF generating MuHF-LBT construct, that was subcloned into pET22-b vector (Novagene) between NdeI/XhoI restriction sites. MuHF-LBT was overexpressed in E. coli BL21 cells upon induction with 1 mM IPTG at OD_{600nm} = 0.6. Cells were harvested by centrifugation 16 hours post induction at 37°C. Cells from 1 L culture were resuspended in 20 mM HEPES pH 7.5, containing 300 mM NaCl, 1 mM TCEP, and a cOmplete™ Mini Protease Inhibitor Cocktail Tablet (Roche). Cells were disrupted by sonication and the soluble fraction was purified by heat treatment at 80°C for 10 minutes. Denatured *E. coli* proteins were removed by centrifugation at 15000 rpm at 4°C for 1 hour. The soluble protein was further purified by ammonium sulfate precipitation. The precipitated fraction at 70% ammonium sulfate, was resuspended in 20 mM HEPES, 50 mM $MgCl_2$, pH 7.5 and dialyzed versus the same buffer. As final purification step, the protein was loaded onto

a HiLoad 26/600 Superdex 200 pg column previously equilibrated in the same buffer using an ÄKTA-Prime system (GE Healthcare). Purified protein was concentrated to obtain the final protein preparation of 1 mg/mL and protein concentration was calculated by measuring the UV absorption spectrum using an extinction coefficient of 32400 M^{-1} cm^{-1} . The expected molecular weight of 22662 Da was confirmed by MALDI-TOF Mass Spectrometry.

Fluorescence spectroscopy

Fluorescence spectra and titrations were performed using FluoroMax 4 (Horiba) spectrofluorimeter with a Haake D8 refrigerated bath at 20 °C. Emission spectra were recorded between 450 and 560 nm, in order to include the luminescent maxima of Tb(III) (490 and 545 nm). The excitation wavelength was chosen at 295 nm to minimize the overlap of second order diffraction (570 nm) with the Tb(III) emission at 545 nm. Emission spectra were taken with excitation and emission band passes of 4 and 8 nm and corrected for the blank contribution and the instrument response at 295 nm in a quartz cell of 1 cm pathlength. Emission spectra were normalized to 1 at 545 nm. Fluorescence measurements were performed using 1 μM apoMuHF-LBT and apoMuHF as a control, in 100 mM MES buffer pH 6.4. A 50 mM $TbCl_3$ (Sigma-Aldrich) stock solution was also prepared in MES buffer at pH 6.4. Fluorescence spectra were recorded upon 30 min incubation with excess of $TbCl_3$ in buffer solution followed by a washing step to remove the

unbound metal. Titration data were obtained by addition of aliquots of $TbCl_3$ in buffer solution under continuous stirring.

Cryo-Electron microscopy

Holey-gold grids from Quantifoil R1.2/1.3 (Quantifoil Micro Tools GmbH) were prepared as described (Russo & Passmore, 2014); their surfaces were treated with plasma cleaning using a mixture of H_2 and O_2 and 3 µl of a solution containing 1 µM MuHF-LBT Tb(III) complex was applied to the holey-gold grids. After 30 seconds waiting time, grids were blotted for 3 seconds at 100% humidity with filter paper and vitrified by rapidly plunging into liquid ethane at −180 °C, with a Vitrobot Mark IV (FEI). Data acquisition was performed using a FEI Titan Halo (FEI, Eindhoven) operating at 300 kV, while the specimen was maintained cold using liquid nitrogen. Datasets were collected with an automated data collection system (Suloway et al., 2005) on a K2 Summit direct detector camera (Gatan, Pleasanton) operating in super-resolution mode, with a calibrated pixel size of 1.15 Å on the object scale and a magnification of 59 000×. Images were typically recorded with a defocus range between -0.7 and -3.0 µm and a dose of electrons on the specimen plane between 10 and 20 electrons/$Å^2$. Data analysis was carried out using RELION 2.0 (Scheres, 2012).CTF correction was done using MotionCor2 (Zheng et al. 2016). Particles were picked and extracted from the original micrographs with the reference based automated particle picking procedure implemented in RELION and were 2D classified using 100 classes. 2D good classes were selected and then subjected to 3D

classification with 8 classes, using as reference model the MuHF structure (PDB code 3WNW). Resulting good classes were refined with the 3D autorefine procedure in RELION. The final 3D map resolution is calculated with the FSC = 0.143 criterion, based on the "gold-standard" protocol (Rosenthal & Henderson, 2003), using soft masks with a 4 pixel soft edge, and has been corrected for the effects of the mask on the FSC curve using high-resolution noise substitution (Chen et al., 2013). This final map was visualized using UCSF Chimera (Pettersen et al., 2004).

Crystallization and X-ray structure determination

Crystals of apoMuFH-LBT and MuHF-LBT-Tb(III) were obtained by mixing in a 2 μl hanging drop the purified protein at 15 mg/ml with a solution containing 1.8/2.0 M ammonium sulphate and 0.1 M Tris pH 8.5, at 25° C within a week. Cryoprotection of crystals was attained by extensively washing the crystals in sodium malonate followed by flash-freeze in liquid nitrogen. Diffraction data were collected at XRD1 beamline at the Elettra Synchrotron, Trieste, Italy. Data were processed with XDS (Kabsch, 2010) and scaled with Aimless (Winn et al., 2011) at a final resolution of 2.85 and 2.65 Å. The structures were solved by Molecular Replacement with Phaser (Adams et al., 2010; McCoy et al., 2007) using the structure of MuHF (PDB code 3WNW) as search model. Model building and refinement were done using Coot (Emsley & Cowtan, 2004) and Refmac5 (Murshudov et al., 2011), respectively. The final model was analyzed with PROCHECK (Laskowski et al., 1993) and Molprobity (Chen et al., 2010). Ramachandran

plot showed more than 98 % residues were in preferred regions and no outlier was observed in both structures. The validation of metal binding sites was performed using CheckMyMetal web server (Zheng et al., 2017). The final atomic coordinates and structure factors were deposited with the PDB Data Bank (www.rcsb.org) with accession code (5OBA and 5OBB, respectively).

Cell cultures and ferritins internalization

Human prostate cancer cell line DU-145 (ATCC® HTB81TM), human colorectal cancer cell line HCT-116 (ATCC® CCL-247TM), human breast cancer cell line MDA-MB-231 (ATCC® HTB-26TM) and human ovarian cancer cell line SKOV-3 (ATCC® HTB-77TM) were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing 10% FBS, 100 μg/ml streptomycin and 100 U/ml penicillin G in a humidified 37 °C incubator. The internalization assay was performed as it follows: cells were detached using trypsin, then washed twice with PBS and resuspended in non-supplemented DMEM containing FITC - ferritin nanoparticles (MuHF-LBT or MuHF as a control) at a final concentration of 0.5 mg/ml for 1 hour in a humidified 37°C incubator. After incubation cells were washed three times with PBS and subjected to confocal microscopy and flow cytometry analysis.

Confocal microscopy

Following internalization step described above, cells were seeded into 8-well Nunc™ Lab-Tek™ Chambered Coverglass with 200 μl DMEM, containing

10% FBS, 100 μg/ml streptomycin and 100 U/ml penicillin G per well. Plates with cells were then incubated on ice until visualization at the microscope. Images were acquired using an inverted confocal microscope IX70 FV 500 (Olympus), with 488 nm laser, 20x objective lens and emission filter 505-560 nm. Image processing was performed using ImageJ software (National Institutes of Health, www.imagej.nih.gov/ij/).

Flow cytometry analysis

Cells were incubated with FITC-ferritin nanoparticles as described previously then washed three times with PBS, and resuspended in FACS buffer (2% FBS, 1 mM EDTA). Internalization of ferritins before and after treatments was measured at the BD FACS AriaTM III equipped with a 488 nm laser. Cells were first gated by forward and side scatter area (FSC-A and SSC-A) plot and for singlets population (FSC-H and SSC-A), then detected in the channel for FITC expression (530/30nm filter) and side scatter parameter. The gate for the final detection was set according to the gate set on the control sample. Data were analyzed using BD FACS DIVA™ and FlowJo softwares.

4 Results and discussion

4.1 Archaeal ferritin's permeation properties

Ferritins cages of archaeal origin have been proposed as privileged scaffolds for nanotechnological applications due to their increased thermo-stability, and high yield expression in recombinant *E.coli* expression systems and, in the case of *Archaeoglobus fulgidus* ferritin (AfFt), for its uncommon structure and association/dissociation properties, that result in the display of four large triangular pores (45Å wide) in the protein shell. For all ferritins, the pores and channels on the protein shell are responsible for the entry and exit of cations during mineralization/demineralization. This characteristic has been exploited in the development of ferritin based drug-carriers and pores have been proposed as main drug loading routes. However molecular diffusion in and out of the ferritin cavity appears to be a complex phenomenon only partially understood.

A set of mutated ferritins from *A. fulgidus* and *P. furiosus* were designed to assess the accessibility of the protein cage's interior. These display significantly different quaternary structures - whereas the PfFt 24-mer has the canonical 432 point-group symmetry, the AfFt 24-mer displays a 23 point-group symmetry typical of smaller 12-mer ferritin-like proteins. This unusual assembly does not display the 4-fold channels and constrains the quaternary structure thus leading to the appearance of four large triangular openings about 45 Å wide in the protein shell. In total, four mutants were designed to bear cysteine residues in selected topological positions inside or outside the

ferritin shell: two PfFt mutants, one bearing cysteines facing the cavity – PfFt-G52C – and another with surface exposed cysteines – PfFt-P77C; two AfFt mutants, both bearing cysteines facing the cavity, one “open”, i.e., displaying the four large pores – AfFt-M54C – and one ‘closed’, where design mutations induce the canonical ferritin assembly – AfFt-M54C, AfFt-M54C/K150A/R151A (Figure 4.1).

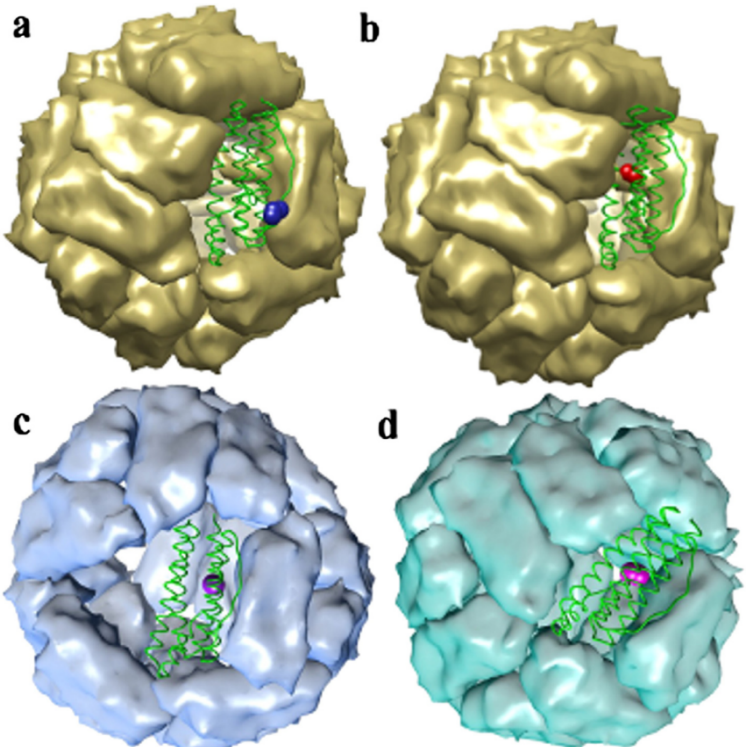


Figure 4.1 – Threedimensional structures of the ferritin mutants. A monomer is depicted as green ribbon and the cysteine residues are represented in CPK style for clarity. **(A)** Pf-FtP77C: the external cysteine is shown as blue spheres. **(B)** Pf-FtC52C: the internal cysteine is depicted in red. Models A and bB built on Pf-Ft structure (PDB code: 2JD6). **(C)** Af-Ft54C: the internal cysteine is depicted in purple. **(D)** Af-FtM54C/K150A/R151A: the internal cysteine in magenta. Mmodels C and D built on AfFt structure, (PDB code: 1SQ3). Molecular graphics and analyses were performed with the UCSF Chimera package. (Pettersen et al., 2004)

Cage accessibility was evaluated by comparing the reactivity of the thiol reactive compound DTNB to strategically located thiol groups at the mutated cysteines.

In order to assess the association state of the the engineered ferritins, DLS and SEC studies as a function of $MgCl_2$ concentration were performed. Although PfFt mutants are stable 24-mers the association state of AfFt mutants is salt dependent and $MgCl_2$ was used as polymerizing salt. At a working concentration of 20 mM $MgCl_2$, all mutants showed to be stable 24-mers, with similar SEC elution profiles (Figure 4.2)

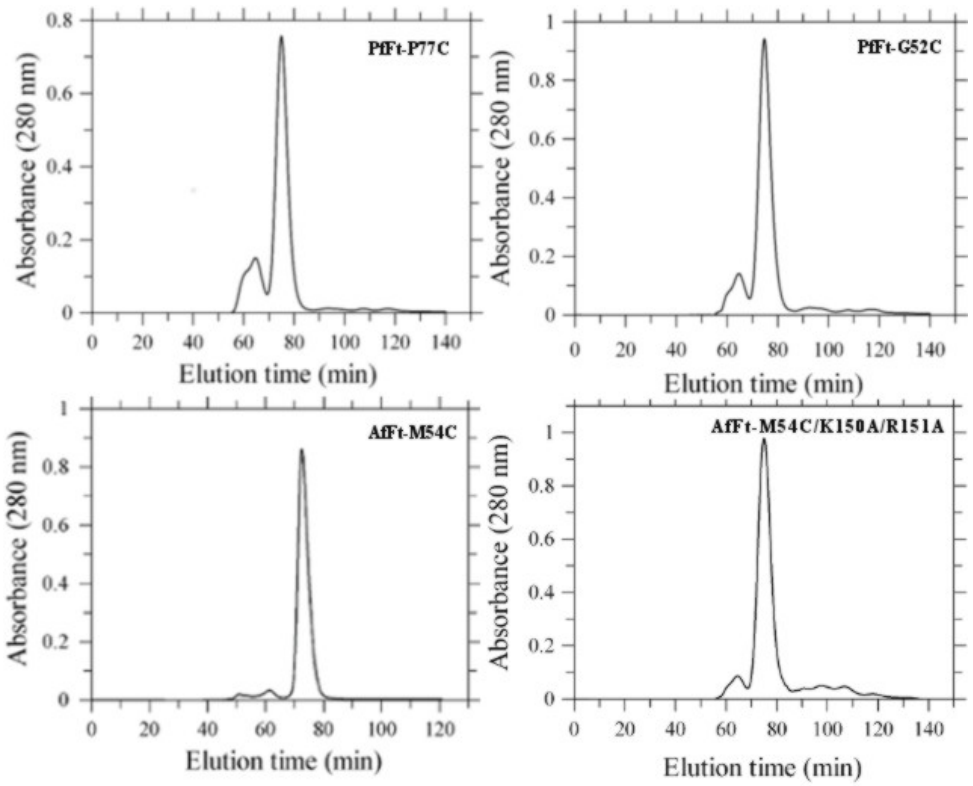


Figure 4.2 – Size exclusion chromatography profiles of ferritin mutants. The characteristic SEC profiles obtained at 20mM HEPES pH. 7.4, 20mM $MgCl_2$, for mutants PfFt-P77C, PfFt-G52C, AfFt-M54C and AfFt-M54C/K150A/R151A (right to left, top to bottom) are represented.

The kinetics of the disulfide exchange reaction of DTNB with cysteines residues were carried using time-resolved UV spectroscopy, using stopped-flow technique. Kinetic traces were obtained under pseudo-first order conditions and followed up to the maximum instrumentally achievable (1000s). As illustrated in Figure 4.3, the time scales of the four mutants differ significantly: the PfFt-P77C mutant bearing the surface exposed cysteine displayed the fastest reactivity, followed by the AfFt-M54C the "open" mutant, then the "closed" AfFt-M54C/K150A/R151A and finally the PfFt-G52C with the inner cysteine. Within the observed time frame all traces could be fit to a simple relaxation process with apparent second order rate constants as indicated in Table 4.1, with the notable exception of AfFt-M54C that was fitted by A bi-exponential process in which the fastest second-order rate constant was of the same order of magnitude as the PfFt-P77C mutant.

Table 4.1
Apparent second-order constants for ferritin mutants as determined by the DTNB reaction

Protein	Cysteine position	K_{obs} ($M^{-1}s^{-1}$)
PfFt-P77C	External	908 ±122
PfFt-G52C	Internal	5 ±1
AfFt-M54C "open"	Internal	Fast: 198 ±65 Slow: 51 ±21
AfFt-M54C/K150A/R151A "closed"	Internal	26 ±2

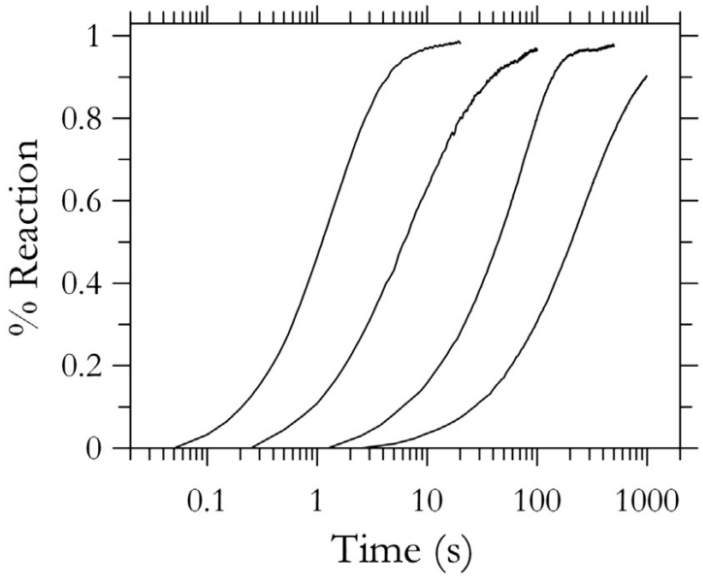


Figure 4.3 – Kinetics of DTNB binding to AfFt and PfFt mutants. Comparison of the reaction time scales of all mutants; from left to right, Pf-FtP77C, Af-FtM54C, Af-FtM54C/K150A/ R151A, and Pf-FtG52C. All mutants were 5 μM except Af-FtM54C which was 4 μM, and DTNB was 0.7 mM.

To ensure and quantify the reaction completion, all mutants were submitted to LC-MS analysis before and after titration with DTNB. In all samples reaction was completed, with the exception of PfFt-G52C in with 18% of cysteines remained unreacted.

Overall these results show that the diffusion of molecules across the protein shell can be accurately measured and that bulky molecules, such as the DTNB, can permeate the protein shell much better than expected. Even fully assembled PfFt-G52C and closed AfFt-M54C/K150A/R151A mutant are able

to bind stoichiometrically the bulky, negatively charged DTNB. This indicate that negatively charged of 8-10 Å length and 5-6 Å diameter are able to cross the protein shell, in contrast with the previously described impermeability of vertebrate ferritins to negatively charged molecular species. The charge selectivity reported in HoSF and HuHF has attributed to the nature of the threefold channels, endowed with hydrophilic, negatively charged residues (Desideri et al., 1991). Moreover, kinetic studies of permeation using small nitroxide spin probes where a negatively charged nitroxide probe was completely excluded from the cavity, and mutation of HuHF threefold lining residues from negatively charged glutamate to histidines partially allowed the entry of negatively charged species provide strong evidence of determining role of the threefold channels ad charge-selective avenues for cavity entry (Yang et al., 2000). In the case of PfFt, channels are similar to those of HuHF in the outer side, that is rich in negative aminoacids (Glu109, Glu110 and Glu111), whereas they differ in the central region due to the presence of polar and positively charged residues (Tyr114 and Arg117). Lastly, the inner opening of the channel displays Ala118 and Glu121 in the topological position which is deemed essential for the transfer of Fe(II) to the ferroxidase center in mammalian ferritins (Behera & Theil, 2014). In the case of AfFt mutants the cluster of negative charges characteristic of HuHF is not present. The threefold channels are lined by a mix of hydrophobic and hydrophilic aminoacids: only one glutamate (E131 is located at the entrance and the two acidic residues D131 and E134 lining HuHF channels are replaced by neutral and positively charged residues Y119 and N120). Presumably this might due to the fact that is not necessary to drive Fe(II) atoms towards the ferroxidase

site since they are readily available through the large pores present in the protein shell. Hence, the threefold channels of archaeal ferritins are less negatively charged than the mammalian ones, possibly allowing the entrance of negatively charged compounds. Logically in the AfFt-M54C "open" mutant, the large triangular pores are the most likely preferential paths for bulky ligand entry, as denoted by the eight-fold faster DTNB reaction rate when compared with AfFt-M54C/K150A/R151A.

Thus, on the basis of the observed DTNB binding rates and available crystallographic structures it is possible to conclude that archaeal ferritins from *P. furiosus* and *A. fulgidus* are able to incorporate negatively charged, modestly sized diffusants, even in their fully "closed" forms, likely through the threefold channels whose nature appears less restrictive with respect to that of vertebrate proteins.

It must also be pointed out that the dimensions of DTNB slightly exceed the diameter of the threefold channel inferred from crystal structure coordinates. A mechanism considering rotameric adjustment of involved aminoacid side chains is thus necessary in order to allow for the entry/exit of small organic molecules. Recent demonstration of multiple conformers in aminoacids lining the inner entrance to the ferritin cavity may explain the necessary plasticity of the threefold channels in ferritins (Behera & Theil, 2014).

4.2 Humanization of an archaeal ferritin

Human ferritin has been extensively investigated as a protein carrier for biotechnological applications and represents a promising alternative to viral carriers for targeted delivery of anticancer drugs and imaging agents. Moreover, ferritins are naturally targeted toward ubiquitously expressed TfR1 transferrin receptors. Such properties have been widely exploited for the efficient delivery of anti-tumor drugs to iron-avid, fast replicating, tumor cells overexpressing TfR1. However, the most relevant limitation to their application is the need for acidic experimental conditions to induce its disassembly, a fundamental step in order to achieve encapsulation of the cargo within the internal cavity. These extreme experimental conditions could affect both drug stability and compromise complete reassembly of the ferritin cage when pH is restored to neutrality, resulting in poor cargo load efficiency. To overcome this issue, the unique scaffold of *A. fulgidus* ferritin was genetically engineered to mimic the peptidic sequence of the surface exposed BC loop of the analogous human H-chain ferritin. This new chimeric protein – HumAfFt – was shown to maintain the unique, salt-dependent association/dissociation properties of AfFt while exhibiting the typical HuHF recognition by TfR1.

4.2.1 "Humanized" *A. fulgidus* ferritin design

HuHF and AfFt display 31% sequence identity. The structural alignment of HuHF and AfFt monomers (PDB code 3AJO and 1S3Q, respectively) reveals

a shorter N-terminus and shorter loops between three of the four helices, with the exception of the long loop connecting B and C helices (Figure 4.3).

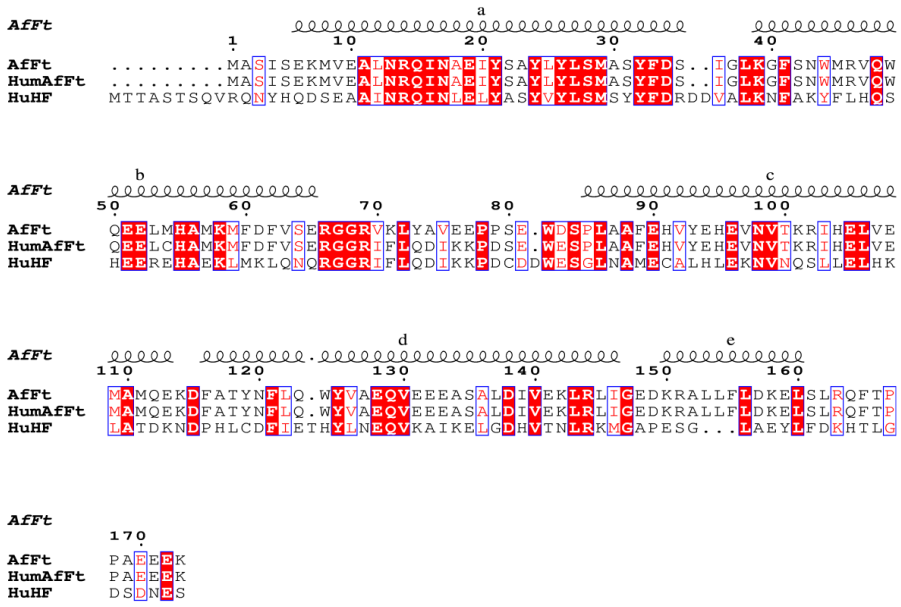


Figure 4.3 – Sequence alignment of AfFt, HumAfFt and HuHF. Elements of secondary structure for the AfFt are shown on the top. White characters in a red background indicate strict conservation while residues with poor conservation are drawn in black on a white background. Alignments were made using CLUSTAL Omega, and the figure was generated using ENDSCRIPT.(Robert & Gouet, 2014)

The BC loops of two adjacent subunits run in an antiparallel fashion establishing mutual interactions. Thus, BC loops of adjacent subunits display the same overall geometry in both HuHF and AfFt, though each couple of loops adopts a different symmetry along the spherical surface of the protein

cage, given the different dimer-dimer positioning within each complete 24-mer structure (Figure 4.4).

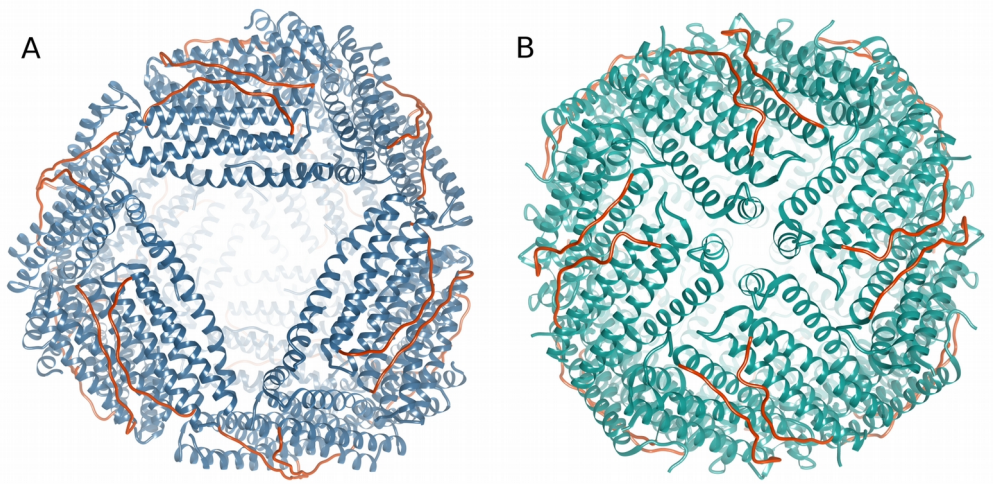


Figure 4.4 – Three dimensional structures determined by X-ray crystallography. A cartoon representation of **(A)** HumAfFt **(B)** HuHF (PDB code 3AJO). Models are colored blue and green, respectively and external loop connecting the helices B and C of each monomer is shown in red. Images were made using the UCSF Chimera software (Pettersen et al., 2004).

In order to preserve the unique assembly properties of AfFt, while implementing a potential cellular uptake, the the key residues involved in the stabilization of the symmetric protein architecture were determined using Pisa software (Krissinel & Henrick, 2007) and analyzing the crystal structure as a guide. Hydrophobic interactions are very largely spread and deeply contribute for the stabilization of 24-meric state. Analysis of the interactions stabilizing dimers in the 24-meric state identified two H bonds, involving residues Lys150 and Lys151, as critical for the stabilization of the "open"

pore conformation. In fact, Sana et al, obtained a “close” conformation by mutating these two residues (Sana et al., 2013). A summary of hydrogen bonds and salt bridges formed between chain A and B and the other subunits of the 24-mer are listed in Table 4.2.

Table 4.2
Interactions of chains A and B with the remaining chains of the 24-mer AfFT (PDB:1S3Q)

Position 1	Position 2	Distance (Å)	Type
Chain A			
Ile146 (O)	F: Lys150 (NZ)	2.9	H bond
Asp149 (OD1)	F: Lys151 (N)	3.8	H bond
Glu108 (OE1)	E: Lys114	2.7	Salt bridge
Glu108 (OE2)	E: Lys114	3.8	Salt bridge
Lys114	D: Glu108 (OE1)	3.3	Salt bridge
Lys114	D: Glu108 (OE2)	3.9	Salt bridge
Lys150 (NZ)	H: Mse111(O)	2.8	H bond
Chain B			
Arg151 (N)	D: Asp149	3.7	H bond
Mse111 (O)	G: Lys150	3.2	H bond
Tyr119	L: Arg151	2.7	Salt bridge

After these considerations, the external BC loop was determined as the most favorable location for redesigning the AfFt scaffold with the corresponding

residues in HuHF, taking in account the sequence alignment shown in figure 4.3. The external loop spans from residue 66 to 83 (AfFt numbering) and has a highly conservative region between archaea and human ferritin (RGGR) at the beginning, and a conservative tryptophan at the end of the loop. A total of nine aminoacids differ between HuHF and AfFt in the nineteen residue long loop, and these were mutated in order to match HuHF's sequence. In addition, a cysteine residue in position 54 (AfFt numbering) has been introduced by point mutation in order to provide a conjugation site for potential thiol reactive derivatives moieties into the cavity.

4.2.2 Structural characterization

The HumAfFt crystallized in a different condition with respect to wild type AfFt. Crystals were exposed to x-rays from a synchrotron in order to confirm the tetrahedral symmetry reported for archaeal ferritins. The protein was crystallized in the presence of Mg^{2+} in order to maintain the 24-mer assembly. The structure of HumAfFt has been determined by X-ray crystallography at a 2.87 Å resolution. It crystallized in the $C222_1$ space group, as observed also for wild type AfFt. The asymmetric unit (ASU) contains twelve identical subunits with a solvent content of 64.3 %. Data collection and refinement statistics are reported in Table 4.3.

The overall folding corresponded to the wild type AfFt structure (pdb code: 1S3Q) with a root mean square deviation (rmsd) value of 0.4 Å and displayed four wide triangular pores on the surface (Fig 4.4.A). B-factors analysis shows a mean B factor of 66 $Å^2$ with the exception of the loop region

between the D and E helices, which displays higher B-factors and a poor electron density on the side chains from 146 to 151 residues. Conversely, the BC loop is well organized and the analysis of the difference electron density (Fo-Fc) map clearly showed the presence of the mutated residues in the loop between helices B and C as well as the M54C mutation, pointing towards the inner cavity.

Table 4.3
Data collection and refinement statistics for the crystal structure of HumAfFt.
*Numbers in parentheses refer to data in the highest resolution shell.
*R_{free} was computed omitting 5% of the reflections as a test set.

Data collection	
Space Group	$C222_1$
Cell Dimensions a, b, c (Å)	185.80 190.65 176.54
Resolutions (Å)	48.86 - 2.94 (3.00 - 2.94)
R_{merge}	0.134 (0.751)
Unique reflections	59926 (4378)
$I/\sigma I$	7.8 (1.8)
Completeness (%)	95.8 (95.6)
Redundancy	4.2 (4.3)
Refinement	
Resolution (Å)	49.69 - 2.94 (3.00 - 2.94)
No. reflections	56823 (346)
R_{factor}	0.27 (0.61)
R_{free}	0.30 (0.59)
Ions (Mg^{2+})	2
Water molecules	13
r.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.99

Two magnesium ions have been positioned and successfully refined in the ASU, located in the ferroxidase site of two different subunits. In both, one magnesium ion is coordinated with both OE1 and OE2 of Glu19, (at 2.5 Å and 2.8 Å distance, respectively), with Glu52 (OE1) at 2.5 Å, with Gln129 (OE1) at 2.6 Å and with a water molecule at 2.6 Å distance. In the other chains, a water molecule has been modeled in the Fo-Fc map and successfully refined in the same position (Figure 4.5).

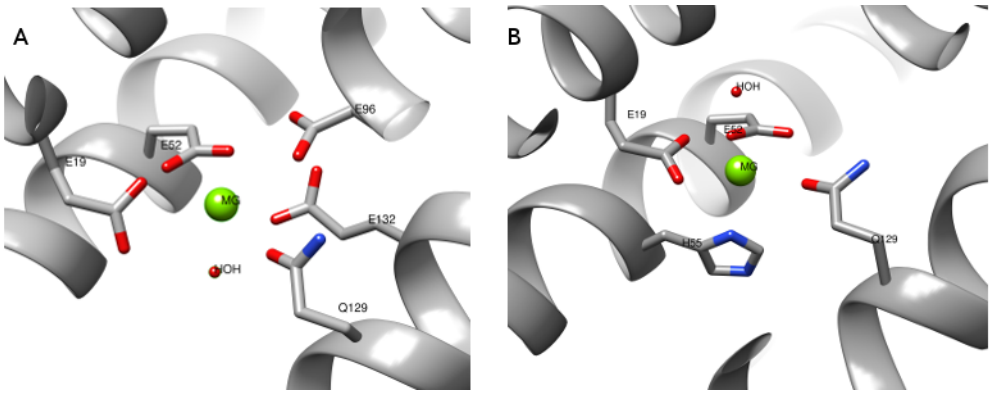


Figure 4.5 – Magnesium ions at modeled at HumAfFt ferroxidase center. **(A)** Mg sphere of coordination at chain I with aminoacids E19, E52, E96, E132, Q129 and a water molecule. **(B)** Mg sphere of coordination at chain F with E19, E52, Q129, H55 and a water molecule. Images were made using UCSF Chimera software. (Pettersen et al., 2004)

The loop region, including the conserved terminal turns, spans from aminoacid 68 to aminoacid 86. The structural superposition between HumAfFt and the HuHF, as well as between HumAfFt and AfFt, are shown

in Fig. 4.6. At the dimeric interface between the two antiparallel BC loops, the main interactions are a hydrogen bond between Arg69(NH2) and Ser80(O) at 2.8 Å distance and two salt bridges, namely Lys71(NZ) - Glu77(OE1) at 3.0 Å distance and Glu81(OE2) - Arg69(NH1) at 2.7 Å distance. A weak salt bridge is established by Glu81(OE2) and Arg69(NH2), at a distance of 3.0 Å. The salt bridge between Lys71 and Glu77 observed in HuHF is absent in HumAfFt since these positions were mutated into a phenylalanine (Phe71) and a lysine (Lys77). Other interactions are conserved between HuHF and HumAfFt and displayed the same distances.

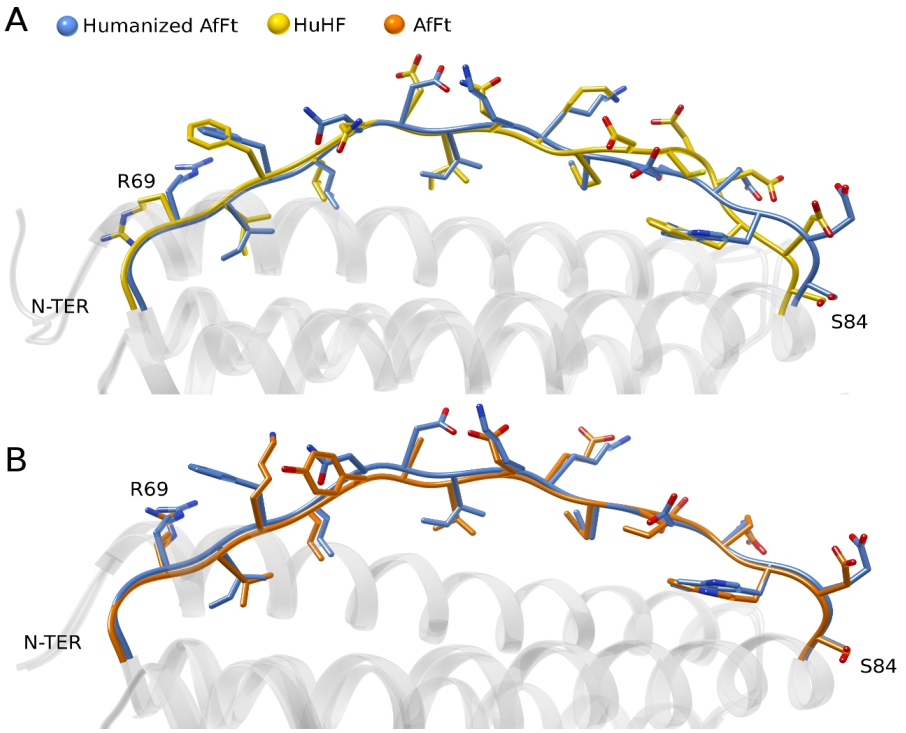


Figure 4.6 – Structural superposition of the region from R69 to S84 (AfFt numbering). **(A)** HumAfFt vs. HuHF **(B)** HumAfFt vs. wt AfFt. Residues are depicted as stick, with N atoms shown in blue, O atoms in red and S atoms in yellow. Images were made using UCSF Chimera software.(Pettersen et al., 2004)

Furthermore, HumAfFt structure in solution was probed by Cryo-EM, in order to asses the three-dimensional assembly in a near native environment. Single ferritin particles were visualized at a nominal resolution of 33 Å demonstrating that their shape and dimensions correspond to those of AfFt. Moreover, the triangular pores on the protein surface were observable in bi-dimensional views and clearly visible in three-dimensional reconstructions (Figure 4.7)

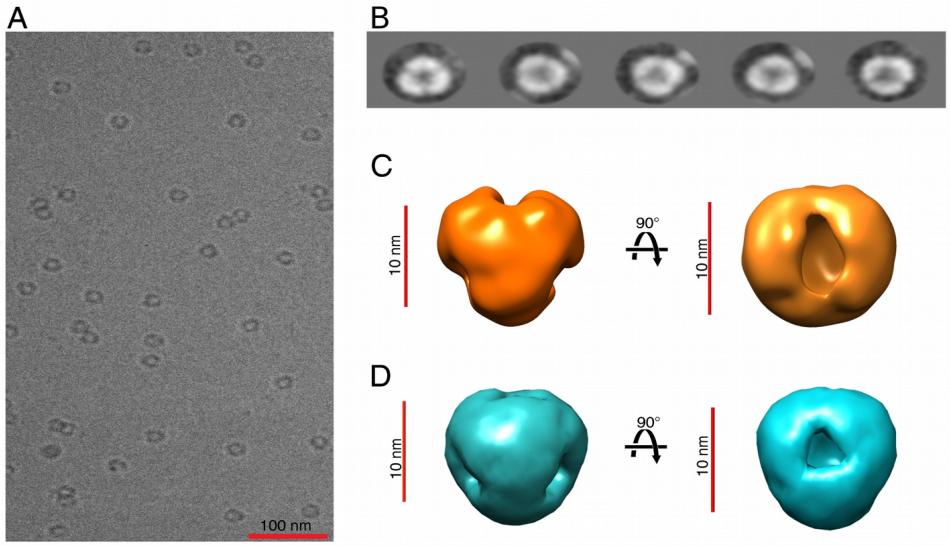


Figure 4.7 – Cryo-EM characterization of HumAfFt. **(A)** Sample micrograph of the HumAfFt. Scale bar: 100 nm. **(B)** Five representative 2D class averages obtained with RELION. **(C)** 3D reconstruction of HumAfFt obtained with RELION. Map final resolution: 33.1 Å. Scale bars: 10 nm. Left: side view. Right: top view. **(D)** *AfFt* crystal structure (PDB code 1S3Q) filtered to 30 Å, shown for comparison. Scale bars: 10 nm. Left: side view. Right: top view.

4.2.3 Cellular uptake of Humanized ferritin

After demonstranting that HumAfFt maintains its structure with large open pores we aimed to verify that the chimeric version was susceptible to internalization by eukaryotic cancer cells such as HeLa cells. It is known that HuHF is recognized and internalized by the TfR1, which is overexpressed in many types of tumor cells but not in normal cells and healthy tissues. To validate the effect of introduced mutations on the external loop in influencing

the uptake efficiency by HeLa cells, time course experiments on cells treated with the same amount (30 μg ml^{-1}) of AfFt, HumAfFt and transferrin labeled with FITC fluorescent moiety have been performed and analysis by flow cytometry carried. As a baseline for FITC fluorescence, control cells not incubated with FITC-ferritins were used. Moreover, to exclude any signal generated from outside particles sticking on the cell membrane due to unspecific binding or remains from the washing steps, trypan blue quenching was performed before using FACS. In figure 4.8, the FACS analysis is summarized, shown as the percentage of cells internalizing the ferritin nanoparticles at different times. These data highlighted that HumAfFt nanoparticles are efficiently taken up by HeLa cells already after one hour incubation with a much higher percentage compared to AfFt (81% and 5% respectively). After a longer incubation time (20 hours), the FITC-positive cells for the humanized samples increased to more than 90% whereas for native AfFt they were still less than 20%. The latter increment is possibly due to the unspecific uptake by pinocytosis.

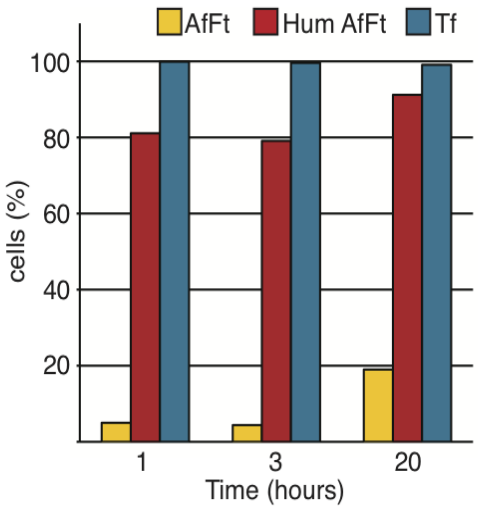


Figure 4.8 – Internalization of ferritins by Hela cells, measured using Flow Cytometry. HumAfFt is internalized with higher efficiency than the wild type. Cells have been treated with 30 μg ml^{-1} of AfFt-FITC, HumAfFt-FITC and Tf-FITC. The percentage of cells internalizing the nanoparticles at the time indicated is shown. For each sample 30 000 events gated on live cells have been acquired.

Moreover, RNAi experiment against the transferrin receptor were performed and analyzed the cellular uptake by FACS, obtaining around 50% HumAfFt uptake reduction after 48 hours of siRNA transfection, confirming the involvement of TfR1 in the internalization process (Figure 4.9)

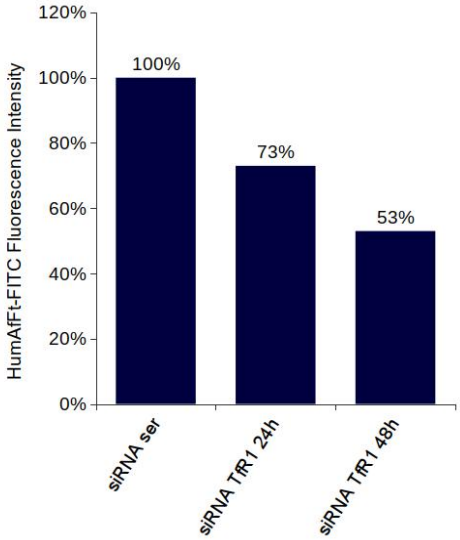


Figure 4.9 – HeLa cells transfected with control (scrambled) and specific anti-TfR1 siRNA were incubated with Hum AfFt-FITC for 24 and 48 hour,. After 3 hours cells were collected, washed and analyzed at the cytometer for the FITC fluorescent intensity. Compared to the FITC intensity in control cells (scrambled) an uptake reduction of almost 50% after 48 hours of RNAi was obtained.

In order to visualize uptaken HumAfFt–FITC nanocages, internalization assay directly on an ibidi 8-well μ-slide was executed, where cells were incubated for 20 hours before confocal microscopy. The properties of the HeLa TagRFP cell lines available in the lab were exploited for obtainment of a reference fluorescence signal confirming that imaging was occurring inside the cell. This cell line contains a TagRFP-FUS protein under the control of a doxycycline-inducible promoter that allows for a controlled expression of the protein and hence permits visualization of the nucleus in the red channel. The

cells were not selected to eliminate the untransfected cells, allowing to visualize different levels of expression and unstained nuclei. Just before imaging, the cells were washed to eliminate the unbound FITC-nanoparticles and then acquired by confocal laser-scanning microscopy. Confocal representative images of the entire field of view of live HeLa TagRFP cells alone (control) or incubated with HumAfFt-FITC or Tf-FITC are shown in figure 4.10 A. A detailed view of the boxed region in panel A is shown in figure 4.10 B. Images confirmed the high extent of HumAfFt internalization and highlighted a cellular distribution in the cytoplasm and in the perinuclear space comparable to that observed for transferrin, thus suggesting a typical clathrin-coated endocytosis pathway, mediated by TfR1.

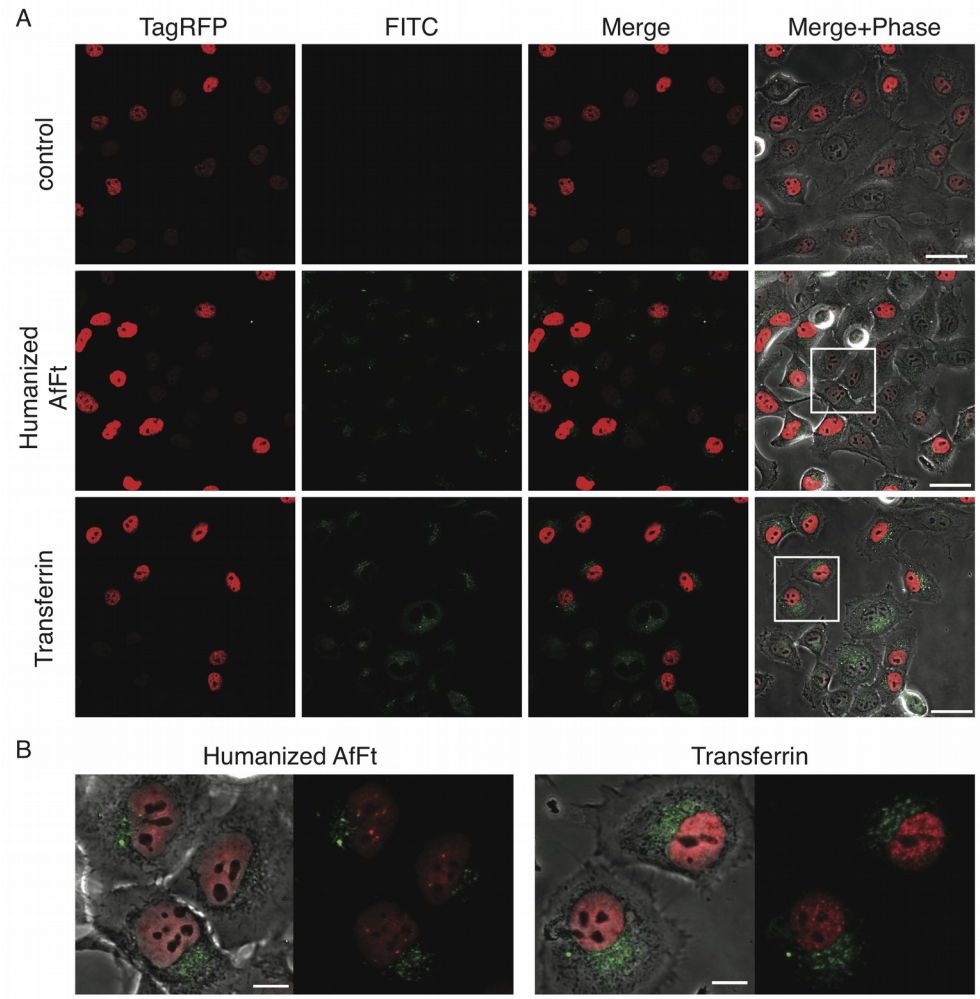


Figure 4.10 – Ferritin internalization observed at the confocal microscope. Cells were plated on an 8-well µ-slide (ibidi), induced with doxycycline to express TagRFP and then incubated with 30 µg ml^{-1} of HumAfFt-FITC or Tf-FITC for 20 h. After washing, cells were live-imaged using a con- focal microscope. **(A)** Confocal images of live HeLa TagRFP cells are shown as single, merged channels and overlay images with the phase contrast. Scale bars: 40 µm. **(B)** Images are a magnified view of the region highlighted by the white inset in panel A; merged channels and the overlay with the phase contrast images are shown. Scale bars: 10 µm.

4.2.4 Humanized ferritin self-assembly

HumAfFt was designed to retain most structural features of its scaffold, such as the tetrahedral 32 symmetry that results in the appearance of four large triangular apertures in the protein shell, as well as the larger cavity volume of ~325 Å with respect to ~261 Å for canonical ferritins. But more importantly it was designed to maintain the unique self-assembly/disassembly properties observed in AfFt. This protein it is stable as a dimer in neutral low ionic strength conditions and self-assembles at high ionic strength or in the presence of divalent cations (Sana et al., 2015). However, key thermodynamics and kinetics properties related to the assembly of HumAfFt are only partially understood.

In order to confirm the salt induced self-assembly, DLS measurements were carried in the absence and presence of 30 mM $MgCl_2$, for HumAfFt and AfFt, and the population distribution by diameter obtained. The hydrodynamic diameter (d_h) corresponds to the size of a hypothetical hard sphere that diffuses at equal speed as the particle under examination. Both proteins show similar population size-distribution before and after addition of $MgCl_2$ with a clear transition from small-size population of particles to another with significantly bigger particles. The calculated d_h of HumAfFt increases from 5.8 to 14.8 nm, in a similar fashion to AfFt whose d_h goes from 5.6 to 12.9 nm. These results are compatible with the transition from an unassembled state, most likely a dimer, to a fully assembled 24-meric ferritin. Additionally,

the absence of other major peaks, indicates that intermediates species formed during assembly are likely to be highly unstable.

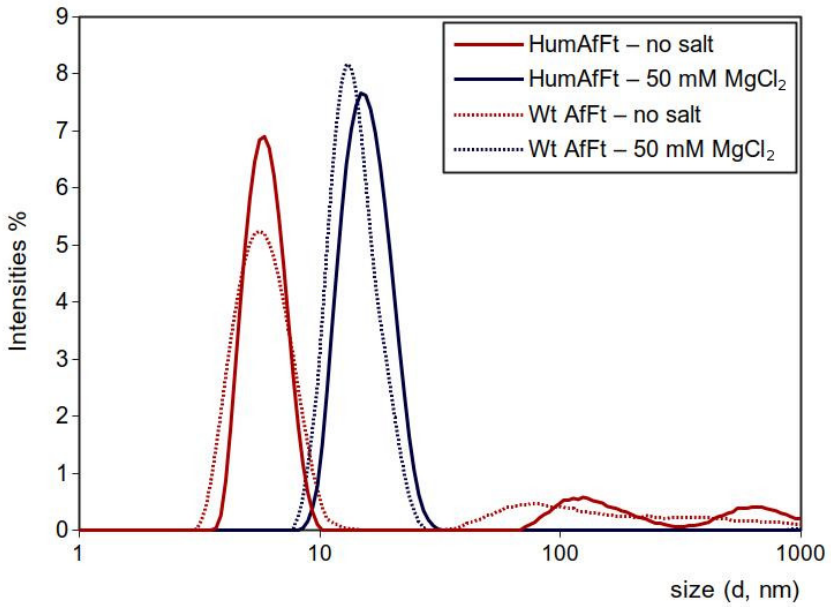


Figure 4.11 – Evaluation of population size-distribution by DLS. Diagram of intensities per size, comparing the HumAfFt and AfFt, before and after addition of 30mM $MgCl_2$. Protein concentration was kept fixed at 2mg/mL.

The the thermodynamics of the effect of $MgCl_2$ on HumAfFt's assembly was investigated by means of ITC. To a solution of protein, increasing concentrations of $MgCl_2$ were added and the heat transfer measured upon each injection. A representative titration profile is reported in figure 4.12, in which data was tentatively fitted with a 'one-site binding model' for calculation of thermodynamic signature of the phenomenon (Table 4.4). This model assumes a unique binding site per monomer, an assumption neither

proved or disproved with the current experiment, however the model find its utility in estimating thermodynamic parameters. Furthermore, a ratio of eight magnesium ions per each monomer was calculated as the needed amount to initiate the observed transition. The thermodynamic profile shows a positive and small ΔH contribution, mostly unfavorable, with a positive and favorable TΔS, of higher magnitude. This strongly suggests that hydrophobic forces dominate the assembly reaction. The huge entropic favorable contribution may report to the disruption of an organized hydration shell around the apolar patches on the solvent exposed surface with the subsequent release of water molecules. This has been been reported for other events involving the formation of supramolecular structures, like the case of self-assembling peptides and viral capsids (Hagan, 2014; Kabiri & Unsworth, 2014).

Table 4.4
Thermodynamic parameters calculated with a one-site binding model for curve fitting. *Values reported are the mean of three independent experiments.

ΔH (cal/mol)	**ΔS** (cal/mol/ deg)	**ΔG** (cal/mol)	$\mathbf{K_{app}}$ (mM)	**n**
323.3±46.2	20.3±1.2	-5739.1±405.9	0.09±0.02	7.8±0.8

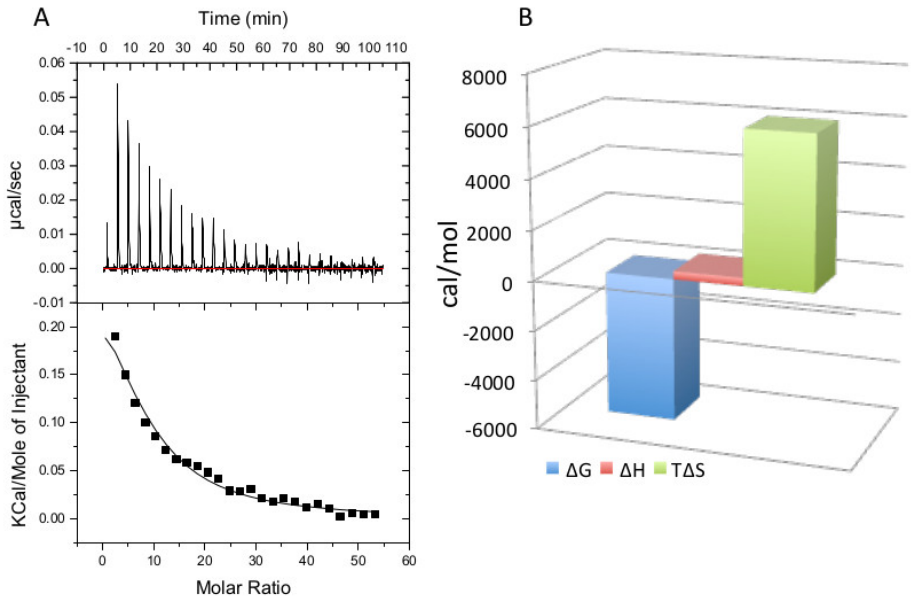


Figure 4.13 – Microcalorimetric titrations of HumAfFt with $MgCl_2$: **(A)** Raw ITC data of titration of 15 µM protein monomer solution with 4 mM $MgCl_2$ (upper panel) and, in the lower panel, integrated peak areas (black squares) fitted with the one-binding-site model. **(B)** Thermodynamic signature of the titration reported in (A) whose thermodynamic parameters are listed in Table 4.4.

In order to obtain more detailed structural information on the assembly between dimer and 24-mer in HumAfFt, the SAXS profiles obtained at increasing $MgCl_2$ concentrations were obtained and displayed in figure 4.14. SAXS is able to provide structural information of native proteins in solution and the structural changes in response to varying conditions. It can probe the macromolecular size, shape and oligomeric state of a complex without recurring to labeling (Koch et al., 2016; Svergun & Koch, 2003).

The shapes of the obtained SAXS curves reveal a significant shape change corresponding to a structural transition from a dimer, to a spheric like final state, compatible with a 24-mer, in concordance with the theoretical scattering curves calculated using CRYSOL (Figure 4.14) (Svergun et al., 1995)

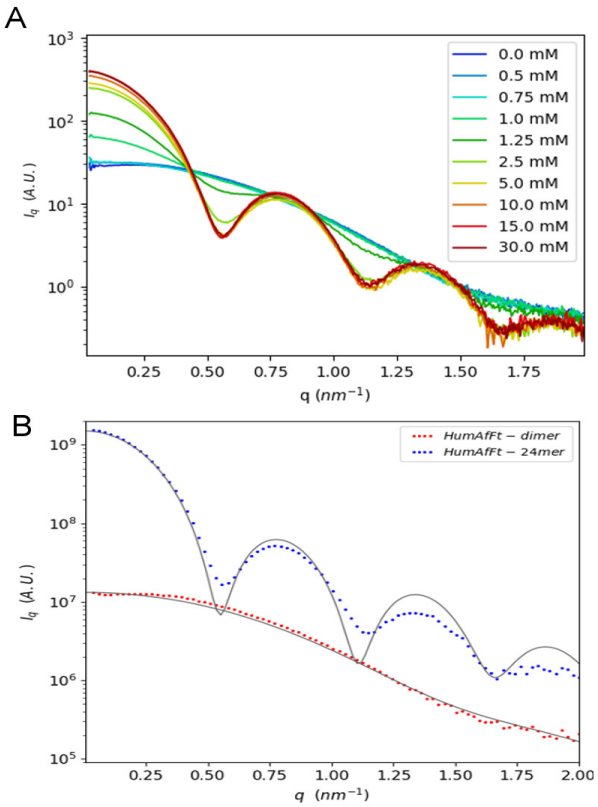


Figure 4.14 – HumAfFt SAXS profiles. **(A)** Changes in SAXS curves of HumAfFt at increasing $MgCl_2$ concentrations. Curves were obtained at 2.5mg/mL protein, and $MgCl_2$ concentrations are reported in the legend. **(B)** Comparison of theoretical scattering curves with the corresponding first and last experimental curves (0 and 30mM $MgCl_2$, respectively). Theoretical scattering curves were obtained with CRYSOL (Svergun et al., 1995)

From these experimental scattering profiles, the radius of gyration (R_g) and the forward scattering intensity I_0 were estimated with the Guinier approximation and plotted in Figure 4.15 . In a monodisperse solution, R_g is a parameter that reflects the size of that particle population, however in a complex system where polydispersity is inherent, the R_g is not simply an arithmetic average of the R_g of all present population but rather biased towards to the values of larger scattering particles. Conversely, I_0 is proportional to the mass concentration, the molecular weight, the beam intensity, and the square of the difference between the particle and the solvent (Putnam et al., 2007). In this system, both I_0 and R_g increased in function of $MgCl_2$ concentration with slightly different profiles.

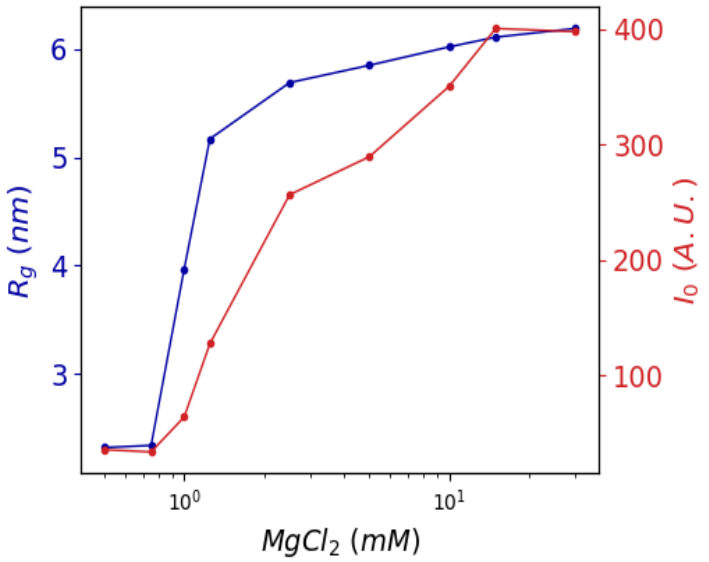


Figure 4.15 – HumAfFT $MgCl_2$ dependent assembly. Protein at 2.5mg/mL was titrated with increasing $MgCl_2$ concentrations and the SAXS profiles collected, at equilibrium. R_g and I_0 were determined using the Guinier approximation. R_g values plotted in blue and I_0 in red.

The SAXS data, in accordance with the DLS data, indicate that dimer and 24-mer are the dominant species at equilibrium and that the population of intermediates is negligible. Analysis of each single SAXS curve can in fact fully accounted for by the sum of the scattering profiles corresponding to the calculated scattering curved for the dimeric and 24-meric structures.

The sharp sigmoidal shapes of the curves hints to the high cooperativeness of the assembly process. Cooperativity can be the result of the formation of mutually supporting interactions, or a much more subtle consequence of conformational changes that occur upon binding. In the case of HumAfFt, the

cooperativeness observed is likely to be due to the the formation of new interfaces with every binding event between subunits. Nevertheless, the triggering effect of $MgCl_2$ on inducing small structural transient states more prone to bind and undergo full assembly, is an hypothesis that cannot be excluded here.

To investigate the $MgCl_2$ dependent kinetics of the HumAfFt assembly SAXS curves were collected at different time-delays upon reaction triggering (Figure 4.16). At each time-delay, the resulting scattering curve is a product of all present intermediates SAXS curves weighted by their molar concentrations. Although in theory, deconvolution of the SAXS curves could provide structural information on the intermediates and their respective concentrations, in reality the lack of information on possible/expected intermediates combined with the lack of adequate structural models for those intermediates makes this task virtually impossible. Nevertheless, by following the R_g in function of time, useful information on the system kinetics can be extracted that can provide clues into the complexity of the assembly mechanism. In figure 4.16, evolution of the R_g in function of time, for three selected $MgCl_2$ concentrations, reveals a sigmoidal profile typical of multistep reactions (Endres & Zlotnick, 2002). Moreover, the overall reaction rate is shows dependency on $MgCl_2$ concentration, as the kinetic traces are shifted to the left with increasing $MgCl_2$ concentrations. In fact, it a bigger portion of the initial part of kinetic trace is lost with increasing $MgCl_2$. Despite the lack of information at shorter time-delays, these profiles are compatible with a nucleation-growth mechanism, similarly to what has been

proposed by Zlotnick et al. for the assembly of HBV capsids (Zlotnick et al., 1999). In this model, the assembly is described as a cascade of low-order reactions and the kinetics predicted to be sigmoidal where intermediates hardly accumulate during the greater part of the reaction. The rate limiting step is the formation of a discrete oligomer more prone to assemble – the 'nucleus' - whose size can be inferred from the concentration dependency of the initial assembly rates (Zlotnick et al., 1999; Zlotnick et al., 2000). Based on these Hagan et al., determined that the critical nucleus size can be determined from the concentration dependence of the assembly half-life and that the elongation time can be revealed by the length of the lag phase (Hagan & Elrad, 2010).

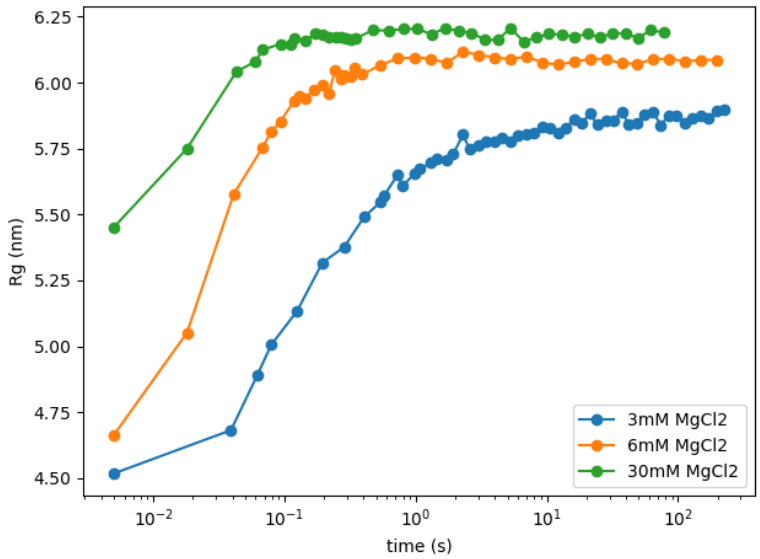


Figure 4.16 –Time-dependent changes during the HumAfFt assembly. R_g time-dependent increase upon rapid mixing of 3 mM, 6 mM and 30 mM $MgCl_2$. Experiements were carried at 2.5 mg/mL in 20 mM HEPES pH 7.4.

Fundamentally, most quaternary interactions are shared between AfFt and canonical ferritins up to the hexameric assembly i.e., one-quarter of the tetraeicosameric shell. As a consequence, the 3-fold symmetric channel formed at the center of each hexamer is conserved among all ferritins. The key difference thus resides on how the hexamers are joined to form the tetraeicosamer. In canonical ferritins, the hexamers are joined by 4-fold symmetry elements whereas in AfFt, the hexamers are faced in a 2-fold symmetry. Thus, specific amino acid residues at the hexamer–hexamer assembly are the most likely locations that govern whether the symmetry is tetrahedral or octahedral. In AfFt, the side chains of Lys150 and Arg153 appear to function as molecular doorstops, impairing the formation of the canonical octahedral symmetry. In contrast, a set of AfFt exclusive H-bonding and hydrophobic interactions contributes to the stabilization of the unique hexamer–hexamer interface thus giving rise to the unusual tetraeicosameric structure (Johnson et al., 2005; Sana et al., 2015). Nevertheless, the available structural information on the HumAfFt smaller assembly units, the dimer and hexamer, is strongly biased towards the rigid conformation of the 24-mer, as increased flexibility leads to instability and usually results in difficulty in crystallization and low resolution crystal diffraction (Yu et al., 2000)

For the assembly to occur spontaneously, the final formed cage must be lower in free energy than the initial state with free subunits. The assembly is expected to reduced translational and rotational entropy and thus must be

driven by formation of favorable subunits and the components that overcome this penalty. Thus, as with most protein-protein interactions, ferritin assembly is likely driven by a complex combination of hydrophobic, electrostatic, van der Waals and hydrogen bonding interactions (Hagan, 2014)

In many ways, the HumAfFt assembly resembles that of a viral capsid. The thermodynamic signature obtained with ITC showed that assembly is largely driven by hydrophobic forces. Similarly, Ceres and Zlotnick verified that the thermodynamic stability of HBV viral capsids increases, on one hand, with temperature, suggesting that hydrophobic interactions are the dominating driving forces; and on the other hand, with ionic strength implying that salts play a role in screening the repulsing electrostatic interaction opposing to subunit association (Ceres & Zlotnick, 2002). Models based on these assumptions have successfully been able to reproduce the dependence of protein-protein interactions on ionic strength (Kegel & van der Schoot, 2006; Šiber et al., 2012) and it has been proposed that higher salt concentrations could enhance assembly by favoring conformations active for assembly (Ceres & Zlotnick, 2002; Hagan, 2014). Assembly kinetics too show resemblance to those described for viral capsid assembly (Chen, et al., 2008; Endres & Zlotnick, 2002; Johnson et al., 2005; Zlotnick et al., 2000). Viral capsid and ferritin assembly share a set of common features such as sigmoidal kinetics of capsid growth starting with a lag-phase followed by a rapid increase that then reaches a plateau, intermediates are not accumulated during assembly.

The present data offer a clear frame for the understanding of the key features of HumAfFt $MgCl_2$ dependent assembly, that pave the way to proposing a relevant mechanism of assembly for this unusual ferritin.

4.3 Lanthanide binding ferritin

Among all the ferritin-based systems for bioimaging that have been created only a few find application in fluorescence based techniques. To meet these needs, a novel construct has been engineered in which a lanthanide binding tag (LBT) was genetically fused to the C-terminal end of mouse H-ferritin (MuHF). Fluorescent probes based on trivalent lanthanide ions are becoming widespread due to their photo-physical properties. One of their main advantages is their long lifetime of the excited state, that allows the specific signal from the lanthanide to be temporally separated from background signal. Lanthanide fluorescence life-times are in the milisecond time-range, while fluorescence arising from other sources such as reagents, plastics, cell environment are in the pico to microsecond time-range. This allows for the use of a technique often referred as time-resolved fluorescence, where the lanthanide fluorescence emission signal is acquired after the background signal has completely decayed. Other advantages of lanthanide include: their large Stokes's shift, i.e., excitation of lanthanides occurs in the UV region while emission happens at longer wavelengths, normally greater then 200 nm; and their narrow emission peak, that simplify the simultaneous use of multiple labels as emission peaks are less likely to overlap (Handl & Gillies,

2005). Despite this, lanthanides have inherent low extinction coefficients, and fluorescence emission is quenched in non-radiative processes (Bünzli, 2015). These issues have been addressed by the use of small organic fluorophores, in Forster resonance energy transfer (FRET) - like systems, where the a sensitizer fluorphores absorbs light in the UV region and transfer it to the lanthanide ion emitting in the visible region, in a phenomenon often referred as “antenna effect” (Selvin, 2002). Among lanthanides, Tb(III) is one of the most interesting due to their more intense microsecond fluorescence in the visible region (Bünzli & Piguet, 2005). A straightforward, potentially generalizable approach has been designed for the incorporation of lanthanide ions into biomolecules consisting of genetically engineering proteins to co-express a short peptidic sequence able to bind lanthanides – a lanthanide-tag. These have been developed on the basis of known properties of calcium binding loops and resulted in the development of short sequences with 20 or less amino acids with low-nanomolar affinities to target ions and selectivity towards lanthanides over other metals (Franz et al., 2003; Martin et al., 2007).

Native apoferritins have been report to naturally bind Tb(III) within their ferroxidase center at the three fold channel and at a third unidentified location likely corresponding to the nucleation site (Bou-abdallah et al., 2003). However, the distance between Terbium ions and aromatic moieties in ferritins made the FRET efficiency poor and suboptimal for fluorescence/luminescence based measurements. To improve both lanthanide binding to ferritins and their fluorescence properties, a short amino acid

sequence, initially designed by Martin et al. (Martin et al., 2007), has been engineered at the C-terminal end of MuHF, pointing to the cavity interior, with a tryptophan strategically located for FRET purposes and resulting chimeric protein called MuHF-LBT (Figure 4.17). The successful introduction was cofirmed by MALDI-TOF Mass spectrometry with an expected molecular weight of 22662 Da.

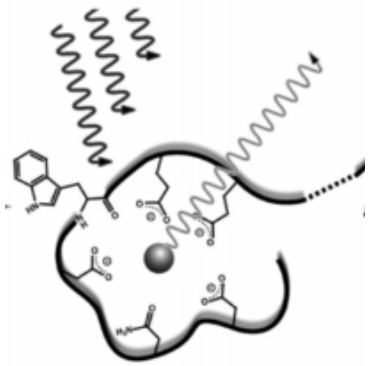


Figure 4.17 – Representation of Lanthanide binding site adapted from Martin et al (Martin et al., 2007)

4.3.1 Binding of Terbium to MuHF-LBT

In order to assess the binding of terbium to MuHF-LBT, the fluorescence emission spectra was recorded and compared to wild type MuHF, upon excitation at 295 nm (Figure 4.18). The intensity of the fluorescence at 544 nm was at least two orders of magnitude higher with respect to MuHF, showing that the introduction of the LBT was successful in improving Tb(III) fluorescence properties. The characteristic band at 544 nm in the MuHF-LBT loaded with Tb(III) was visible down to 4 pM of protein concentration and

neither dialysis nor size-exclusion chromatography were able to remove bound Tb(III), revealing high affinity binding.

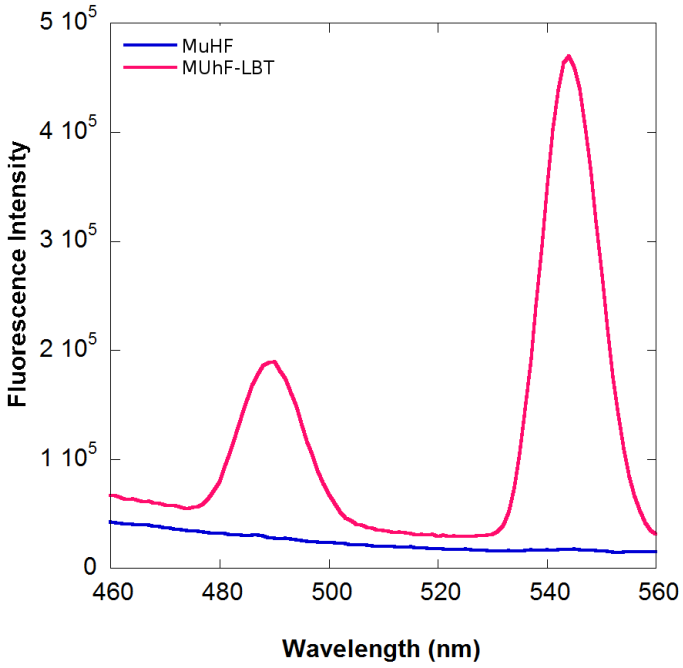


Figure 4.18 – Fluorescence spectra of MuHF-LBT Tb(III) complex (pink line) and MuHF Tb(III) complex (blue line) at the same protein concentration (1 µM). Spectra were recorded after 1 h incubation time with $TbCl_3$ in 0.1 M MES buffer pH 6.4.

Furthermore, titration of MuHF with increasing Tb(III) concentrations were performed, where a 1.7 equivalent amount of Tb(III) per monomer was calculated as endpoint. Carrying the same titration using wt MuHF demonstrated a negligible contribution to fluorescence of Tb(III) located at the ferroxidase center and three-fold channels. These results point out that the

introduced LBT is able to bind Tb(III) and efficiently increase the fluorescence of terbium, using the tryptophan as a FRET donor to terbium, resulting in two narrow emission bands at 490 and 544nm. Knowing that Tb(III) is able to bind at the ferroxidase center in average one Tb(III) per monomer, and that the full 24-mer is able to bind Tb(III) at the three-fold channels, and assuming one Tb(III) per LBT, one would expect maximum reached fluorescence to be maximal around 2.3 equivalent amount. However the lower 1.7 indicates that despite the high affinity binding, the LBT are likely not fully occupied. This suggests that the ferritin cage environment affect the structure of the introduced LBT.

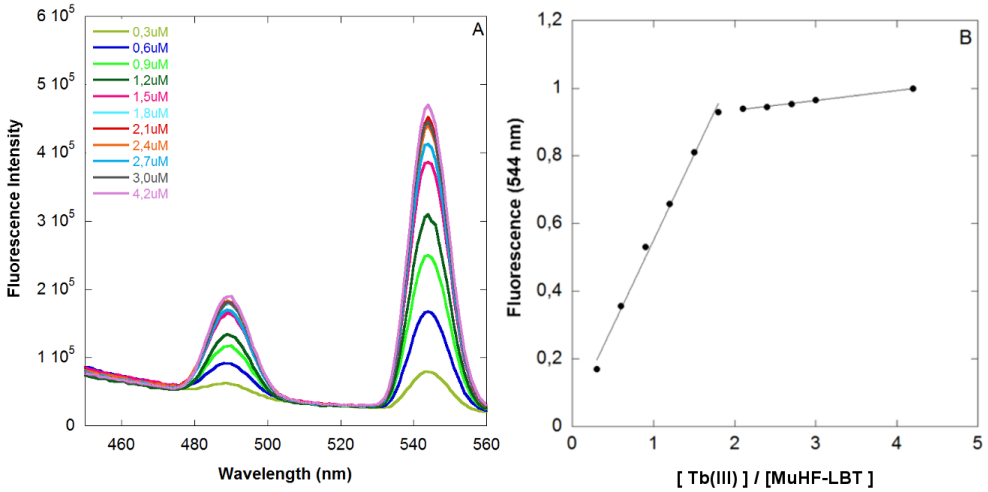


Figure 4.19 – Fluorescence titration of MuHF-LBT with Tb(III). **(A)** Fluorescence titration of MuHFt-LBT (1 μM) with incremental concentration of Tb(III) (0–4 equivalents) in 0.1 M MES buffer pH 6.4. Emission spectra were recorded in 1 cm pathlength cuvette upon excitation at 295 nm. **(B)** Fluorescence intensity of MuHFt-LBT Tb(III) complex as a function of the Tb(III)/MuHFt-LBT ratio. Fluorescent intensity was recorded at 545 nm and normalized to the emission maximum.

4.3.2 Structural Analysis of MuHF-LBT with and without Tb(III)

The structures of apoMuHF-LBT and MuHF-LBT-Tb(III) were determined by X-ray crystallography at a 2.85 Å and 2.65 Å resolution, respectively. They both crystallized in I222 space group with 24 identical subunits in the ASU with a solvent content of 64.7%. The overall X-ray-structure confirms that the LBT does not affect the protein scaffold which superposes to the native MuHF (pdb code 3WNW) with a rmsd value of 0.1 Å. However, the LBT loop was not visible in any of the two structures. This is thought to be due to high flexibility in the LBT. Temperature factors analysis shows a mean B factor of 57 $Å^2$ calculated for the main chain of the subunits in both structures and 78 $Å^2$ for the Tb(III) modeled in MuHF-LBT Tb(III) .

In order to confirm the presence or absence of Tb(III) in the crystals, X-ray emission scans from 4.0 to 21.0KeV were performed. Characteristic peaks corresponding to the L and M emission lines of Tb(III) were clearly identified in the MuHF-Tb(III) crystals and, as expected, absent in the MuHF-LBT crystals. This allow confident modeling of Tb(III) in the electron density maps.

In MuHF-LBT Tb(III) structure, 24 Terbium ions were positioned and successfully refined with a 75% occupancy in each ferroxidase site of the protein scaffold and eight Terbium ions were positioned with 100 % occupancy in each threefold axes. In the ferroxidase center; each Terbium ion is located in a trigonal planar coordination to OE1-Gln141 and to OE1 and

OE2-Glu62, and to OE2-Glu27, in a range 2.6-3.2 Å distance (Figure 4.20 A). In the threefold center, the Terbium ion is tetrahedrally coordinated to OE1- Glu134 of the three subunits, (at 2.2-2.4 Å distance) (Figure 4.20 B). In the 4-fold channels one water molecule each were positioned and successfully refined.

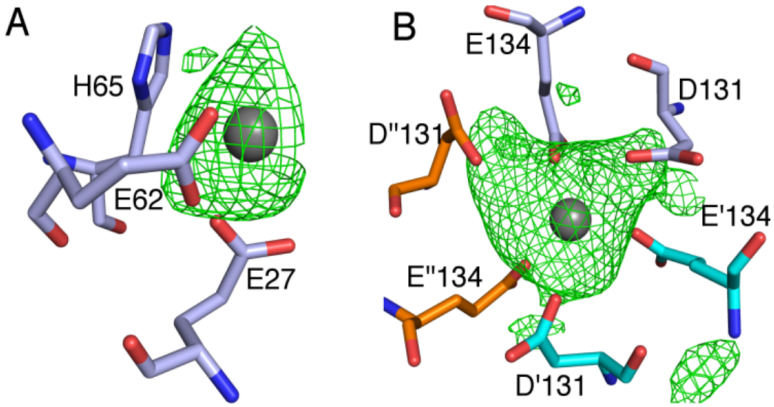


Figure 4.20 – Terbium binding sites from X-ray crystallography. The omit map contoured at 3 σ is shown as a green mesh **(A)** In the ferroxidase center: a Tb(III), is shown as a grey sphere and **(B)** in the threefold axes: the residues Glu131 and Glu134 of three different monomers are depicted as sticks in light blue, cyan and orange, respectively.

In an attempt to better describe the MuHF-LBT-Tb(III) a low resolution 7.1Å electron density map was obtained with Cryo-EM. The final 3D map of MuHF-LBT-Tb(III) corresponds to an overall structure that conforms to the high resolution crystallographic data indicating perfect matching of the subunit assembly and helix axes. It is worth considering, from the analysis of the selected 2D classes, that significant densities appear within the internal cavity that can be attributed to the LBT. Nevertheless, considerable

heterogeneity of the observed signals does not allow even a partial reconstruction of at least a single conformation of the 24 LBTs. Only, a kinked fragment of the terminal tail can be modeled, which corresponds to a segment of 6-7 aminoacids (Figure 4.21).

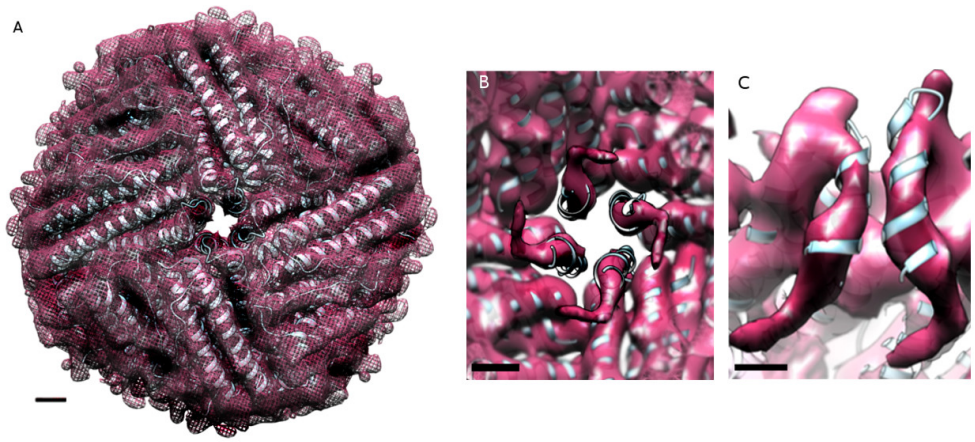


Figure 4.21 – 3D postprocess final maps of MuHF-LBT Tb(III) from Cryo EM analysis. Data were obtained with RELION (Scheres, 2012). Light blue: X-ray diffraction data of the crystal Terbium structure. Pink: experimental Cryo-EM electronic density map. Map resolution: 7.1 Å . Scale bar: 10 Å . **(A)** external view; **(B)** internal view; **(C)**: focus on one internal C-terminal helix.

Analysis of the three-dimensional structure of the MuFt-LBT Tb(III) complex, by X-ray crystallography and cryo-EM demonstrates that the presence of the C-terminal tag does not affect the overall assembly of the protein and that the genetically fused tags point to the interior cavity. Furthermore, the peptide arm connecting the lanthanide binding loop to the C-terminal sequence is flexible and does not allow for a complete resolution of the local structure.

4.3.3 Uptake of MuHF-LBT by tumor cell lines

After demonstration that MuHF-LBT-Tb(III) maintained the overall structure of MuHF, its potential use as a imaging agent was tested. It is known that MuHF is recognized and internalized through the TfR1, which is overexpressed in most cancer cells lines but not in healthy tissues. Experiments on cells treated with the same amount (0.5 mg/mL) of MuHF-LBT-Tb(III) or MuHF were undertaken in order to study the uptake efficiency of cancer cells lines by flow cytometry and confocal microscopy. However attempts to image live cells by direct excitation of the tryptophan residue at 290-375 nm, yielded poor results due to very high background noise within Tb(III) emission interval. Thus, uptake efficiency was assessed using FITC labeled ferritins. As a baseline for FITC fluorescence, control cells untreated with FITC-ferritins were used. In order to visualize MuHF-LBT-Tb(III) nanocages within cytoplasm, cells were observed under confocal microscopy. Confocal representative images of entire field of view of live cells incubated with MuHF-LBT Tb(III) are shown in Figure 4.22 A. These show that all tested cell lines (DU-145, HCT-116, MDA-MB-231 and SKOV-3) successfully internalized both MuHF-LBT and MuHF-LBT to a similar extent and highlight similar cellular distribution in the cytoplasm and in the perinuclear space. These were further confirmed by flow-citometry analysis, were similar percentage of cells with internalized ferritin were found (Figure 4.22 B).

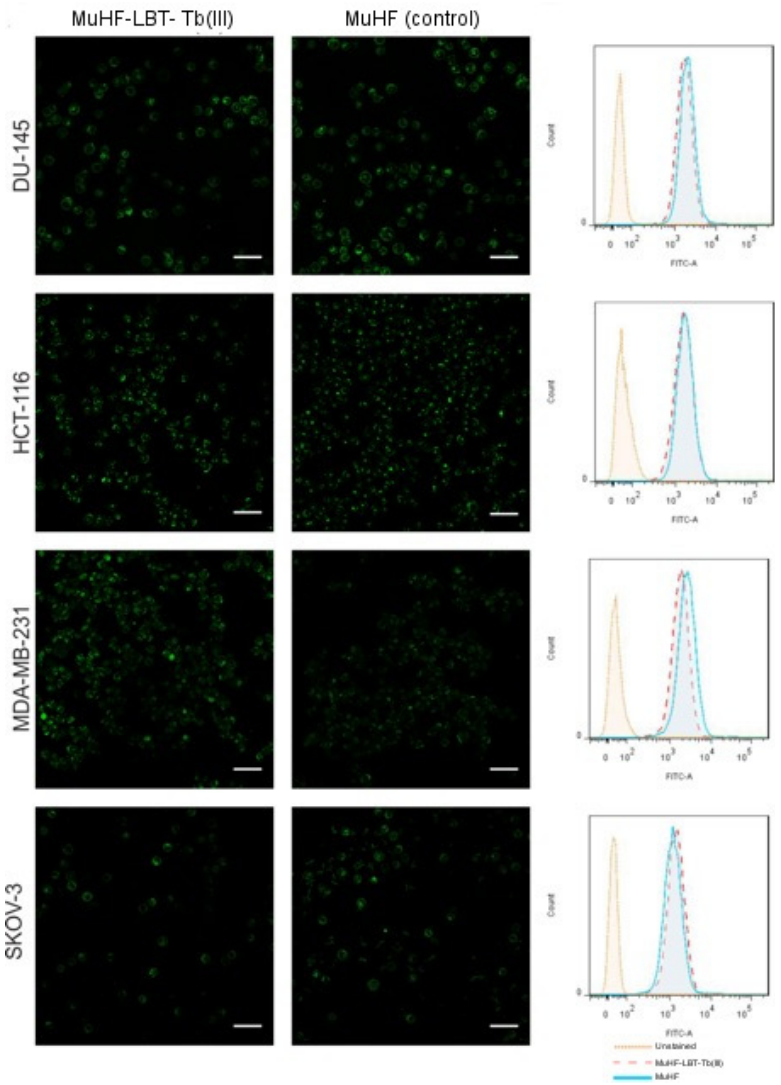


Figure 4.21 – Confocal microscopy images and flow cytometry analysis of MuHF-LBT Tb(III) uptake by selected tumor cell lines. DU-145, HCT-116, MuHF- LBT Tb(III) or MuHF (0.5 mg/ml) for 60 min. A) Images acquired by confocal microscopy showing side by side comparison of cellular distribution of MuHF-LBT-Tb(III) and MuHF conjugated with FITC. Scale bar: 50 μm. (B) Flow cytometry analysis of MuHF-LBT-Tb(III) and MuHF cellular uptake.

In conclusion, ferritin nanocage can be engineered to improve metal binding properties by addition of an appropriate metal binding tag facing the cavity's interior at topologically selected locations. This approach is an alternative to the quasi random metal cluster insertions into the ferritin cavity that has been used either by free diffusion of metal ions through the open pores on the protein shell or by disassembly and reassembly of the protein cage. The rational design of metal binding sites foster more advanced applications such as the formation of up-converting nanoparticles (Wang et al., 2011) or ultrabright fluorescent organic polymer for single molecule detection (Grazon et al., 2013) – emerging techniques for live cell imaging.

5 Conclusions

The body of experimental results of the present work highlights the notable properties related to small molecule entry into prototypic and genetically engineered archaeal ferritin homopolymers that are relevant both to the general mechanism of payload encapsulation into ferritins and to the widespread nanotechnological applications of these proteins. The current view on ferritin three-dimensional structure has always been consistent with the idea of an apparently closed, impermeable structure, in which the communication between the internal cavity and the outer space is uniquely governed by the gating effect of threefold channels, endowed with hydrophilic, negatively charged tunnels involved in iron uptake. This vision is consistent with the analysis of accessible surfaces on the ferritin shell based on crystallographic data that envisage the protein matrix as an almost uniform barrier that impairs even the diffusion of water molecules. In contrast, the present data indicate clearly that the protein matrix in archaeal ferritins does not provide a significant barrier against bulky ligands such as DTNB, which, due to its dimensions (8-10 Å) and the net negative charge, would never be able to cross the hydrophilic threefold channel, even though the channels in archaeal ferritins are less negative than in the mammalian ones. These findings indicate that molecular diffusion into ferritin is a complex phenomenon that most likely entails the formation of transient openings at the intersubunit interfaces due to possible synchronous motions of aminoacid side chains or, as recently reported (Zhang et al., 2017), to subunit rearrangements ascribable to “breathing motions” of the whole 24-

mer. In this framework, the small molecule entry process can be envisaged as a fully reversible one and fluxes of molecules in and out must simply obey concentration gradient between the internal cavity and the external environment. In our experimental approach, the small molecule release outside the cavity was deliberately abolished by formation of a covalent disulfide bond of the sulfidryl reactive molecule and a cysteine residue, inserted in topologically selected position inside the cavity. In the experimental set up, the kinetics of disulfide bond formation were in the order of tens of seconds, a time frame likely to reflect the slow crossing of the protein matrix, given the tens of millisecond time frame required for bimolecular disulfide bond formation. At present, the kinetics of small molecule release has not been investigated, but is likely to be much slower, possibly due to multiple interactions of the molecule within the surface of the ferritin internal cavity. In summary, the so-called “encapsulation” process in which ferritins are able to capture small molecules inside the cavity should be rather considered as a “matrix-assisted binding” of the small molecule to the surface of the internal cavity rather than a simple physical confinement.

Within this vision, the technological effort of protein engineering for payload delivery may be most conveniently addressed to modification of the properties of the surface of the internal cavity rather than to possible rearrangements of classical threefold channels. In this framework, the unique properties of AfFt are a prominent example of ferritin versatility and offer the opportunity of completely bypassing complications due to protein cavity accessibility and 24-mer disassembly. AfFt in fact represents a uniquely

suitable scaffold for incorporating a range of diverse substructures inside the protein cavity, either by assembly/disassembly process at neutral pH or by diffusion through the large triangular pores on the surface. Complete understanding of AfFt assembly process is however crucial in order to exploit the potentiality of this protein for biomedical applications. In this work, the effect of divalent cation on protein assembly has been clearly singled out, thus demonstrating that physiological Mg^{2+} concentrations are necessary and sufficient to promote full assembly of the 24-mer. At present, however, the effect of divalent cations on polymerization has not been explained as yet in terms of mechanism. The most likely hypothesis might entail cations binding to the iron binding sites on the dimeric species that promotes a conformational change leading to exposure of sticky ends capable of interdimers association into higher molecular weight species. In this framework, AfFt polymerization may possibly follow the same pathway identified for cations induced viral capsid assembly.

Nevertheless, one of the key properties of ferritin nanocages in biomedical applications is the possibility of targeting receptors on human cells, thus allowing the delivery of the desired payload within the cytoplasm. In this respect, HumAfFt, here described as humanized archaeal ferritin, combines the versatility of assembly and cargo incorporation of AfFt with binding to TfR1 and cellular uptake of HuHF. As a matter of fact, the term “humanized” is a misnomer, as HumAfFt contains just a stretch of aminoacids displayed in the external loop of HuHF, that is necessary and sufficient for TfR1 binding.

Further mutations will be necessary in order to make HumAfFt immunologically similar, if ever possible, to HuhF.

On a different strategy, grafting diverse functionalities on mammalian ferritins, while preserving immunological properties of the native protein, can be allowed by insertions of polypeptide segments within the internal cavity. In particular, the short C-terminal helix, located in proximity of the fourfold symmetry axis, represents a privileged topological region for adding novel functional properties through simple extension of the aminoacid sequence. Thus, a well characterized lanthanide binding sequence has been engineered at the C-terminal end of the MuHF in order to build a new carrier endowed with the notable property of chelating a terbium atom inside the cavity within a defined spatial position and stoichiometry and not quasi-randomly oriented as for iron oxide during the iron core formation. Besides the obvious use of lanthanide containing ferritin for diagnostic (e.g. Gadolinium for MRI imaging) or therapeutic (^{177}Lu as a radioemitting tool), the rationale for such engineering consisted in the first step for the set up of a new scaffold for further multifunctional polymetallic nanoparticles. The spatial confinement of 24 lanthanide atoms in proximity of the fourfold axis will in fact allow the possibility of placing other different elements that are capable of efficient energy transfer with the lanthanide itself. Such elements might be fluorescent organic antenna systems endowed with high photon absorption cross section (to compensate the limited light harvesting properties of lanthanide) encapsulated within the ferritin cavity. This strategy could be used to generate highly absorbing hot spots that transfer energy to the lanthanide ion

and exploit the intrinsic narrow band emission of the lanthanide itself. In addition, according to the same strategy, multi lanthanides could be hosted and generate up-converting photosystems through a two photon absorption process (by one species of lanthanide) whose energy is transferred to a single emitting ion at doubled energy. The synthesis of such innovative up-converting nanoparticles carrying with them the water solubility, stability and cell targeting properties of ferritins is currently ongoing and promises to become a smart tool not only in cell imaging applications but also in most advanced photonics applications (water soluble nanophosphors).

6 Abbreviations

ADC Antibody-drug conjugates

AfFt *Archaeoglobus. fulgidus* ferritin

ASU Asymmetric unit

Bfr Bacterioferritin

CCMV Cowpea chloroticmoottle virus

CD Cluster of differentiation

cryo-ET Electron cryotomography

cryo-TEM Cryo transmission electron microscopy

CTF Contrast transfer function

DDS Drug-delivery systems

DLS Dynamic light scattering

DMEM Dulbecco's Modified Eagle Medium

Dps DNA-binding protein from starved cells

DTNB 5,5'-dithio-bis-(2-nitrobenzoic acid)

EDTA Ethylenediaminetetraacetic acid

EPR Enhanced permeability and retention

ESRF European Synchrotron Radiation Facility

FACS Fluorescence activated cell sorting

FBS Fetal bovine serum

FITC Fluorescein-isothiocyanide

FOC Ferroxidase center

FRET Forster resonance energy transfer

FSC Fourier shell correlation

GSPG4 Chorndroitin sulfate proteoglycan 4

HA Heamagglutinin

HBV Hepatitis B virus

HoSP Horse spleen ferritin

Hsp Heatshock Protein

HuHF Human H-chain ferritin

HuLF Human L-chain ferritin

HumAfFt Humanized *A. fulgidus* ferritin

IPTG Isopropyl β-D-1-thiogalactopyranoside

ITC Isothermal titration calorimetry

LB Luria-Bertani

LBT Lanthanide binding peptide

LC-MS Liquid Chromatography – Mass Spectrometry

mAb Monoclonal antibodies

MRI Magnetic resonance imaging

MuHF Mouse H chain ferritin

PAA Polyacrylic acid

PBS Phosphate-buffered saline

PCR Polymerase chain reaction

PfFt *Pyrococcus furiosus* ferritin

rmsd root mean square deviation

ROS Reactive oxygen species

SAXS Small-angle X-ray scattering

Scara5 mouse scavenger receptor class A member 5

SEC Size-exclusion chromatography

TB Trypan Blue

TCEP Tris (2-carboxyethyl) phosphine

Tf Transferrin

TfR1 Human transferrin receptor 1

TIM-2 T-cell immunoglobuline and mucin domain containing receptor 2

TNB 5-thio-2-nitro-benzoic acid

US-FDA United States Food and Drug Administration

7 References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., … Zwart, P. H. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 213–221.

Aisen, P. (2004). Transferrin receptor 1. *The International Journal of Biochemistry & Cell Biology*, *36*(11), 2137–2143.

Alkhateeb, A. A., & Connor, J. R. (2010). Nuclear ferritin: A new role for ferritin in cell biology. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1800*(8), 793–797.

Alkhateeb, A. A., & Connor, J. R. (2013). The significance of ferritin in cancer: Anti-oxidation, inflammation and tumorigenesis. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1836*(2), 245–254.

Alkhateeb, A. A., Han, B., & Connor, J. R. (2013). Ferritin stimulates breast cancer cells through an iron-independent mechanism and is localized within tumor-associated macrophages. *Breast Cancer Research and Treatment*, *137*(3), 733–744.

Andrews, S. C. (2010). The ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1800*(8), 691–705.

Andrews, S. C., Smith, J. M. A., Yewdall, S. J., Guest, J. R., & Harrison, P. M. (1991). Bacterioferritins and ferritins are distantly related in evolution Conservation of ferroxidase-centre residues. *FEBS Letters*, *293*(1–2), 164–168.

Arosio, P., Elia, L., & Poli, M. (2017). Ferritin, cellular iron storage and regulation. *IUBMB Life*, *69*(6), 414–422.

Behera, R. K., & Theil, E. C. (2014). Moving Fe^{2+} from ferritin ion channels to catalytic OH centers depends on conserved protein cage carboxylates. *Proceedings of the National Academy of Sciences*, *111*(22), 7925–7930.

Bou-abdallah, F., Arosio, P., Sonia, L., Janus-Chandler, C., & Chasteen, N. D. (2003). Defining metal ion inhibitor interactions with recombinant human H- and L-chain ferritins and site-directed variants : an isothermal titration calorimetry study. *Journal of biological inorganic chemistry :JBIC : a publication of the Society of Biological Inorganic Chemistry, 8(4),* 489–497.

Bünzli, J.-C. G. (2015). On the design of highly luminescent lanthanide complexes. *Coordination Chemistry Reviews*, *293–294*, 19–47.

Bünzli, J.-C. G., & Piguet, C. (2005). Taking advantage of luminescent lanthanide ions. *Chemical Society Reviews*, *34*(12), 1048.

Burdo, J. R., & Connor, J. R. (2003). Brain iron uptake and homeostatic mechanisms: An overview. *BioMetals*, *16*(1), 63–75.

Calzolari, A., Oliviero, I., Deaglio, S., Mariani, G., Biffoni, M., Sposi, N. M., … Testa, U. (2007). Transferrin receptor 2 is frequently expressed in human cancer cell lines. *Blood Cells, Molecules, and Diseases*, *39*(1), 82–91.

Cann, A. J. (2012). *Principles of Molecular Virology*. *Principles of Molecular Virology*.

Carvalho, C., Santos, R. X., Cardoso, S., Correia, S., Oliveira, P. J., Santos, M. S., & Moreira, P. I. (2009). Doxorubicin : The Good , the Bad and the Ugly Effect. *Current Medicinal Chemistry*, *16*, 3267–3285.

Ceres, P., & Zlotnick, A. (2002). Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids. *Biochemistry*, *41*(39), 11525–11531.

Chen, C., Kao, C. C., & Dragnea, B. (2008). Self-Assembly of Brome Mosaic Virus Capsids: Insights from Shorter Time-Scale Experiments. *The Journal of Physical Chemistry A*, *112*(39), 9405–9412.

Chen, S., McMullan, G., Faruqi, A. R., Murshudov, G. N., Short, J. M., Scheres, S. H. W., & Henderson, R. (2013). High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy*, *135*, 24–35.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., … Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, *66*(1), 12–21.

Cho, K., Wang, X., Nie, S., Chen, Z., & Shin, D. M. (2008). Therapeutic Nanoparticles for Drug Delivery in Cancer. *Clinical Cancer Research*, *14*(5), 1310–1316.

Danev, R., Buijsse, B., Khoshouei, M., Plitzko, J. M., & Baumeister, W. (2014). Volta potential phase plate for in-focus phase contrast transmission electron microscopy. *Proceedings of the National Academy of Sciences*, *111*(44), 15635–1564

Daniels, T. R., Delgado, T., Helguera, G., & Penichet, M. L. (2006). The transferrin receptor part II: Targeted delivery of therapeutic agents into cancer cells. *Clinical Immunology*, *121*(2), 159–176.

Davis, A. V, Yeh, R. M., & Raymond, K. N. (2002). Supramolecular assembly dynamics. *Proceedings of the National Academy of Sciences*, *99*(8), 4793–4796.

Desideri, A., Stefanini, S., Polizio, F., Petruzzelli, R., & Chiancone, E. (1991). Iron entry route in horse spleen apoferritin. Involvement of the three-fold channels as probed by selective reaction of cysteine-126 with the spin label 4-maleimido-tempo. *FEBS Letter.* 287(1-2),10-14.

Dill, K. A., & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature Structural & Molecular Biology*, *4*(1), 10–19.

Douglas, T., & Ripoll, D. R. (1998). Calculated electrostatic gradients in recombinant human H-chain ferritin. *Protein Science*, *7*(5), 1083–1091.

Dufes, C., Al, M., & Somani, S. (2013). Transferrin and the transferrin receptor for the targeted delivery of therapeutic agents to the brain and cancer cells. *Therapeutic Delivery*, *4*(5), 629–640.

Easty, G. C., Yarnell, M. M., & Andrews, R. D. (1964). The Uptake of Proteins by Normal and Tumour Cells In Vitro. *British Journal of Cancer*, *18*(2), 354–367.

Emsley, P., & Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, *60*(12 I), 2126–2132.

Endres, D., & Zlotnick, A. (2002). Model-based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophysical Journal*, *83*(August), 1217–1230.

Falvo, E., Tremante, E., Arcovito, A., Papi, M., Elad, N., Boffi, A., … Ceci, P. (2016). Improved Doxorubicin Encapsulation and Pharmacokinetics of Ferritin–Fusion Protein Nanocarriers Bearing Proline, Serine, and Alanine Elements. *Biomacromolecules*, *17*(2), 514–522.

Falvo, E., Tremante, E., Fraioli, R., Leonetti, C., Zamparelli, C., Boffi, A., … Giacomini, P. (2013). Antibody–drug conjugates: targeting melanoma with cisplatin encapsulated in protein-cage nanoparticles based on human ferritin. *Nanoscale*, *5*(24), 12278.

Fan, K., Cao, C., Pan, Y., Lu, D., Yang, D., Feng, J., … Yan, X. (2012). Magnetoferritin nanoparticles for targeting and visualizing tumour tissues. *Nature Nanotechnology*, *7*(7), 459–464.

Fan, K., Gao, L., & Yan, X. (2013). Human ferritin for tumor detection and therapy. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, *5*(4), 287–298.

Fan, R., Chew, S. W., Cheong, V. V., & Orner, B. P. (2010). Fabrication of Gold Nanoparticles Inside Unmodified Horse Spleen Apoferritin. *Small*, *6*(14), 1483–1487.

Fargion, S., Arosio, P., Fracanzani, A. L., Cislaghi, V., Levi, S., Cozzi, A., … Fiorelli, G. (1988). Characteristics and expression of binding sites specific for ferritin H-chain on human cell lines. *Blood*, *71*(3), 753–757.

Flenniken, M. L., Willits, D. A., Harmsen, A. L., Liepold, L. O., Harmsen, A. G., Young, M. J., & Douglas, T. (2006). Melanoma and Lymphocyte Cell-Specific Targeting Incorporated into a Heat Shock Protein Cage Architecture. *Chemistry & Biology*, *13*(2), 161–170.

Franz, K. J., Nitz, M., & Imperiali, B. (2003). Lanthanide-binding tags as versatile protein coexpression probes. *ChemBioChem*, *4*(4), 265–271.

Frolow, F., Kalb, A. J., & Yariv, J. (1994). Structure of a unique twofold symmetric haem-binding site. *Nature Structural & Molecular Biology*, *1*(7), 453–460.

Frolow, F., Kalb (Gilboa), A. J., & Yariv, J. (1993). Location of haem in bacterioferritin of *E . coli*. *Acta Crystallographica Section D Biological Crystallography*, *49*(6), 597–600.

Geninatti Crich, S., Cadenazzi, M., Lanzardo, S., Conti, L., Ruiu, R., Alberti, D., … Aime, S. (2015). Targeting ferritin receptors for the selective delivery of imaging and therapeutic agents to breast cancer cells. *Nanoscale*, *7*(15), 6527–6533.

Grant, R. A., Filman, D. J., Finkel, S. E., Kolter, R., & Hogle, J. M. (1998). The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nature Structural Biology*, *5*(4), 294–303.

Grazon, C., Rieger, J., Me, R., Charleux, B., Clavier, G., Lyon, U., … Bat, T. L. (2013). Ultrabright Fluorescent Polymeric Nanoparticles Made from a New Family of BODIPY Monomers. *Macromolecules,*46(13), 5167-5176.

Hagan, M. F. (2014). Modeling Viral Capsid Assembly. In *Advanced Chemistry Physics,*155,1–68.

Hagan, M. F., & Elrad, O. M. (2010). Understanding the Concentration Dependence of Viral Capsid Assembly Kinetics—the Origin of the Lag Time and Identifying the Critical Nucleus Size. *Biophysical Journal*, *98*(6), 1065–1074.

Haikarainen, T., & Papageorgiou, A. C. (2010). Dps-like proteins: structural and functional insights into a versatile protein family. *Cellular and Molecular Life Sciences*, *67*(3), 341–351.

Haikarainen, T., Thanassoulas, A., Stavros, P., Nounesis, G., Haataja, S., & Papageorgiou, A. C. (2011). Structural and Thermodynamic Characterization of Metal Ion Binding in *Streptococcus suis* Dpr. *Journal of Molecular Biology*, *405*(2), 448–460.

Han, J., Seaman, W. E., Di, X., Wang, W., Willingham, M., Torti, F. M., & Torti, S. V. (2011). Iron Uptake Mediated by Binding of H-Ferritin to the TIM-2 Receptor in Mouse Cells. *PLoS ONE*, *6*(8), e23800.

Handl, H. L., & Gillies, R. J. (2005). Lanthanide-based luminescent assays for ligand-receptor interactions. *Life Sciences*, *77*(4), 361–371.

Hardy, J., & Selkoe, D. (2002). The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. *Science*, *297*(5580), 353–356.

Harrison, P. M., & Arosio, P. (1996). The ferritins: molecular properties, iron storage function and cellular regulation. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1275*(3), 161–203.

He, D., & Marles-Wright, J. (2015). Ferritin family proteins and their use in bionanotechnology. *New Biotechnology*, *32*(6), 651–657.

Huang, X., Chisholm, J., Zhuang, J., Xiao, Y., Duncan, G., Chen, X., … Hanes, J. (2017). Protein nanocages that penetrate airway mucus and tumor tissue. *Proceedings of the National Academy of Sciences*, *114*(32), E6595–E6602.

Huard, D. J. E., Kane, K. M., & Tezcan, F. A. (2013). Re-engineering protein interfaces yields copper-inducible ferritin cage assembly. *Nature Chemical Biology*, *9*(3), 169–176.

Ilari, A., Stefanini, S., Chiancone, E., & Tsernoglou, D. (2000). The dodecameric ferritin fro Listeria innocua contains a novel intersubunit iron-binding site. *Nature Structural Biology*, *7*(1), 38–43.

Jacobs, J. F., Hasan, M. N., Paik, K. H., Hagen, W. R., & van Loosdrecht, M. C. M. (2010). Development of a bionanotechnological phosphate removal system with thermostable ferritin. *Biotechnology and Bioengineering*, *105*(5), 918–923.

Johnson, E., Cascio, D., Sawaya, M. R., Gingery, M., & Schröder, I. (2005). Crystal structures of a tetrahedral open pore ferritin from the hyperthermophilic Archaeon *Archaeoglobus fulgidus*. *Structure*, *13*(4), 637–648.

Johnson, J. M., Tang, J., Nyame, Y., Willits, D., Young, M. J., & Zlotnick, A. (2005). Regulating self-assembly of spherical oligomers. *Nano Letters*, *5*(4), 765–770.

Kabiri, M., & Unsworth, L. D. (2014). Application of isothermal titration calorimetry for characterizing thermodynamic parameters of biomolecular interactions: Peptide self-assembly and protein adsorption case studies. *Biomacromolecules*, *15*(10), 3463–3473.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 133–144.

Kanekiyo, M., Wei, C.-J., Yassine, H. M., McTamney, P. M., Boyington, J. C., Whittle, J. R. R., … Nabel, G. J. (2013). Self-assembling influenza nanoparticle vaccines elicit broadly neutralizing H1N1 antibodies. *Nature*, *499*(7456), 102–106.

Kang, Y. J., Park, D. C., Shin, H.-H., Park, J., & Kang, S. (2012). Incorporation of Thrombin Cleavage Peptide into a Protein Cage for Constructing a Protease-Responsive Multifunctional Delivery Nanoplatform. *Biomacromolecules*, *13*(12), 4057–4064.

Kegel, W. K., & van der Schoot, P. (2006). Physical Regulation of the Self-Assembly of Tobacco Mosaic Virus Coat Protein. *Biophysical Journal*, *91*(4), 1501–1512.

Kim, M., Rho, Y., Jin, K. S., Ahn, B., Jung, S., Kim, H., & Ree, M. (2011). pH-Dependent Structures of Ferritin and Apoferritin in Solution: Disassembly and Reassembly. *Biomacromolecules*, *12*(5), 1629–1640.

Kish-Trier, E., & Hill, C. P. (2013). Structural Biology of the Proteasome. *Annual Review of Biophysics*, *42*(1), 29–49.

Kitagawa, T., Kosuge, H., Uchida, M., Dua, M. M., Iida, Y., Dalman, R. L., … McConnell, M. V. (2012). RGD-Conjugated Human Ferritin Nanoparticles for

Imaging Vascular Inflammation and Angiogenesis in Experimental Carotid and Aortic Disease. *Molecular Imaging and Biology*, *14*(3), 315–324.

Koch, M. H. J., Vachette, P., & Svergun, D. I. (2003). Small-angle scattering : a view on the properties, structures and structural changes of biological macromolecules in solution. *Quarterly Reviews of Biophysics*. 36(2), 147-227.

Kondo, K., Noguchi, M., Mukai, K., Matsuno, Y., Sato, Y., Shimosato, Y., & Monden, Y. (1990). Transferrin Receptor Expression in Adenocarcinoma of the Lung as a Histopathologic Indicator of Prognosis. *Chest*, *97*(6), 1367–1371.

Kramer, R. M., Li, C., Carter, D. C., Stone, M. O., & Naik, R. R. (2004). Engineered Protein Cages for Nanomaterial Synthesis. *Journal of the American Chemical Society*, *126*(41), 13282–13286.

Krissinel, E., & Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*, *372*(3), 774–797.

Kuruppu, A. I., Zhang, L., Collins, H., Turyanska, L., Thomas, N. R., & Bradshaw, T. D. (2015). An Apoferritin-based Drug Delivery System for the Tyrosine Kinase Inhibitor Gefitinib. *Advanced Healthcare Materials*, *4*(18), 2816–2821.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(2), 283–291.

Lawrence, C. M., Sanjoy, R., Babyonyshev, M., Galluser, R., Borhani, D. W., & Harrison, S. C. (1999). Crystal Structure of the Ectodomain of Human Transferrin Receptor. *Science*, *286*(5440), 779–782.

Lee, E. J., Lee, N. K., & Kim, I.-S. (2016). Bioengineered protein-based nanocage for drug delivery. *Advanced Drug Delivery Reviews*, *106*, 157–171.

Lee, E. J., Lee, S. J., Kang, Y. S., Ryu, J. H., Kwon, K. C., Jo, E., … Lee, J. (2015). Engineered proteinticles for targeted delivery of siRNA to cancer cells. *Advanced Functional Materials*, *25*(8), 1279–1286.

Levi, S., Cesareni, G., Arosio, P., Lorenzetti, R., Soria, M., Sollazzo, M., … Cortese, R. (1987). Characterization of human ferritin H chain synthetized in *Escherichia coli*. *Gene*, *51*(2–3), 269–274.

Levy, R. H., & Bauer, L. A. (1986). Basic Pharmacokinetics. *Therapeutic Drug Monitoring*, *8*(1), 47–58.

Levy, Y. (2017). Protein Assembly and Building Blocks: Beyond the Limits of the LEGO Brick Metaphor. *Biochemistry*, *56*(38), 5040–5048.

Li, J. Y., Paragas, N., Ned, R. M., Qiu, A., Viltard, M., Leete, T., … Barasch, J. (2009). Scara5 Is a Ferritin Receptor Mediating Non-Transferrin Iron Delivery. *Developmental Cell*, *16*(1), 35–46.

Li, L., Fang, C. J., Ryan, J. C., Niemi, E. C., Lebron, J. A., Bjorkman, P. J., … Seaman, W. E. (2010). Binding and uptake of H-ferritin are mediated by human transferrin receptor-1. *Proceedings of the National Academy of Sciences*, *107*(8), 3505–3510.

Maassen, S. J., van der Ham, A. M., & Cornelissen, J. J. L. M. (2016). Combining Protein Cages and Polymers: from Understanding Self-Assembly to Functional Materials. *ACS Macro Letters*, *5*(8), 987–994.

MaHam, A., Tang, Z., Wu, H., Wang, J., & Lin, Y. (2009). Protein-Based Nanomedicine Platforms for Drug Delivery. *Small*, *5*(15), 1706–1721.

Maity, B., Fukumori, K., Abe, S., & Ueno, T. (2016). Immobilization of two organometallic complexes into a single cage to construct protein-based microcompartments. *Chem. Commun.*, *52*(31), 5463–5466.

Marsh, J. A., & Teichmann, S. A. (2015). Structure, Dynamics, Assembly, and Evolution of Protein Complexes. *Annual Review of Biochemistry*, *84*(1), 551–575.

Martin, L. J., Hähnke, M. J., Nitz, M., Wöhnert, J., Silvaggi, N. R., Allen, K. N., … Imperiali, B. (2007). Double-lanthanide-binding tags: Design, photophysical properties, and NMR applications. *Journal of the American Chemical Society*, *129*(22), 7106–7113.

Mateu, M. G. (2013). Assembly, stability and dynamics of virus capsids. *Archives of Biochemistry and Biophysics*, *531*(1–2), 65–79.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*, *40*(4), 658–674.

Meyron-Holtz, E. G., Moshe-Belizowski, S., & Cohen, L. A. (2011). A possible role for secreted ferritin in tissue iron distribution. *Journal of Neural Transmission*, *118*(3), 337–347.

Millán, J. G., Brasch, M., Anaya-Plaza, E., de la Escosura, A., Velders, A. H., Reinhoudt, D. N., … Cornelissen, J. J. L. M. (2014). Self-assembly triggered by

self-assembly: Optically active, paramagnetic micelles encapsulated in protein cage nanoparticles. *Journal of Inorganic Biochemistry*, *136*, 140–146.

Milne, J. L. S. (2002). Molecular architecture and mechanism of an icosahedral pyruvate dehydrogenase complex: a multifunctional catalytic machine. *The EMBO Journal*, *21*(21), 5587–5598.

Molino, N. M., & Wang, S.-W. (2014). Caged protein nanoparticles for drug delivery. *Current Opinion in Biotechnology*, *28*, 75–82.

Moore, A. D., Björklund, Å. K., Ekman, D., Bornberg-Bauer, E., & Elofsson, A. (2008). Arrangements in the modular evolution of proteins. *Trends in Biochemical Sciences*, *33*(9), 444–451.

Munns, Yaxley, Coomer, Lavin, Gardiner, & Watters. (1998). Evaluation of the potential of transferrin-adriamycin conjugates in the treatment of bladder cancer. *BJU International*, *82*(2), 284–289.

Murshudov, G. N., Skubák, P., Lebedev, A. a., Pannu, N. S., Steiner, R. a., Nicholls, R. a., … Vagin, A. a. (2011a). *REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D Biological Crystallography*, *67*(4), 355–367.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., … Vagin, A. A. (2011b). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 355–367.

Neckers, L. M., & Trepel, J. B. (1986). Transferrin Receptor Expression and the Control of Cell Growth. *Cancer Investigation*, *4*(5), 461–470.

Noguchi, C. T., & Schechter, A. N. (1981). The intracellular Polimerization of Sickle Hemoglobin and Its Relevance to Sickle Cell Disease. *Blood*, *58*(6), 1057–1068.

Perlmutter, J. D., & Hagan, M. F. (2015). Mechanisms of Virus Assembly. *Annual Review of Physical Chemistry*, *66*(1), 217–239.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612.

Prevelige, P. E., Thomas, D., & King, J. (1993). Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells. *Biophysical Journal*, *64*(3), 824–835.

Prutki, M., Poljak-Blazi, M., Jakopovic, M., Tomas, D., Stipancic, I., & Zarkovic, N. (2006). Altered iron metabolism, transferrin receptor 1 and ferritin in patients with colon cancer. *Cancer Letters*, *238*(2), 188–196.

Putnam, C. D., Hammel, M., Hura, G. L., & Tainer, J. a. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics*, *40*(3), 191–285.

Riddles, P. W., Blakeley, R. L., & Zerner, B. (1979). Ellman's Reagent: 5,5'-Dithiobis(2-nitrobenzoic Acid) -a Reexamination. *Analytical Biochemistry.*94, 75–81.

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(1), 320–324.

Rosenthal, P. B., & Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology*, *333*(4), 721–745.

Russo, C. J., & Passmore, L. A. (2014). Ultrastable gold substrates for electron cryomicroscopy. *Science*, *346*(6215), 1377–1380.

Ryschich, E., Huszty, G., Knaebel, H. P., Hartel, M., Büchler, M. W., & Schmidt, J. (2004). Transferrin receptor is a marker of malignant phenotype in human pancreatic cancer and in neuroendocrine carcinoma of the pancreas. *European Journal of Cancer*, *40*(9), 1418–1422.

Ryser, H., Aub, J. C., & Caulfield, J. B. (1962). Studies on protein uptake by isolated tumor cells II. Quantitative Data on the Adsorption and Uptake of I131-Serum Albumin by Ehrlich Ascites Tumor Cells. *The Journal of Cell Biology*, *15*, 437–449.

Safari, J., & Zarnegar, Z. (2014). Advanced drug delivery systems: Nanotechnology of health design A review. *Journal of Saudi Chemical Society*, *18*(2), 85–99.

Saibil, H. R. (2008). Chaperone machines in action. *Current Opinion in Structural Biology*, *18*(1), 35–42.

Sana, B., Johnson, E., & Lim, S. (2015). The unique self-assembly/disassembly property of *Archaeoglobus fulgidus* ferritin and its implications on molecular release from the protein cage. *Biochimica et Biophysica Acta - General Subjects*, *1850*(12), 2544–2551.

Sana, B., Johnson, E., Magueres, P. Le, Criswell, A., Cascio, D., & Lim, S. (2013). The role of nonconserved residues of A*rchaeoglobus fulgidus* ferritin on its unique structure and biophysical properties. *Journal of Biological Chemistry*, *288*(45), 32663–32672.

Sasaki, E., Böhringer, D., van de Waterbeemd, M., Leibundgut, M., Zschoche, R., Heck, A. J. R., … Hilvert, D. (2017). Structure and assembly of scalable porous protein cages. *Nature Communications*, *8*, 14663.

Sato, D., Ohtomo, H., Yamada, Y., Hikima, T., Kurobe, A., Fujiwara, K., & Ikeguchi, M. (2015). Ferritin Assembly Revisited: A Time-resolved Small-angle X-ray Scattering Study. *Biochemistry*, *55*(2), 287-293.

Scheres, S. H. W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530.

Selvin, P. R. (2002). Principles and Biophysical Applications of Lanthanide-Based Probes. *Annual Review of Biophysics and Biomolecular Structure*, *31*(1), 275–302.

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

Šiber, A., Božič, A. L., & Podgornik, R. (2012). Energies and pressures in viruses: contribution of nonspecific electrostatic interactions. *Phys. Chem. Chem. Phys.*, *14*(11), 3746–3765.

Simsek, E., & Akif Kilic, M. (2005). Magic ferritin: A novel chemotherapeutic encapsulation bullet. *Journal of Magnetism and Magnetic Materials*, *293*(1), 509–513.

Stefanini, S., Vecchini, P., & Chiancone, E. (1987). On the mechanism of horse spleen apoferritin assembly: a sedimentation velocity and circular dichroism study. *Biochemistry*, *26*(7), 1831–1837.

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., … Carragher, B. (2005). Automated molecular microscopy: The new Leginon system. *Journal of Structural Biology*, *151*(1), 41–60.

Surguladze, N., Patton, S., Cozzi, A., Fried, M. G., & Connor, J. R. (2005). Characterization of nuclear ferritin and mechanism of translocation. *The Biochemical Journal*, *388*(3), 731–40.

Sutter, M., Boehringer, D., Gutmann, S., Günther, S., Prangishvili, D., Loessner, M. J., … Ban, N. (2008). Structural basis of enzyme encapsulation into a bacterial nanocompartment. *Nature Structural & Molecular Biology*, *15*(9), 939–947.

Svergun, D., Barberato, C., & Koch, M. H. (1995). CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography*, *28*(6), 768–773.

Svergun, D. I., & Koch, M. H. J. (2003). Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics*, 66(10), 1735–1782.

Theil, E. C., Behera, R. K., & Tosha, T. (2013). Ferritins for chemistry and for life. *Coordination Chemistry Reviews*, *257*(2), 579–586.

Theil, E. C., Matzapetakis, M., & Liu, X. (2006). Ferritins: iron/oxygen biominerals in protein nanocages. *Journal of Biological Inorganic Chemistry*, *11*(7), 803–810.

Thompson, K. J., Fried, M. G., Ye, Z., Boyer, P., & Connor, J. R. (2002). Regulation, mechanisms and proposed function of ferritin translocation to cell nuclei. *Journal of Cell Science*, *115*(10), 2165–77.

Todd, T. J., Zhen, Z., & Xie, J. (2013). Ferritin nanocages: great potential as clinically translatable drug delivery vehicles. *Nanomedicine*, *8*(10), 1555–1557.

Todorich, B., Zhang, X., Slagle-Webb, B., Seaman, W. E., & Connor, J. R. (2008). Tim-2 is the receptor for H-ferritin on oligodendrocytes. *Journal of Neurochemistry*, *107*(6), 1495–1505.

Tortorella, S., & Karagiannis, T. C. (2014). Transferrin Receptor-Mediated Endocytosis: A Useful Target for Cancer Therapy. *The Journal of Membrane Biology*, *247*(4), 291–307.

Uchida, M., Kang, S., Reichhardt, C., Harlen, K., & Douglas, T. (2010). The ferritin superfamily: Supramolecular templates for materials synthesis. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1800*(8), 834–845.

Uchida, M., Klem, M. T., Allen, M., Suci, P., Flenniken, M., Gillitzer, E., … Douglas, T. (2007). Biological Containers: Protein Cages as Multifunctional Nanoplatforms. *Advanced Materials*, *19*(8), 1025–1042.

Vannucci, L., Falvo, E., Fornara, M., Di Micco, P., Benada, O., Krizan, J., … Ceci, P. (2012). Selective targeting of melanoma by PEG-masked protein-based multifunctional nanoparticles. *International Journal of Nanomedicine*, *7*, 1489–1509.

Wang, M., Abbineni, G., Clevenger, A., Mao, C., & Xu, S. (2011). Upconversion nanoparticles: synthesis, surface modification and biological applications. *Nanomedicine: Nanotechnology, Biology and Medicine*, *7*(6), 710–729.

Wang, Q., Mercogliano, C. P., & Löwe, J. (2011). A Ferritin-Based Label for Cellular Electron Cryotomography. *Structure*, *19*(2), 147–154.

Wang, W., Knovich, M. A., Coffman, L. G., Torti, F. M., & Torti, S. V. (2010). Serum ferritin: Past, present and future. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1800*(8), 760–769.

Weeratunga, S. K., Gee, C. E., Lovell, S., Zeng, Y., Woodin, C. L., & Rivera, M. (2009). Binding of *Pseudomonas aeruginosa* Apobacterio ferritin-Associated Ferredoxin to Bacterioferritin B Promotes Heme Mediation of Electron Delivery and Mobilization of Core Mineral Iron. *Biochemistry*, *48*(31), 7420–7431.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., … Wilson, K. S. (2011). Overview of the *CCP* 4 suite and current developments. *Acta Crystallographica Section D Biological Crystallography*, *67*(4), 235–242.

Wu, H., Engelhard, M. H., Wang, J., Fisher, D. R., & Lin, Y. (2008). Synthesis of lutetium phosphate–apoferritin core–shell nanoparticles for potential applications in radioimmunoimaging and radioimmunotherapy of cancers. *Journal of Materials Chemistry*, *18*(15), 1779.

Yang, X., Arosio, P., & Chasteen, N. D. (2000). Molecular Diffusion into Ferritin : Pathways , Temperature Dependence , Incubation Time , and Concentration Effects. *Biophysical Journal*, *78*(4), 2049–2059.

Yang, X., & Chasteen, N. D. (1996). Molecular diffusion into horse spleen ferritin: a nitroxide radical spin probe study. *Biophysical Journal*, *71*(3), 1587–1595.

Yang, Z., Wang, X., Diao, H., Zhang, J., Li, H., Sun, H., & Guo, Z. (2007). Encapsulation of platinum anticancer drugs by apoferritin. *Chemical Communications*, *7345*(33), 3453.

Yu, L., Reutzel-Edens, S. M., & Mitchell, C. A. (2000). Crystallization and Polymorphism of Conformationally Flexible Molecules: Problems, Patterns, and Strategies. *Organic Process Research & Development*, *4*(5), 396–402.

Zhang, S., Zang, J., Chen, H., Li, M., Xu, C., & Zhao, G. (2017). The Size Flexibility of Ferritin Nanocage Opens a New Way to Prepare Nanomaterials. *Small*, *13*(37), 1701-45.

Zhang, Y., & Orner, B. P. (2011). Self-Assembly in the Ferritin Nano-Cage Protein Superfamily. *International Journal of Molecular Sciences*, *12*(12), 5406–5421.

Zhao, G., Ceci, P., Ilari, A., Giangiacomo, L., Laue, T. M., Chiancone, E., & Chasteen, N. D. (2002). Iron and Hydrogen Peroxide Detoxification Properties of DNA-binding Protein from Starved Cells. *Journal of Biological Chemistry*, *277*(31), 27689–27696.

Zheng, H., Cooper, D. R., Porebski, P. J., Shabalin, I. G., Handing, K. B., & Minor, W. (2017). CheckMyMetal: A macromolecular metal-binding validation tool. *Acta Crystallographica Section D: Structural Biology*, *73*, 223–233.

Zheng, S., Palovcak, E., Armache, J.-P., Cheng, Y., & Agard, D. (2016). Anisotropic Correction of Beam-induced Motion for Improved Single-particle Electron Cryo-microscopy. *bioRxiv*, 1–30.

Zhou, K., Li, L., Tan, Z., Zlotnick, A., & Jacobson, S. C. (2011). Characterization of Hepatitis B Virus Capsids by Resistive-Pulse Sensing. *Journal of the American Chemical Society*, *133*(6), 1618–1621.

Zlotnick, A., Johnson, J. M., Wingfield, P. W., Stahl, S. J., & Endres, D. (1999). A Theoretical Model Successfully Identifies Features of Hepatitis B Virus Capsid Assembly . *Biochemistry*, *38*(44), 14644–14652.

Zlotnick, A. (1994). To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. *Journal of Molecular Biology*. 241, 59-67

Zlotnick, A, Aldrich, R., Johnson, J. M., Ceres, P., & Young, M. J. (2000). Mechanism of capsid assembly for an icosahedral plant virus. *Virology*, *277*(2), 450–456.

Zwanzig, R., Szabo, A., & Bagchi, B. (1992). Levinthal's paradox. *Proceedings of the National Academy of Sciences*, *89*(1), 20–22.

Appendix I

Biochimica et Biophysica Acta 1861 (2017) 450–456

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Probing bulky ligand entry in engineered archaeal ferritins

Lorenzo Calisti [a], Irene Benni [a], Matilde Cardoso Trabuco [a], Paola Baiocco [b], Barbara Ruzicka [c], Alberto Boffi [a,d], Elisabetta Falvo [d], Francesco Malatesta [a], Alessandra Bonamore [a,*]

[a] Department of Biochemical Sciences "Alessandro Rossi Fanelli", Sapienza University of Rome, P.le Aldo Moro 5, I-00185 Rome, Italy
[b] Center for Life Nano Science@Sapienza, Istituto Italiano di Tecnologia, V.le Regina Elena 291, Rome I-00185, Italy
[c] Istituto dei Sistemi Complessi del Consiglio Nazionale delle Ricerche (ISC-CNR) Sede Sapienza and Dipartimento di Fisica, Sapienza University of Rome, P.le Aldo Moro 5, I-00185 Rome, Italy
[d] Istituto di Biologia e Patologia Molecolari, Consiglio Nazionale delle Ricerche (IBPM-CNR) Sede Sapienza University of Rome, P.le Aldo Moro 5, I-00185 Rome, Italy

ARTICLE INFO

Article history:
Received 31 May 2016
Received in revised form 1 September 2016
Accepted 11 October 2016
Available online 15 October 2016

Keywords:
Self-assembly
Binding kinetics
Nano-scaffold
Ferritin

ABSTRACT

Background: A set of engineered ferritin mutants from *Archaeoglobus fulgidus* (Af-Ft) and *Pyrococcus furiosus* (Pf-Ft) bearing cysteine thiols in selected topological positions inside or outside the ferritin shell have been obtained. The two apo-proteins were taken as model systems for ferritin internal cavity accessibility in that Af-Ft is characterized by the presence of a 45 Å wide aperture on the protein surface whereas Pf-Ft displays canonical (three-fold) channels.
Methods: Thiol reactivity has been probed in kinetic experiments in order to assess the protein matrix permeation properties towards the bulky thiol reactive DTNB (5,5′-dithiobis-2-nitrobenzoic acid) molecule.
Results: Reaction of DTNB with thiols was observed in all ferritin mutants, including those bearing free cysteine thiols inside the ferritin cavity. As expected, a ferritin mutant from Pf-Ft, in which the cysteine thiol is on the outer surface displays the fastest binding kinetics. In turn, also the Pf-Ft mutant in which the cysteine thiol is placed within the internal cavity, is still capable of full stoichiometric DTNB binding albeit with an almost 200-fold slower rate. The behaviour of Af-Ft bearing a cysteine thiol in a topologically equivalent position in the internal cavity was intermediate among the two Pf-Ft mutants.
Conclusions and general significance: The data thus obtained indicate clearly that the protein matrix in archaea ferritins does not provide a significant barrier against bulky, negatively charged ligands such as DTNB, a finding of relevance in view of the multiple biotechnological applications of these ferritins that envisage ligand encapsulation within the internal cavity.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Ferritins from a variety of species have emerged as versatile scaffolds for a number of diverse nanotechnological applications spanning from the synthesis of metal nanoparticles to drugs or diagnostics delivery. The most notable property of ferritins resides in their ability to sequester metals as well as small molecules within their internal cavities. Ferritin proteins are by far the best-studied biomineralisation scaffolds in that these proteins are able to accommodate up to 4500 iron atoms in an iron (III) oxide form within the central cavity. Iron (II) is oxidised within the ferroxidase centers located at the entrance of ferritin pores (threefold channels on the protein surface) and subsequently transferred to the central cavity and mineralised as iron (III) oxide nanocrystals. Thereafter, under physiological reducing conditions, the iron (III) oxide can be reverted to iron (II) and diffuses out of the cavity, most likely through negatively charged pores in the ferritin shell, formed between subunits [1]. These channels allow for the entry and exit of cations during mineralisation and demineralisation and display a relatively broad selectivity thus allowing for the accumulation of a variety of metal ions, with a preference for divalent cations. These properties have been used to develop ferritin as a drug delivery platform [2]. At physiological pH ferritin exists as a stable 24-mer, whereas in highly acidic or basic solutions it disassembles reversibly thereby spontaneously reassembling as neutral pH is restored. The reversible subunit assembly has been used to trap molecules in solution within its cavity simply by changing pH in the presence of the desired molecule. This property has been used to load the cavity with metal containing drugs, such as the cancer drug cisplatin [3,4], and the iron chelator desferrioxamine B [5,6], as well as a variety of organic and inorganic compounds, including metal nanoparticles.

The incorporation of non-metal-containing drugs within ferritin is however challenging due to the limited interactions with the ferritin shell, and the diffusion of these molecules through the surface pores. Strategies to overcome these problems have focused on complexing

* Corresponding author.
E-mail addresses: alessandra.bonamore@uniroma1.it, alessandra.bonamore@gmail.com (A. Bonamore).

http://dx.doi.org/10.1016/j.bbagen.2016.10.007
0304-4165/© 2016 Elsevier B.V. All rights reserved.

drugs with transition metals, such as Cu(II), prior to their internalisation [7], or the addition of charged accessory molecules such as poly-L-aspartic acid to optimise loading of ferritin with drugs [8]. By combining the loading of ferritin with drugs and surface modification with peptide epitopes and labels, ferritin can be specifically targeted to particular cell types and tumours for efficient delivery of therapeutic agents [9]. Thus far, ferritins and other protein nanocages show a great deal of promise that will hopefully transfer to a clinical setting.

Nevertheless, molecular diffusion in and out of the ferritin cavity appears to be a complex phenomenon that is only partially understood. The pathway of iron entry inside the ferritin cavity is characterized by the presence of eight protein ion channels positioned around the three-fold symmetry axes of the cage and delimited by three proximal subunits. Such ion channels are about 15 Å in length and 5–6 Å in diameter. Because hydrated ions have diameters of about 6.5 Å, partial dehydration is thought to occur for ion passage through the channel. On this basis, possible entry of larger ligands of physiological interest (e.g. iron III reductants) or other small organic molecules may seem unlikely. However, a large number of experimental observations, mostly carried out on mammalian ferritins heteropolimers or recombinant H and L homopolimers, pointed out that small organic molecules can permeate the protein shell, possibly through the same negatively charged three-fold channels that govern metal ions fluxes [10–12]. The permeation of small molecules into ferritins at physiological temperature and pH has been shown to be a charge-selective process in both native, H and L-chains of several mammalian proteins having similar channel structures [13]. A number of studies thus supported the hypothesis that the threefold channels are indeed the primary avenues of entry into the protein cavity of small molecules endowed with cationic properties [14]. The complete exclusion of the negatively charged probes from the interior of these proteins has been elegantly demonstrated by means of spin labeled molecules [11]. At the same time, the permeability of the same molecules in mutated proteins bearing positively charged groups along the three fold channels has been demonstrated [15,16]. More recently, however, many different ferritins from phylogenetically distinct trees have been identified and characterized. Among these, highly thermostable ferritins from Archaea emerged as privileged scaffolds in view of their remarkable thermal stability, easy expression in high yields in common *E. coli* cells and, at least in a few examples, uncommon association-dissociation properties. In particular, archaeal ferritins from *Archaeoglobus fulgidus* (Af-Ft) and *Pyrococcus furiosus* (Pf-Ft) emerged as most interesting tools for diverse applications. In spite of the high sequence similarity (50 and 70% amino acid sequence identity and similarity, respectively), Pf-Ft and Af-Ft do not share the same quaternary assembly. In fact, whereas the Pf-Ft 24-mer has the canonical 432 point-group symmetry, the Af-Ft 24-mer displays a 23 point-group symmetry typical of smaller 12-mer ferritin-like proteins. This unusual assembly does not display the 4-fold channels and constrains the quaternary structure thus leading to the appearance of four large triangular openings about 45 Å wide in the protein shell [17]. To date, Af-Ft assembly, here referred to as an "open" structure, is considered unique among all other known structures of tetraeicosameric ferritins. The stability of this tetrahedral configuration is governed by two critical residues in the helix E of the 4-helices bundle, namely K150 and R151. Indeed, the structure of Af-Ft K150A/R151A mutant reported by Sana et al. [18], shows a typical "closed" octahedral symmetry.

However, diffusion of small molecules within archaeal ferritins have been little investigated. Differences in the nature, shape and properties of open pores within these ferritins thus provide novel possible routes for small molecules entry/incorporation thus expanding the scope of possible biotechnological applications of these proteins. In the present work, we have engineered a set of Pf-Ft, Af-Ft ("open") and Af-Ft K150A/R151A ("closed") mutants by placing reactive cysteine residues in the same topological positions either inside or outside the internal cavity. The reactivity of the bulky, negatively charged DTNB molecule has been probed in the set of available mutants.

2. Materials and methods

2.1. Point mutations and protein expression

The genes encoding for bacterial ferritin from Af-Ft and Pf-Ft were cloned into the expression vector pET22b (Novagen). Point mutants Af-FtM54C, Af-FtM54C/K150A/R151A, Pf-FtG52C and Pf-FtP77C were obtained by PCR using QuickChange Mutagenesis kit (Stratagene). The recombinant plasmids were transformed into *E. coli* TOP 10 cells and the resulting colonies were screened by DNA sequencing. Plasmids bearing the desired mutations were transformed into BL21(DE3) *E. coli* strain for protein expression. For each mutant, protein over-expression was obtained as follows: 1 L LB broth medium was inoculated with 2 ml overnight culture of a single colony and the gene expression was induced with 1 mM IPTG when the absorbance at 600 nm reached 0.6. Cells were harvested by centrifugation after overnight induction at 37 °C and the cell pellets were stored at −20 °C.

2.2. Protein purification

Harvested cells form 1 l colture over-expressing Af-FtM54C and Af-FtM54C/K150A/R151A mutants were resuspended in 20 ml buffer A (25 mM HEPES pH 7.5, 20 mM $MgCl_2$) containing a cOmplete™ Mini Protease Inhibitor Cocktail Tablet (Roche) and disrupted by sonication. The soluble fraction was thermally purified by heating at 85 °C for 10 min followed by removal of denatured proteins by centrifugation at 14000 rpm for 30 min at 4 °C. The supernatant was fractioned by ammonium sulfate precipitation. 70% ammonium sulfate pellet containing highly purified protein was resuspended in buffer A, dialysed versus the same buffer, sterile filtered and stored at 4 °C. Cells over-expressing Pf-FtG52C and Pf-FtP77C were sonicated in 25 mM HEPES buffer at pH 7.5 containing 0.5 mM EDTA, 0.3 M NaCl and cOmplete™ Mini Protease Inhibitor Cocktail Tablet. After sonication, the crude bacterial extract was digested with DNase for 1 h at 37 °C, heated at 55 °C for 10 min and then at 80 °C for 8 min. Heat treatment was followed by centrifugation to remove insoluble material and ammonium sulfate precipitation. 70% ammonium sulfate pellet was resuspended in 20 mM HEPES pH 7.5 plus 150 mM NaCl, dialysed versus the same buffer and loaded onto a HiLoad 26/600 Superdex 200 pg column (GE Healthcare). Fractions containing highly purified protein were pooled, sterile filtered and store at 4 °C.

2.3. Preparation of ferritin-DTNB adducts

All mutants were reduced with 3 mM TCEP (tris(2-carboxyethyl)phosphine) in their storage buffers and then loaded onto a desalting column (GE Healthcare) to remove the reducing agent. Each mutant was reacted with 40-fold molar excess of Ellman's Reagent, (DTNB) per cysteine for 3 h at room temperature. Stock DTNB solutions were prepared in ethanol. The excess (non-reacted) reagent was removed by ultra-filtration using 100 kDa Amicon Ultra-15 centrifugal devices (Millipore Corporate). The Ferritin–DTNB samples were analyzed by mass spectrometry as described below.

2.4. Stopped flow experiments

Kinetic measurements were carried out on a thermostated Applied Photophysics stopped-flow apparatus (Leatherhead, UK) by mixing 8–10 μM protein solutions, previously reduced with TCEP, with solutions containing different concentrations of DTNB (from 0.2 to 0.7 mM after mixing) in 20 mM HEPES, 20 mM $MgCl_2$ pH 7.5. In order to avoid interference of the instrument phototube from the high concentrations of DTNB and the released chromophore 5-thio-2-nitro-benzoic acid (TNB), the reaction was followed at 430 nm and the extinction coefficient calculated to be 12205 mM^{-1} cm^{-1}, as determined from the extinction coefficient of 14150 mM^{-1} cm^{-1} at 412 nm [19]. All fitting

procedures were carried out by using the Matlab software (Mathworks, USA). Experimental traces were fitted by non-linear regression to either exponential of biexponential processes by using an Levenberg-Marquardt algorithm.

2.5. Self-assembly study

$MgCl_2$ -mediated self-assembly of ferritin mutants was studied by incubating aliquots of proteins (1 mg/ml) with different salt concentrations in 25 mM HEPES buffer, pH 7.5. Molecular sizes of Af-FtM54C, Af-FtM54C/K150A/R151A, Pf-FtG52C and Pf-FtP77C was determined by size exclusion chromatography (SEC) using HiPrep 16/60 Sephacryl S300 column (GE Healthcare). The column was equilibrated with 25 mM HEPES, pH 7.5, containing $MgCl_2$ at the same concentration in which the protein was pre-incubated and the same buffer was used as mobile phase. Molecular weight of each mutant were determined by comparing their elution volumes with the elution volumes of standard proteins in the same salt concentration.

Dynamic light scattering measurements (DLS) were performed using an ALV-5000 logarithmic correlator in combination with a standard optical set-up based on a He–Ne ($\lambda = 632.8$ nm) 10 mW laser and a photomultiplier detector. The intensity autocorrelation functions were directly obtained as $g_2(q, t) = \langle I(q, t)I(q, 0)\rangle / \langle I(q, 0)\rangle^2$, where q is the modulus of the scattering vector defined as $q = (4\pi n/\lambda)\sin(\theta/2)$ ($\theta = 90°$ in the present experiment). The raw measurements as directly obtained, without any data corrections are shown in Fig. 1 for different samples without added salt (black symbols) and at different $MgCl_2$ concentrations between 5 and 30 mM (colored symbols as described in the legend. Quantitative analysis of the measurements was obtained through a fit of the data with a single exponential expression: $g_2(q, t) = 1 + be^{-t/\tau}$ where b is the coherence factor and τ is the relaxation time related to the motion of the particles, specifically to the diffusion coefficient [20].

2.6. Protein LC–MS

LC–MS was performed on protein samples before and after DTNB titration, after dialysis in distilled water in the presence of 0.1 mM EDTA, using a Waters AcquityuPLC connected to Waters Acquity Single Quad Detector. A Hypersil Gold C4 column was used: 1.9 μm, 2.1 × 50 mm at 254 nm observation wavelength; mobile phase: 95:5 water (0.1% formic acid):MeCN (0.1% formic acid); gradient over 6 min (to 5:75 water (0.1% formic acid):MeCN (0.1% formic acid); flow rate: 0.4 ml min^{-1}; MS mode was set at a scan range: $m/z = 250–2{,}000$ (ES+); scan time: 0.25 s. Data were obtained in continuum mode by setting the electrospray source of the MS with a capillary voltage of 3.5 kV and a cone voltage of 50 V. N_2 gas was used as nebulizer and desolvation gas at a total flow of 300 l/h. Ion series were generated by integration of the ultraviolet-absorbance (at 254 nm) chromatogram over 1.2–1.8 min range. Mass spectra were subsequently reconstructed for proteins from the ion series using the MaxEnt 1 algorithm on MassLynx software program.

3. Results

3.1. Assessment of ferritin mutants assembly

In order to assess the association state of the engineered ferritin mutants Af-FtM54C, Af-FtM54C/K150A/R151A, Pf-FtG52C and Pf-FtP77C, all proteins were studied by DLS and SEC as a function of $MgCl_2$ concentration. A full characterization of Af-FtM54C association state is reported in Fig. 1. Data relative to the other mutants are reported as Supplementary materials (Figs. S1, S2 and S3). Fig. 1 clearly shows that the shape of the curve without added salt is significantly different from that observed in the presence of added salt. This behavior is confirmed by data fitting obtained as described under Methods section. The fitting

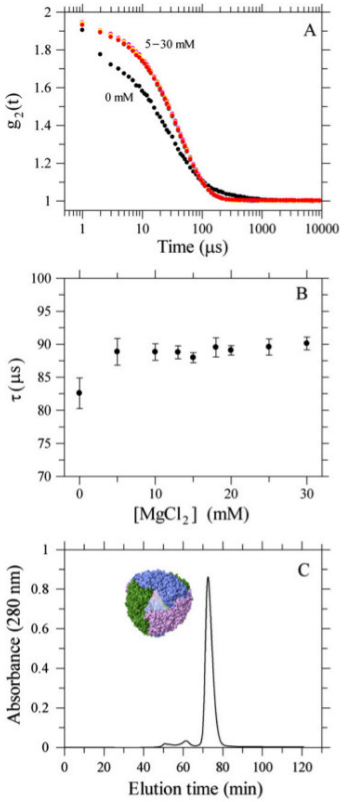


Fig. 1. Polymeric assembly of *Archaeoglobus fulgidus* ferritin mutants. A) Dynamic Light Scattering intensity correlation functions of Af-FtM54C at different $MgCl_2$ concentrations indicated in the legend. The curves at different $MgCl_2$ are all superimposed. B) Dynamic Light scattering relaxation time of Af-FtM54C as a function of $MgCl_2$ concentration. C) Gel filtration profiles of Af-FtM54C at 20 mM $MgCl_2$.

curves appears to interpolate very well experimental data in presence of salts whereas are less accurate for samples without added salt, where the addition of polydispersity is necessary.

Pf-Ft mutants are stable 24-mers (~490 kDa) independent of salt concentration, whereas the self-assembly of Af-Ft mutants was strongly dependent on ionic strength. $MgCl_2$ was used as polymerizing salt at variance with previous reports [18] in which NaCl was used. As a result, it appeared that at our working concentration (20 mM $MgCl_2$) all the proteins are structured as a stable 24-meric cage. It is noticeable that the 24-mer to dimer ratio increases with increasing $MgCl_2$ concentration from 0 to 20 mM, whereas at least 100 mM NaCl concentration of NaCl are needed to reach full polymerization (data not shown). Thus, it appears that divalent cations such as Mg^{2+} and Ca^{2+} are more effective in promoting the 24-mer association with respect to NaCl. The

comparison between SEC and DLS data on Af-FtM54C and Af-FtM54C/K150A/R151A however need further comments. First of all, the heterogeneous, polydisperse population observed in the absence of divalent cations is manifest in the presence of both low molecular weight species (presumably dimers as reported in refs. [16–18]) and distributed high molecular weight polymers (present in much lower amount), apparently in a proportion even higher than the canonical 24-mers. Addition of $MgCl_2$ at 5–10 mM concentration brings about a very sharp transition in both Af-Ft mutants resulting in the formation of homogeneous and stable 24-mer species. At 20 mM $MgCl_2$ concentration, all proteins investigated are consistent with a stable 24-mer assembly.

3.2. Kinetics of DTNB binding to Pf-Ft and Af-Ft mutants

The kinetics of the disulfide exchange reaction of DTNB with cysteine residues on ferritin mutants were carried out by stopped-flow spectroscopy. The kinetics, determined under pseudo-first order conditions, appeared to be multiexponential and were followed to the maximum time possible with the instrument (1000 s). Independent experiments performed by using UV-visible absorption spectroscopy and LC–MS (see below) confirmed the essentially complete cysteine reactivity. As Fig. 2A shows, the time scales of the four ferritin mutants are significantly different. As expected, the Pf-Ft P77C mutant bearing the cysteine residue on the convex outer surface displayed the fastest reactivity as compared to Pf-Ft G52C which carries the mutation inside the ferritin tetraeicosamer. A complete set of time courses of the Pf-Ft P77C mutant has been carried out increasing DTNB concentration and reported in Supplementary Materials (Fig. S4). Within the observed time frame all traces could be fit to a simple relaxation process with apparent second-order rate constants of *ca.* 900 M^{-1} cm^{-1} (see also Table 1). The Af-Ft mutants had an intermediate behaviour in the reaction with DTNB. Quite unexpectedly the Af-Ft M54C mutant, in which the internal cysteine reacts with DTNB in a facilitated way through the four 45-Å large triangular openings in the protein shell, could only be fit by a biexponential process in which the fastest second-order rate constant was of the same order of magnitude determined for the Pf-Ft P77C mutant (Table 1). Finally, closure of the openings by introduction of the K150A/R151A mutations in Af-M54C brought about a significant quenching of the ligand binding rate. The complete set of kinetic traces for each mutant were also fit by either exponential or biexponential relaxations with second-order rate constants as reported in Supplementary materials (Fig. S5). Overall these results indicate that the diffusion of molecules across the proteinaceous barrier of closed ferritin 24-mers can be accurately measured and secondly that contrary to intuition and reported data (see discussion section) the protein shell

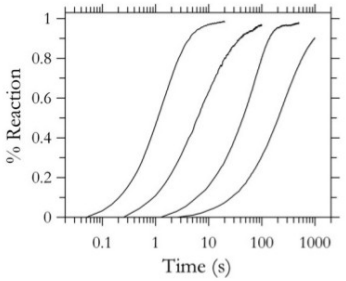


Fig. 2. Kinetics of DTNB binding to Af-Ft and Pf-Ft mutants. A. Comparison of the reaction time scales of all mutants; from left to right, Pf-FtP77C, Af-FtM54C, Af-FtM54C/K150A/R151A, and Pf-FtG52C. All mutants were 5 μM except Af-FtM54C which was 4 μM, and DTNB was 0.7 mM.

Table 1
Apparent second-order constants for ferritin mutants as determined by the DTNB reaction.

Protein	Cysteine position	k_{OBS} (M^{-1} s^{-1})
Pf-FtP77C	External	908 ± 122
Pf-FtG52C	Internal	5 ± 1
Af-FtM54C open	Internal	Fast: 198 ± 65 Slow: 51 ± 21
Af-FtM54C/K150A/R151A closed	Internal	26 ± 2

permeability to bulky molecules such as DTNB is much higher than expected. These results bear relevant implications discussed below in so far as ligand encapsulation and delivery is concerned.

3.3. LC/MS data

LC–MS measurements were performed on the selected point mutants Af-FtM54C, Af-FtM54C/K150A/R151A, Pf-FtG52C and Pf-FtP77C after titration with DTNB in comparison with unreacted proteins. In all samples, ferritins were eluted as monomers of molecular weight of about 21 kDa, and DTNB reacted proteins showed a shifted peak of + 198 ± 2 Da, in agreement with the expected molecular weight of the thio-nitrobenzoic moiety. In all samples reaction was complete with the exception of Pf-Ft G52C protein in which 18 % of unreacted protein was present (Supplementary materials, Fig. S6).

4. Discussion

The present data highlight notable properties related to bulky ligands penetration through prototypic archaeal ferritin homopolymers that are relevant both to the general ligand entry/escape mechanism and to the widespread nanotechnological applications of these proteins [21,22].

The experimental set up was designed around four mutants of Pf-Ft and Af-Ft bearing a cysteine residue per subunit either in topologically equivalent positions inside the 24-mer cavity (Af-FtM54C, Af-FtM54C/K150A/R151A, Pf-FtG52C) or outside the 24-mer cage in the Pf-FtP77C mutant (see Fig. 3). In this framework, access to the reactive sulfidryls is totally unhindered on the protein external surface in Pf-FtP77C, partially hindered in the internal cavity in Af-FtM54C and totally hindered in the internal cavity in Af-FtM54C/K150A/R151A and Pf-FtG52C. The overall picture that emerges from the body of experimental results is that even fully assembled, closed Af-FtM54C/K150A/R151A and Pf-FtG52C structures are capable of stoichiometric binding of the bulky, negatively charged, DTNB ligand. These results indicate that the permeation of negatively charged molecular species of 8–10 Å length and 5–6 Å diameter into archaeal ferritins does occur at variance with the reported exclusive entry of positively charged or neutral species only in vertebrate ferritins [13,14]. Charge selectivity was reported in the case of horse spleen ferritin (HoSF) and human H-chain ferritin (HuHF) and was correlated to the nature of threefold channels, endowed with hydrophilic, negatively charged tunnels involved in iron uptake [23]. Electron paramagnetic resonance spectroscopy and gel permeation chromatography studies with HoSF demonstrated that molecular charge and polarity of the diffusants play a critical role in their permeation into ferritin [14]. Kinetic studies of permeation using small nitroxide spin probes also confirmed the role of these channels as providing a charge-selective pathway for entry into the cavity [10]. A negatively charged nitroxide was completely excluded from the interior of the protein, whereas positively charged and polar nitroxide radicals penetrated the protein shell to interact with the iron core. Mutated HuHF (D131H/E134H), where the negatively charged glutamate and aspartate residues lining the threefold channels are replaced by histidines, partially allowed the entry of negatively charged species (4-carboxyTEMPO radical) thus providing strong evidence that the negatively charged threefold channels are the principal pathways for molecular

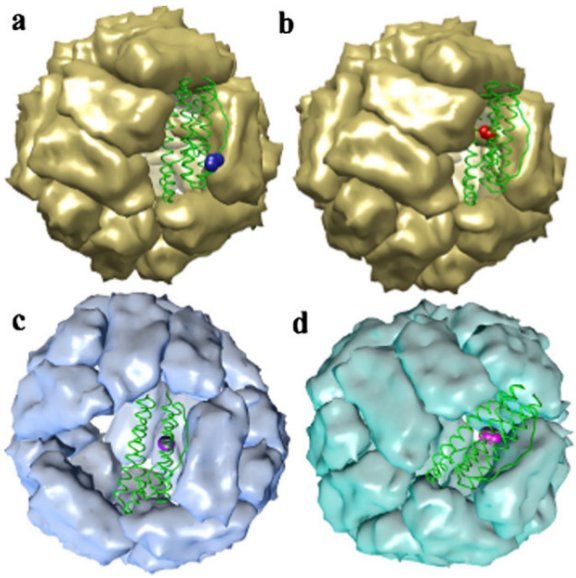


Fig. 3. Three-dimensional structures of the ferritin mutants. A monomer is depicted as green ribbon and the cysteine residues are represented in CPK style for clarity. a) Pf-FtP77C: the external cysteine is shown as blue spheres. b) Pf-FtC52C: the internal cysteine is depicted in red (models a) and b) built on Pf-Ft structure, PDB ID: 2JD6). c) Af-Ft54C: the internal cysteine is depicted in purple. d) Af-FtM54C/K150A/R151A: the internal cysteine in magenta (models c) and d) built on Af-Ft structure, (PDB ID: 1SQ3)). Molecular graphics and analyses were performed with the UCSF Chimera package [25].

diffusion into ferritin [10]. Interestingly, first-order half-lives for permeation of positively charged compounds in mammalian ferritins are in the same time range of those observed in DTNB uptake in fully closed Af-FtM54C/K150A/R151A and Pf-FtG52C archaeal ferritins although Af-Ft and Pf-Ft have been reported to show a different amino acid arrangement in the threefold channels that suggested a different ligand incorporation mechanism.

In particular, the eight threefold channels in each protein are shaped by the C-terminal segment of C helix and the N-terminus of D helix in the 4-helix bundle of three adjacent monomers [24]. Thus, C-terminal ends of helices C define the outer entrance to the channel and the N-termini of helices D define the inner entrance. In the case of Pf-Ft, channels are similar to those of HuHF in the outer side, that is rich in negative aminoacids (Glu109, Glu110 and Glu111), whereas they differ in the central region due to the presence of polar and positively charged residues (Tyr114 and Arg117). Lastly, the inner entrance of the channel is characterized by the presence of Ala118 and Glu121, in the topological position which is deemed essential for the transfer of iron (II) to the ferroxidasic center in mammalian ferritins [26].

In the case of Af-Fts, namely Af-FtM54C ("open" form) and Af-FtM54C/K150A/R151A ("closed" form), the threefold channels are lined by a mix of hydrophobic and hydrophilic aminoacids: only one glutamate (Glu113; PBD ID: 1SQ3 and 3KX9) is located to the entrance and two acidic residues lining HuHF channels (D131 and E134) (PDB ID:2FHA) are replaced by neutral and positively charged residues (Tyr119 and Asn 120). Thus, in the case of Af-Ft, the cluster of negative charges characteristic of HuHF is not present, presumably because it is not necessary to drive iron (II) atoms towards the ferroxidasic site, readily available through the large pores present in this protein. Hence, the threefold channels of archaeal ferritins are less negatively charged than the mammalian ones, possibly allowing the entrance of negatively charged compounds (see Fig. 4). Of course, the "open" Af-FtM54C, here investigated, displays the four 45 Å wide triangular pores, defined by C and D helices, characterized by the presence of a group of positive charged residues located at the three apices [17]. These large openings are most likely preferential paths for large ligand entry into the cavity as demonstrated by the 8 fold faster DTNB reaction rate with the internal thiol with respect to the closed Af-FtM54C/K150A/R151A mutant (see Table 1). In turn, the DTNB entry rate into Af-FtM54C/K150A/R151A mutant is 5 fold faster than in the Pf-FtG52C protein. Such difference is significant in view of the fact that the last two proteins are assimilated to closed cages with thiols in topologically equivalent positions.

As a last comment it is worth considering the possible role of the so-called fourfold channels. As matter of fact, the quaternary assembly of the closed Af-FtM54C/K150A/R151A mutant is similar to that of Pf-FtG52C and other mammalian Ft cages displaying a canonical octahedral conformation with fourfold channels of 4–5 Å size. The six fourfold channels of Pf-Ft (PDB ID: 2JD6) are polar and hydrophilic. Moreover, Lys145 in the DE loop (corresponding to Lys 150 in Af-Ft), which forms an outer gate to the fourfold channel, makes this entrance more polar and charged than that of HuHF, which is basically uncharged. Nevertheless, the dimensions of the fourfold channel in both mammalian and archaeal ferritins is found to be too small even to metal ions with the possible exception of protons and cannot be taken as a possible option for the entry of organic molecules.

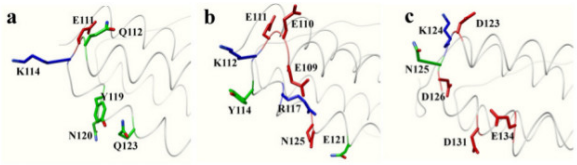


Fig. 4. Profile view of the lining amino acids in the threefold channel in a) Af-Ft (PDB ID: 1SQ3) and b) Pf-Ft (PDB ID: 2JD6) with respect to c) HuHF (PDB ID: 2FHA). The exterior of the shell lies on the left side and the inner cavity on the right side of each cartoon as shown schematically. Positive, negative and polar residues are depicted as blue, red and green sticks, respectively.

Thus, on the basis of the observed DTNB binding rates and available crystallographic structures we conclude that archaeal ferritins from *P. furiosus* and *A. fulgidus* are able to incorporate negatively charged, modestly sized diffusants, even in their fully "closed" forms, most likely through the threefold channels whose nature appears less restrictive with respect to that of vertebrate proteins. Possibly, the presence of positively charged residues in the middle of the threefold channel may favour entrance of negatively charged species in the archaeal proteins. As a last comment, it must also be pointed out that the dimensions of the currently used probe slightly exceed the diameter of the threefold channel as inferred from crystal structure coordinates. A mechanism of rotameric adjustment of relevant aminoacid side chains is thus necessary in order to allow for the entry/exit of small organic molecules. Recent demonstration of multiple conformers in aminoacids lining the inner entrance to the ferritin cavity may explain the necessary plasticity of the threefold channels in ferritins [26].

5. Conclusions

The results of the present work indicate that molecular diffusion into archaeal ferritin is a complex phenomenon and that even apparently closed, impermeable structures, ferritins do allow entry of 8–10 Å long organic molecules with no necessity of 24-mer desassembly. The data thus obtained indicate clearly that the protein matrix in archaeal ferritins does not provide a significant barrier against bulky, negatively charged ligands such as DTNB, a finding of relevance in view of the multiple biotechnological applications of these ferritins that envisage ligand encapsulation within the internal cavity. The potential impact of such engineered ferritins on the general topics related to their biotechnological application is wide. On the side of material science, the possibility of disposing of a cations dependent self-assembling cage provides the basis of unique "molecular carpentry" tools. On the side of biomedical applications, the use of archaeal proteins must still be explored. The entry of archaeal ferritins into mammalian cells is under investigation in order to clarify their receptor recognition properties, the pathways of intracellular trafficking, the impact on eukaryotic cell iron homeostasis and/or their toxic effects due to possible generation of oxidative stress under the widely different iron reductive mechanisms typical of eukaryotic cells.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Acknowledgements

EU H_{2020} Project "X-Probe", Grant N° 637295, to A. Bonamore and M.C.T. is gratefully acknowledged. Flagship project "Nanomax" from MIUR (Nanomax MIUR_J120062) to A. Boffi is also acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.bbagen.2016.10.007.

References

[1] K. Honarmand Ebrahimi, P.L. Hagedoorn, W.R. Hagen, Unity in the biochemistry of the iron-storage proteins ferritin and bacterioferritin, Chem. Rev. 115 (2015) 295–326.
[2] Z. Zhen, W. Tang, H. Chen, X. Lin, T. Todd, G. Wang, RGD-modified apoferritin nanoparticles for efficient drug delivery to tumors, ACS Nano 7 (2013) 4830–4837.
[3] N. Pontillo, F. Pane, L. Messori, A. Amoresano, A. Merlino, Cisplatin encapsulation within a ferritin nanocage: a high-resolution crystallographic study, Chem. Commun. (Camb.) 52 (2016) 4136–4139.
[4] E. Falvo, E. Tremante, R. Fraioli, C. Leonetti, C. Zamparelli, A. Boffi, V. Morea, P. Ceci, P. Giacomini, Antibody-drug conjugates: targeting melanoma with cisplatin encapsulated in protein-cage nanoparticles based on human ferritin, Nanoscale 5 (2013) 12278–12285.
[5] X.-T. Ji, L. Huang, H.-Q. Huang, Construction of nanometer cisplatin core-ferritin (NCC-F) and proteomic analysis of gastric cancer cell apoptosis induced with cisplatin released from the NCC-F, J. Proteome 75 (2012) 3145–3157.
[6] J.M. Domínguez-Vera, Iron(III) complexation of desferrioxamine B encapsulated in apoferritin, J. Inorg. Biochem. 98 (2004) 469–472.
[7] A. Maham, Z. Tang, H. Wu, J. Wang, Y. Lin, Protein-based nanomedicine platforms for drug delivery, Small 5 (2009) 1706–1721.
[8] A. Maham, H. Wu, J. Wang, X. Kang, Y. Zhang, Y. Lin, Apoferritin-based nanomedicine platform for drug delivery: equilibrium binding study of daunomycin with DNA, J. Mater. Chem. 21 (2011) 8700–8708.
[9] L. Vannucci, E. Falvo, C.M. Failla, M. Carbo, M. Fornara, R. Canese, S. Cecchetti, L. Rajsiglova, D. Stakheev, J. Krizan, A. Boffi, G. Carpinelli, V. Morea, P. Ceci, In vivo targeting of cutaneous melanoma using a melanoma stimulating hormone-engineered human protein cage with fluorophore and magnetic resonance imaging tracers, J. Biomed. Nanotechnol. 11 (2015) 81–92.
[10] X. Yang, P. Arosio, N.D. Chasteen, Molecular diffusion into ferritin: pathways, temperature dependence, incubation time, and concentration effects, Biophys. J. 78 (2000) 2049–2059.
[11] X. Yang, N.D. Chasteen, Molecular diffusion into horse spleen ferritin: a nitroxide radical spin probe study, Biophys. J. 71 (1996) 1587–1595.
[12] B. Zhang, R.K. Watt, N. Gálvez, J.M. Domínguez-Vera, G.D. Watt, Rate of iron transfer through the horse spleen ferritin shell determined by the rate of formation of prussian blue and Fe-desferrioxamine within the ferritin cavity, Biophys. Chem. 120 (2006) 96–105.
[13] R.R. Crichton, J.P. Declercq, X-ray structures of ferritins and related proteins, Biochim. Biophys. Acta 1800 (2010) 760–818.
[14] T. Douglas, D.R. Ripoll, Calculated electrostatic gradients in recombinant H-chain ferritin, Protein Sci. 7 (5) (1998 May) 1083–1091.
[15] P.M. Harrison, A. Treffry, T.H. Lilley, Ferritin as an iron-storage protein: mechanisms of iron uptake, J. Inorg. Biochem. 27 (1986) 287–293.
[16] A. Treffry, P.M. Harrison, Spectroscopic studies on the binding of iron, terbium, and zinc by apoferritin, J. Inorg. Biochem. 21 (1984) 9–20.
[17] E. Johnson, D. Cascio, M.R. Sawaya, M. Gingery, I. Schröder, Crystal structures of a tetrahedral open pore ferritin from the hyperthermophilic archaeon *Archaeoglobus fulgidus*, Structure 13 (2005) 637–648.
[18] B. Sana, E. Johnson, P. Le Magueres, A. Criswell, D. Cascio, S. Lim, The role of nonconserved residues of *Archaeoglobus fulgidus* ferritin on its unique structure and biophysical properties, J. Biol. Chem. 288 (2013) 32663–32672.
[19] P.W. Riddles, R.L. Blakeley, B. Zerner, Ellman's reagent: 5,5′-dithiobis(2-nitrobenzoic acid) —a reexamination, Anal. Biochem. 94 (1979) 75–81.
[20] B.J. Berne, R. Pecora, Dynamic Light Scattering, Wiley, New York, 1976.
[21] 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. U. S. A. 111 (2014) 14900–14905.

[22] 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 (2013) 52–56.
[23] A. Desideri, S. Stefanini, F. Polizio, R. Petruzzelli, E. Chiancone, Iron entry route in horse spleen apoferritin. Involvement of the three-fold channels as probed by selective reaction of cysteine-126 with the spin label 4-maleimido-tempo, FEBS Lett. 287 (1991) 10–14.
[24] J. Tatur, W.R. Hagen, P.M. Matias, Crystal structure of the ferritin from the hyperthermophilic archaeal anaerobe *Pyrococcus furiosus*, J. Biol. Inorg. Chem. 12 (2007) 615–630.
[25] 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 (2004) 1605–1612.
[26] R.K. Behera, E.C. Theil, Moving Fe^{2+} from ferritin ion channels to catalytic OH centers depends on conserved protein cage carboxylates, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 7925–7930.

Appendix II

Nanoscale

PAPER

Humanized archaeal ferritin as a tool for cell targeted delivery†

Cite this: *Nanoscale*, 2017, **9**, 647

Valeria de Turris,[a] Matilde Cardoso Trabuco,[b] Giovanna Peruzzi,[a] Alberto Boffi,[a,c] Claudia Testi,[a] Beatrice Vallone,[c] Linda Celeste Montemiglio,[d] Amédée Des Georges,[e] Lorenzo Calisti,[d] Irene Benni,[d] Alessandra Bonamore[d] and Paola Baiocco*[a]

Received 8th September 2016,
Accepted 23rd November 2016

DOI: 10.1039/c6nr07129e

www.rsc.org/nanoscale

Human ferritins have been extensively studied to be used as nanocarriers for diverse applications and could represent a convenient alternative for targeted delivery of anticancer drugs and imaging agents. However, the most relevant limitation to their applications is the need for highly acidic experimental conditions during the initial steps of particle/cargo assembly, a process that could affect both drug stability and the complete reassembly of the ferritin cage. To overcome this issue the unique assembly of *Archaeoglobus fulgidus* ferritin was genetically engineered by changing a surface exposed loop of 12 amino acids connecting B and C helices to mimic the sequence of the analogous human H-chain ferritin loop. This new chimeric protein was shown to maintain the unique, cation linked, association–dissociation properties of *Archaeoglobus fulgidus* ferritin occurring at neutral pH values, while exhibiting the typical human H-homopolymer recognition by the transferrin receptor TfR1. The chimeric protein was confirmed to be actively and specifically internalized by HeLa cells, thus representing a unique nanotechnological tool for cell-targeted delivery of possible payloads for diagnostic or therapeutic purposes. Moreover, it was demonstrated that the 12 amino acids' loop is necessary and sufficient for binding to the transferrin receptor. The three-dimensional structure of the humanized *Archaeoglobus* ferritin has been obtained both as crystals by X-ray diffraction and in solution by cryo-EM.

Introduction

Ferritin proteins have been extensively used as nanocarriers for diverse applications due to their hollow cage-like structures and their unique, reversible, 24-mer assembly.[1,2] In more detail, they represent the most convenient alternative to viral carriers for targeted delivery of anticancer drugs and imaging agents and have been successfully utilized as reaction nanovessels for the synthesis of non-native metallic nanoparticles in the inner core, with applications in nanoelectronic devices.[3–5] The external and internal surfaces of ferritin are chemically and genetically modifiable allowing for the attachment site for drugs, nucleic acids, fluorophores or magnetic moieties. Along this line, recent studies further established heavy (H) or light (L)-chain homopolymers as versatile multifunctional nanocarriers for targeted cancer diagnosis and therapy.[6–10] In fact, human ferritins constitute biocompatible nanocarriers that stabilize and shelter the enclosed particles, thus preventing immunogenic responses. Moreover, ferritins are naturally targeted toward ubiquitously expressed TfR1 transferrin receptors (H-chain specific) or hepatic SCARA 5 receptors (L-chain specific). Such properties have been widely exploited for the efficient delivery of antitumor drugs to iron-avid, fast replicating, tumor cells overexpressing the TfR1 receptor.[11] Thus, ferritin based protein cages have been developed as versatile platforms for multiple applications in nanomedicine.

Current development of human ferritin based particle is however facing intrinsic limitations due to the experimental conditions linked to their assembly–disassembly equilibrium, whose control is a prerequisite in order to achieve encapsulation of the cargo within the internal cavity. In the conventional *in vitro* encapsulation procedures, ferritin must be disassembled at extreme pH values (pH = 2.0) and re-assembled in the presence of highly concentrated payload compounds.[12]

[a]*Center for Life Nano Science@Sapienza, Istituto Italiano di Tecnologia, V.le Regina Elena 291, Rome 00161, Italy. E-mail: paola.baiocco@iit.it*
[b]*Molirom srl, via Ravenna 8, 00161 Rome, Italy*
[c]*Institute of Molecular Biology and Pathology, National Research Council, P.le A. Moro, 7, 00185, Rome, Italy*
[d]*Department of Biochemical Sciences "Alessandro Rossi Fanelli", Sapienza University of Rome, P.le A. Moro, 5, 00185, Rome, Italy*
[e]*The City University of New York Advanced Science Research Center 85 St. Nicholas Terrace, New York, NY 10031, USA*
†Electronic supplementary information (ESI) available. See DOI: 10.1039/c6nr07129e

This journal is © The Royal Society of Chemistry 2017

Such a procedure leads to a suboptimal load of cargo material, whose chemical structure is required to be pH resistant. Moreover, the pH jump procedure is only partially reversible and re-assembly may not be complete, depending on complex equilibrium and kinetic parameters.[13] Currently, extensive research efforts are devoted to adjust the assembly properties of ferritin nanocages to the desired applications either by intersubunit interface mutagenesis[14] or by genetic engineering of N- or C-terminal regions.[15]

Recently, novel ferritins from lower eukaryotes, bacteria and archaea, endowed with different polymer association–dissociation thermodynamic and kinetic features have emerged as possible alternatives to human ferritin homopolymers for several biotechnological applications requiring cargo material encapsulation.[16,17] Despite low sequence similarity, these ferritins display a highly conserved quaternary structure consisting of a four-helix bundle namely A, B, C and D and a short E helix at the C-terminus. Helices B and C are connected by a 12 amino acid (aa) long loop involved in stabilizing interactions at the 2-fold dimer interface. Unlike most eukaryotic and prokaryotic ferritins, the archaeal ferritin from *Archaeoglobus fulgidus* (AfFt) is characterized by unique self-assembly properties. In fact, in neutral buffers it is present as dimeric species, easily combining into a non-canonical 24-mer cage in the presence of metal cations.[18] AfFt assembles in a distinctive tetrahedral geometry as a result of particular packing between four hexametric units into a 24-mer structure different from those observed so far. Such unusual assembly results in the formation of four wide triangular pores (45 Å) on the protein shell.[19] As demonstrated by Sana *et al.*,[19] amino acid substitutions in the turn motif that connects D and E helices forming the 4-fold iron channels, namely K150A and R151A, are sufficient to restore the canonical octahedral symmetry observed in vertebrate and bacterial ferritins,[18,19] thus providing a rationale for the atypical tetrahedral architecture of AfFt. Nevertheless, the low sequence similarity of AfFt with mammalian ferritins runs against the possibility of targeting AfFt onto the TfR1 receptors in mammalian cells.

Transferrin receptor TfR1, or CD71, has been reported to be a preferred target for human ferritin, due to the specific interaction of the receptor extracellular moiety with epitopes of the H-ferritin subunit.[20,21] Inspection of the three-dimensional structure of the human H-homopolymer external surface indicates that the most significant accessible area is occupied by the external 12 aa long loop connecting B and C helices (up to 19 aa including the turn regions). Patches of a lesser extent exposed to solvent are formed by the N-terminal regions or by the iron channels within the threefold axis of the intersubunit assembly. However, it has been shown that N-terminal or C-terminal deleted human H-homopolymers are efficiently taken up by target cells, thus suggesting these regions are not required for receptor recognition or uptake.[22] Therefore the BC loop, besides its structural role in stabilization of the interdimer interface,[23] appeared to be the best candidate for TfR1 receptor recognition of the ferritin molecule.

We decided then to engineer the archaeal ferritin to devise a chimeric construct, named humanized *Archaeoglobus* ferritin (HumAfFt), in which the external 12 aa loop connecting the B and C helices was mutated to reproduce an analogous one in the corresponding region of the human H chain homopolymer (HuHF). It was observed that this chimeric protein (HumAfFt) was actively internalized by HeLa cells to an extent comparable to transferrin, the preferred ligand for TfR1.

Results and discussion

"Humanized" *Archaeoglobus* ferritin design

Heavy chain human ferritin and *Archaeoglobus* ferritin display 31% sequence identity. The structural alignment of HuHF and AfFt monomers (pdb code 3AJO and 1S3Q, respectively) reveals a shorter N-terminus and shorter loops between three of the four helices, with the notable exception of the long loop connecting B and C helices.

As shown in Fig. 1, the BC loops of two adjacent subunits run in an antiparallel fashion establishing significant, mutual interactions. Thus the BC loops of adjacent subunits display the same overall geometry in both HuHF and AfFt, though each couple of loops adopts a different symmetry along the spherical surface of the protein cage, given the different dimer–dimer positioning within each complete 24-mer structure.

In order to preserve the unique assembly properties of AfFt, while implementing a potential cellular uptake, we decided to exploit this external loop by mutating 9 residues into the corresponding residues found in HuHF, according to the sequence alignment shown in Fig. 2. In particular, the nine amino acids sequence IFLQDIKKP, typical of a human H ferritin homopolymer, located at the center of the 12 aa loop was inserted in place of VKLYAVEEP (from residue 70 to 79 in AfFt numbering, see Fig. 2). In addition, a cysteine residue in position 54 (AfFt numbering) has been introduced by point mutation in order to provide a conjugation site of potential thiol reactive derivatives into the cavity. As previously mentioned, we will refer to the chimeric protein described above as humanized *Archaeoglobus* ferritin (HumAfFt).

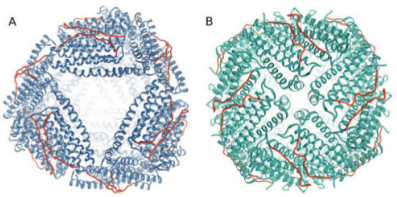


Fig. 1 Three-dimensional structures determined by X-ray crystallography. A cartoon representation of (A) HumAfFt and (B) HuHF (pdb code 3AJO). Models are coloured in blue and green, respectively, and the external loop connecting the helices B and C of each monomer is shown as red ribbons. Molecular graphics were performed using a UCSF Chimera package.[35]

This journal is © The Royal Society of Chemistry 2017

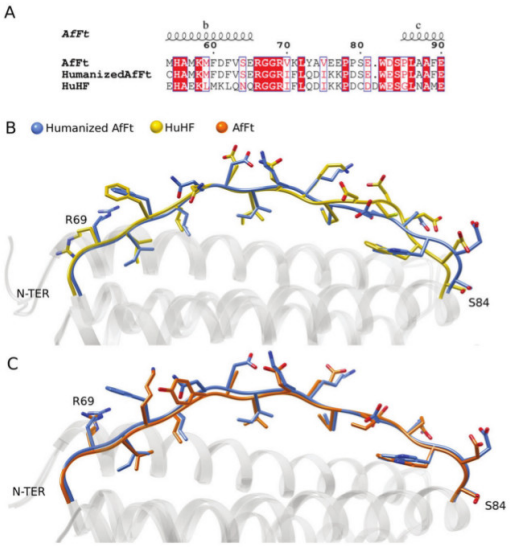


Fig. 2 Structure-based sequence alignment of AfFt, HumAfFt, and HuHF. (A) Close-up view of the sequence alignment. Elements of secondary structure for the AfFt are shown on the top. White characters in a red background indicate strict conservation while residues with poor conservation are drawn in black on a white background. Alignments were made using CLUSTAL Omega, and the figure was generated using ENDSCRIPT.[37] The structural superposition of the region from R69 and S84 (AfFt numbering) is shown in (B) HumAfFt (in blue) *vs.* HuHF (in yellow) and in (C) HumAfFt (in blue) *vs.* AfFt (in orange). Residues are depicted as sticks indicating N atoms in blue, O atoms in red and S atoms in yellow.

Self-assembly assessment

The effect of mutations on the $MgCl_2$-mediated self-assembly of HumAfFt was studied by size exclusion chromatography (SEC), in order to separate different possible oligomers according to their molecular size. Identical peak-positions confirmed that HumAfFt retains the $MgCl_2$-mediated self-assembly properties of native AfFt[24] (see Fig. S2†). As shown in the chromatograms, the increasing of $MgCl_2$ concentration triggered the self-assembly of dimers until they reached a stable polymeric structure around 500 kDa, roughly corresponding to the expected 24-mer cage-like structure, at 20 mM $MgCl_2$. The data highlighted that the chimeric HumAfFt maintained the cation induced association/dissociation properties of archaeal ferritin and is possibly assembled into a 24-mer typical structure.[24]

X-ray diffraction data on HumAfFt show the tetrahedral symmetry of archaeal ferritins

The humanized AfFt crystallized under different conditions with respect to wild type AfFt. Crystals were exposed to the synchrotron light in order to confirm the tetrahedral symmetry reported for archaeal ferritins. The protein was crystallized in the presence of Mg^{2+} in order to maintain the 24-mer assembly.

The structure of HumAfFt has been determined by X-ray crystallography at a 2.87 Å resolution. It crystallized in the $C222_1$ space group, as also observed for wild type AfFt. The asymmetric unit (ASU) contains 12 identical subunits with a solvent content of 64.3%. The overall folding corresponded to the wild type structure (pdb code 1S3Q) with a rmsd value of 0.4 Å and displayed four wide triangular pores on the surface (Fig. 1A). B factor analysis shows a mean B factor of 66 $Å^2$ with the exception of the loop region between the D and E helices, which displays higher B factors and a poor electron density on the side chains from 146 to 151 residues. Conversely, the BC loop is well organized and the analysis of the difference electron density (Fo–Fc) map clearly showed the presence of the mutated residues in the loop between the helices B and C as well as the M54C mutation, pointing towards the inner cavity.

Two magnesium ions have been positioned and successfully refined in the ASU, and they are located in the ferroxidasic site of two different subunits. In both, one magnesium

ion is coordinated with both OE1 and OE2 of Glu19, (at 2.5 Å and 2.8 Å distance, respectively), with Glu52 (OE1) at 2.5 Å, with Gln129 (OE1) at 2.6 Å and with a water molecule at 2.6 Å distance. In the other chains, a water molecule has been modelled in the Fo–Fc map and successfully refined in the same position.

The loop region, including the conserved terminal turns, spans from amino acid 68 to amino acid 86. The sequence alignment and the structural superposition between HumAfFt and the human H-homopolymer, as well as between HumAfFt and AfFt, are shown in Fig. 2B and C. At the dimeric interface between the two antiparallel BC loops, the main interactions are a hydrogen bond between Arg69(NH2) and Ser80 (O) at 2.8 Å distance and two salt bridges, namely Lys71 (NZ)–Glu77 (OE1) at 3.0 Å distance and Glu81 (OE2)–Arg69 (NH1) at 2.7 Å distance. A weak salt bridge is established by Glu81 (OE2) and Arg69 (NH2), at a distance of 3.0 Å. The salt bridge between Lys71 and Glu77 observed in HuHF, is absent in HumAfFt since these positions were mutated into a phenylalanine (Phe71) and a lysine (Lys77). Other interactions are conserved between HuHF and HumAfFt and measured the same distances.

Cryo-electron microscopy confirms the canonical AfFt architecture in solution

Humanized AfFt samples were prepared in thin ice layer and analyzed by cryo-EM in order to assess the three-dimensional structure of the chimeric protein in a near-native environment. Particles were picked with a reference-based automated particle picking procedure. Single ferritin particles were visualized at a nominal resolution of 33 Å demonstrating that their shape and dimensions correspond to that of wild type AfFt in its 24-mer assembly (Fig. 3A). Moreover, the triangular apertures on the protein surface were clearly observable in a bi-dimensional view (Fig. 3B) and better displayed in a three-dimensional reconstruction as shown in Fig. 3C.

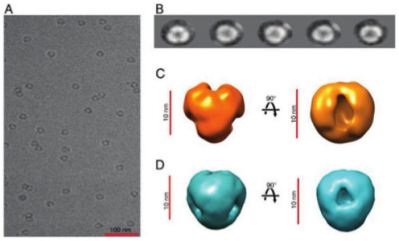


Fig. 3 Electron microscopy characterization of humanized *Archaeoglobus fulgidus* ferritin. (A) Sample micrograph of the HumAfFt data set. Scale bar: 100 nm. (B) Five representative 2D class averages obtained with RELION. (C) 3D reconstruction of HumAfFt obtained with RELION and visualized with UCSF Chimera.[35] Map final resolution: 33.1 Å. Scale bars: 10 nm. Left: side view. Right: top view. (D) *Archaeoglobus fulgidus* ferritin crystal structure (from 16) filtered to 30 Å, shown for comparison. Scale bars: 10 nm. Left: side view. Right: top view.

Flow cytometry analysis shows HumAfFt cellular uptake

After we have demonstrated that our HumAfFt maintained its structure with large open pores and the self-assembly property characteristic of the original AfFt, we aimed to verify that the modified version also gained access to eukaryotic cancer cells such as HeLa cells. It is known that HuHF is recognized and internalized by the TfR1, which is overexpressed in many types of tumor cells but not in normal cells and healthy tissues.[21] To validate the effect of our mutations on the external loop related to the uptake efficiency by HeLa cells, we performed time course experiments on cells treated with the same amount (30 μg ml^{-1}) of AfFt-FITC, HumAfFt-FITC and transferrin-FITC (TF-FITC) and analyzed them by flow cytometry. As a baseline for FITC fluorescence, control cells not incubated with FITC-ferritins were used. Moreover, to exclude any signal generated from outside particles sticking on the cell membrane due to unspecific binding or remains from the washing steps, trypan blue quenching was performed before using FACS. In Fig. 4, the FACS analysis is summarized, shown as the percentage of cells internalizing the nanoparticles at different times. These data highlighted that HumAfFt nanoparticles are efficiently taken up by HeLa cells already after one hour incubation with a much higher percentage compared to AfFt (81% and 5% respectively). After a

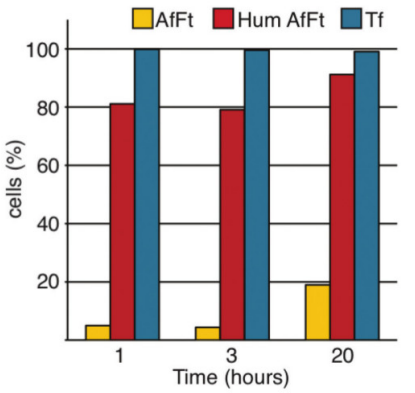


Fig. 4 Humanized AfFt is internalized with higher efficiency than the original ferritin. Ferritins taken up in HeLa cells have been quantified by flow cytometry. Cells have been treated with 30 μg ml^{-1} of AfFt-FITC, HumAfFt-FITC and transferrin-FITC (Tf). The percentage of cells internalizing the nanoparticles at the time indicated is shown. For each sample 30 000 events gated on live cells have been acquired.

This journal is © The Royal Society of Chemistry 2017

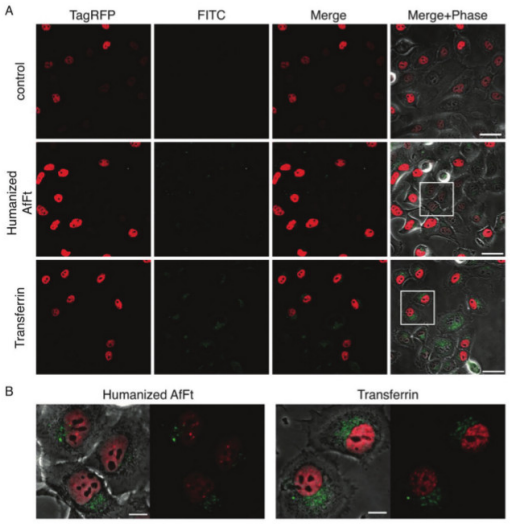


Fig. 5 Ferritin internalization observed at the confocal microscope. Cells were plated on an 8-well μ-slide (ibidi), induced with doxycycline to express TagRFP and then incubated with 30 μg ml^{-1} of humanized AfFt-FITC or transferrin-FITC for 20 h. After washing, cells were live-imaged using a confocal microscope. (A) Confocal images of live HeLa TagRFP cells are shown as single, merged channels and overlay images with the phase contrast. Scale bars: 40 μm. (B) Images are a magnified view of the region highlighted by the white inset in panel A; merged channels and the overlay with the phase contrast images are shown. Scale bars: 10 μm.

longer incubation time (20 hours), the FITC-positive cells for the humanized samples are increased to more than 90% whereas for native AfFt they are still less than 20%. The latter increment is possibly due to the unspecific uptake by pinocytosis. Each acquired plot is shown in Fig. S4.†

Moreover we performed the RNAi experiment against the transferrin receptor and analyzed the cellular uptake by FACS (Fig. S5†) obtaining around 50% HumAfFt uptake reduction after 48 hours of siRNA transfection, confirming the involvement of TfR1 in the internalization process.

Visualization of ferritin nanoparticles inside living cells by confocal microscopy

In order to visualize uptaken HumAfFt–FITC nanocages, we performed the internalization assay directly on an ibidi 8-well μ-slide and incubated the cells for 20 hours before confocal microscopy. We exploited the properties of the HeLa TagRFP cell lines available in the lab in order to have a reference fluorescence signal confirming that we were imaging inside the cell. This cell line contains a TagRFP-FUS protein under the control of a doxycycline-inducible promoter that allows for a controlled expression of the protein and hence permits visualization of the nucleus in the red channel. In this case the cells were not selected to eliminate the untransfected cells, allowing us to visualize different levels of expression and also unstained nuclei. Just before imaging, the cells were washed to eliminate the unbound FITC-nanoparticles and then acquired by confocal laser-scanning microscopy. Confocal representative images of the entire field of view of live HeLa TagRFP cells alone (control) or incubated with HumAfFt-FITC or TF-FITC are shown in Fig. 5A. A detailed view of the boxed region in panel A is shown in Fig. 5B. Images confirmed the high extent of HumAfFt internalization and highlight a cellular distribution in the cytoplasm and in the perinuclear space comparable to that observed in the case of transferrin, thus suggesting a typical clathrin-coated endocytosis pathway, mediated by TfR1.

Conclusion

The present data highlight the remarkable properties of a novel chimeric ferritin nanocage suitable for the design of efficient and versatile scaffolds for the intracellular delivery of bioactive small molecules and/or diagnostic probes.

This journal is © The Royal Society of Chemistry 2017

Archaeoglobus fulgidus ferritin (AfFt) is a prominent example of this versatility due to its unique association/dissociation properties that lead to the presence of stable dimeric species at neutral pH and low ionic strength capable of associating into non-classical 24-mer species in the presence of either monovalent or divalent cations at physiological concentrations (*i.e.* higher than 0.5 M Na^+ or 20 mM Mg^{++}). Moreover, AfFt also displays a unique subunit assembly, based on tetrahedral symmetry, which leads to the formation of four large openings in the protein shell. As such, AfFt represents a uniquely suitable scaffold for incorporating a wealth of diverse substructures inside the protein cavity, either by an assembly/disassembly process at neutral pH or by diffusion through the large triangular pores on the surface. Notable examples have been reported in recent literature.[17] Nevertheless, one of the key properties of ferritin nanocages in biomedical applications is the possibility of targeting receptors on human cells, thus allowing the delivery of the desired payload within the cytoplasm. The engineered HumAfFt described here combines the versatility in assembly and cargo incorporation of AfFt by binding to TfR1 and by cellular uptake of HuHF. At present, further experiments with possible payloads for diagnostic or therapeutic applications are under development. However, *in vivo* applications will have to wait for the evaluation of possible immunological responses against non-human epitopes present on the protein surface. In fact, though the mutated loops appear to account for most of the solvent exposed surface, N-terminal and pore lining regions might still offer windows for non-self-recognition by the immune system.

Materials and methods

"Humanized" *Archaeoglobus* ferritin design

The gene encoding for a mutated ferritin from *Archaeoglobus fulgidus* was synthesised by GeneArt (ThermoFisher) and subcloned into a pET22b vector (Novagen) between the restriction sites NdeI and HindIII at 5′ and 3′ respectively. The recombinant plasmid was transformed into BL21(DE3) *E. coli* strain for protein expression.

Protein expression and purification

E. coli cells, containing the HumAfFt plasmid, were grown and induced with 1 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside) at OD_{600} = 0.6. Cells were harvested by centrifugation 3 hours post induction at 37 °C.

Cells harvested from 1 L culture were resuspended in 20 mM HEPES buffer, pH 7.5, containing 200 mM NaCl, 1 mM TCEP (tris(2-carboxiethyl)phosphine), and a cOmplete™ Mini Protease Inhibitor Cocktail Tablet (Roche). Cells were disrupted by sonication and the soluble fraction was purified by heat treatment at 78 °C for 10 minutes. Denatured *E. coli* proteins were removed by centrifugation at 15 000 rpm at 4 °C for 1 hour. The soluble protein was further purified by ammonium sulfate precipitation. The precipitated fraction at 70% ammonium sulfate was resuspended in 20 mM HEPES, 50 mM $MgCl_2$, pH 7.5 and dialyzed *versus* the same buffer. As a final purification step, the protein was loaded onto a HiLoad 26/600 Superdex 200 pg column previously equilibrated in the same buffer using an ÄKTA-Prime system (GE Healthcare). The purified protein was concentrated to obtain the final protein preparation of 1 mg mL^{-1} and protein concentration was calculated by measuring the UV spectrum using an extinction coefficient of 32 400 M^{-1} cm^{-1}. Protein yield was ∼40 mg L^{-1} culture.

Self-assembly assessment in solution

Size exclusion chromatography $MgCl_2$-mediated self-assembly was studied by size exclusion chromatography (SEC) using a Superdex 200 26/600 GL column (GE Healthcare). The molecular size of HumAfFt was determined under different conditions by comparing the elution volume with that of standard proteins. Composition of the mobile phase was 25 mM HEPES pH = 7.5 with different $MgCl_2$ concentrations accordingly to the composition of the protein buffer.

Crystallization and crystal structure determination

The purified protein was concentrated to 20 mg mL^{-1} and initial crystallization screening was performed using a Phenix Robot. Crystals were obtained by mixing in a 2 μL hanging drop 1 mM of the purified protein with a solution containing 22% (vol/vol) polyacrylic acid PAA, 0.1 M Tris, 0.02 M $MgCl_2$, pH 7.4, at 25 °C within a week, cryo-protected by increasing the precipitant concentration and flash-frozen in liquid nitrogen. Diffraction data have been collected at ID23-2 beamline at the European Synchrotron Radiation Facility (ESRF), Grenoble, France.

Data were processed with XDS[25] and scaled with Aimless (ccp4 suite) at a final resolution of 2.87 Å. The structure was solved by Molecular Replacement with MolRep (ccp4 suite) using the open pore structure AfFt (pdb code 1S3Q) as the search model. Model Building and refinement were done using Coot[26] and Refmac5, respectively. The final model was analyzed with PROCHECK[27] and Molprobity.[28] The Ramachandran plot showed that 97.8% of residues are in preferred regions, 2.2% in allowed regions and no outlier is observed. The final atomic coordinates and structure factors were deposited with the PDB Data Bank (http://www.rcsb.org) with accession code: 5LS9. Complete data collection and refinement statistics are reported in Table S1.†

Cryo-Electron microscopy

Holey-gold grids were prepared as described by Russo and Passmore[29] from Quantifoil R1.2/1.3 (Quantifoil Micro Tools GmbH, Germany). 3 μL of HumAfFt (12 μM) was applied to the holey-gold grids after plasma cleaning with a mixture of H_2 and O_2. Grids were blotted for 4 seconds and vitrified by rapidly plunging into liquid ethane at −180 °C (ref. 30 and 31) with a Vitrobot (FEI).

Data acquisition was done using a FEI Titan Halo (FEI, Eindhoven) operating at 300 kV. Datasets were imaged with a Volta phase-plate (FEI, Eindhoven)[32] and were collected with

 This journal is © The Royal Society of Chemistry 2017

the automated data collection system EPU (FEI, Eindhoven) at a nominal magnification of 59 000× on an FEI Ceta camera (FEI, Eindhoven) with a camera pixel size of 14 μm, corresponding to a calibrated pixel size of 1.49 Å on the specimen scale and with a dose of 50 e^{-} $Å^{-2}$.

Image processing. The particles were picked with the reference-based automated particle picking procedure implemented in RELION 1.3.[33,34] CTF correction was not applied since the data were collected within 200 nm of focus and the first CTF zero crossing was well beyond the achievable resolution of the dataset. Those particles were subjected to 2D classification using RELION with k = 100 classes. Good particles were then subjected to 3D classification using RELION with the number of classes K = 8. Resulting classes were refined with the autorefine procedure in RELION.

Resolution estimation. Reported resolutions are based on the 'gold-standard' protocol with the FSC = 0.143 criterion using soft masks with an 8 pixel soft edge, and were corrected for the effects of the mask on the FSC curve using high-resolution noise substitution.[34]

Maps were visualized using UCSF Chimera.[35]

Cell line generation

The HeLa cell line stably expressing an inducible TagRFP-FUS protein (HeLaTagRFP) was generated by transfection with epB-Puro-TT-RFP-FUS wt plasmid and the piggyBac transposase vector. Plasmid construction is described in ref. 36.

Protein FITC labeling

HumAfFt, AfFt and Olo-transferrin were labeled with fluorescein-isothiocyanide (FITC, ThermoFisher) according to the manufacturer's standard protocol. Briefly, 2 mg mL^{-1} of the purified protein was added with 10-fold molar excess of in protein storage buffer stirring for 2 hours at RT. The non-reacted dye was removed by gel filtration chromatography and the fluorescent dye to protein ratio was determined by UV-spectroscopy. LC-MS spectrometry measurements on HumAfFt-FITC confirmed that >60% of monomers are FITC labeled as reported in Fig. S6.†

Cell cultures and ferritin internalization

HeLa cells were grown at 37 °C in Eagle's MEM supplemented with 10% (v/v) FBS, Glutamax (Invitrogen) and penicillin–streptomycin solution (Sigma). When needed, the cells were induced with doxycycline 0.2 μg mL^{-1}. The internalization assay was performed as follows: after seeding the cells on the relevant substrate depending on the experiment, cells were left one day to attach and then incubated with FITC-ferritin nanoparticles (AfFt-FITC, HumAfFt-FITC or Tf-FITC as specified in each experiment) at the final concentration of 30 μg ml^{-1} for the time indicated (1 h, 3 h or 20 h).

Flow cytometry analysis

For flow cytometry analysis HeLa cells were seeded on multi-well plates. Cells were incubated with FITC-ferritin nanoparticles as described previously, then washed two times with PBS, detached with trypsin-EDTA (Euroclone), washed with PBS and resuspended in BD-FACS flow buffer. Half of each sample was treated with Trypan Blue (TB; Sigma) to quench the FITC signal from membrane-bound nanoparticles that were not internalized. The quenching was performed with 0.04% TB for 10 min on ice. Control cells were treated in the same way but without FITC-ferritin incubation. Internalization of ferritins before and after TB treatments was measured at the BD LSFORTESSA (BD Biosciences, San Jose, CA, USA) equipped with a 488 nm laser and FACSDiva software (BD Biosciences version 6.1.3). Live cells were first gated by forward and side scatter area (FSC-A and SSC-A) plots, then detected in the green channel for FITC expression (530/30 nm filter) and side scatter parameter. The gate for the final detection was set in the control sample. Data were analyzed using FlowJo9.3.4 software (Tree Star, Ashland, OR, USA).

Confocal microscopy of live cells

To visualize ferritin internalization by live cells using a confocal microscope, HeLa TagRFP cells were seeded on a μ-slide 8-well ibiTreat (ibidi) and induced with 0.2 μg mL^{-1} of doxycycline. Cells were then incubated with FITC-ferritin nanoparticles as previously described for 20 h and, before microscopy, cells were washed two times with an imaging medium (DMEM without phenol red, 10% FBS, 10 mM Hepes, Glutamax and penicillin–streptomycin solution) to eliminate the unbound FITC-nanoparticles. The confocal laser-scanning microscope used was an Olympus FV10i platform equipped with a built-in incubator. Images were acquired with a 60×/1.2NA water-immersion objective, LD lasers, 473 nm and 559 nm, and filter sets for FITC and TRITC. Phase-contrast images were acquired simultaneously.

Acknowledgements

We thank Alessandro Rosa and Riccardo De Santis for the HeLa TagRFP cell line. EU H_{2020} Project "X-Probe", Grant No. 637295, to A. B. and M. C. T. is gratefully acknowledged. Flagship Project "Nanomax" to A. B. is acknowledged. Regional project FILAS to B. V. is also acknowledged.

References

1 E. C. Theil, Ferritin protein nanocages use ion channels, catalytic sites, and nucleation channels to manage iron/oxygen chemistry, *Curr. Opin. Chem. Biol.*, 2011, **15**, 304–311.

2 L. Schoonen and J. C. van Hest, Functionalization of protein-based nanocages for drug delivery applications, *Nanoscale*, 2014, **6**, 7124–7141.

3 A. Prastaro, P. Ceci, E. Chiancone, A. Boffi, G. Fabrizi and S. Cacchi, Homocoupling of arylboronic acids and potassium aryltrifluoroborates catalyzed by protein-stabilized

palladium nanoparticles under air in water, *Tetrahedron Lett.*, 2010, **51**, 2550.
4 F. Meng, B. Sana, Y. Li, Y. Liu, S. Lim and X. Chen, Bioengineered tunable memristor based on protein nanocage, *Small*, 2014, **10**, 277–283.
5 M. A. Kostiainen, P. Hiekkataipale, A. Laiho, V. Lemieux, J. Seitsonen, J. Ruokolainen and P. Ceci, Electrostatic assembly of binary nanoparticle superlattices using protein cages, *Nat. Nanotechnol.*, 2013, **8**, 52–56.
6 M. Liang, K. Fan, M. Zhou, D. Duan, J. Zheng, D. Yang, J. Feng and X. Yan, H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 14900–14905.
7 L. Vannucci, E. Falvo, C. M. Failla, M. Carbo, M. Fornara, R. Canese, S. Cecchetti, L. Rajsiglova, D. Stakheev, J. Krizan, A. Boffi, G. Carpinelli, V. Morea and P. Ceci, In Vivo Targeting of Cutaneous Melanoma Using an Melanoma Stimulating Hormone-Engineered Human Protein Cage with Fluorophore and Magnetic Resonance Imaging Tracers, *J. Biomed. Nanotechnol.*, 2015, **11**, 81–92.
8 E. Falvo, E. Tremante, R. Fraioli, C. Leonetti, C. Zamparelli, A. Boffi, V. Morea, P. Ceci and P. Giacomini, Antibody-drug conjugates: targeting melanoma with cisplatin encapsulated in protein-cage nanoparticles based on human ferritin, *Nanoscale*, 2013, **5**, 12278–12285.
9 L. Vannucci, E. Falvo, M. Fornara, P. Di Micco, O. Benada, J. Krizan, J. Svoboda, K. Hulikova-Capkova, V. Morea, A. Boffi and P. Ceci, Selective targeting of melanoma by PEG-masked protein-based multifunctional nanoparticles, *Int. J. Nanomed.*, 2012, **7**, 1489–1509.
10 S. Geninatti Crich, M. Cadenazzi, S. Lanzardo, L. Conti, R. Ruiu, D. Alberti, F. Cavallo, J. C. Cutrin and S. Aime, Targeting ferritin receptors for the selective delivery of imaging and therapeutic agents to breast cancer cells, *Nanoscale*, 2015, **7**, 6527–6533.
11 K. Fan, C. Cao, Y. Pan, D. Lu, D. Yang, J. Feng, L. Song, M. Liang and X. Yan, Magnetoferritin nanoparticles for targeting and visualizing tumour tissues, *Nat. Nanotechnol.*, 2012, **7**, 459–464.
12 Z. Zhen, W. Tang, T. Todd and J. Xie, Ferritins as nanoplatforms for imaging and drug delivery, *Expert Opin. Drug Delivery*, 2014, **11**, 1913–1922.
13 K. Mihee, R. Yecheol, S. J. Kyeong, A. Byungcheol, J. Sungmin, K. Heesoo and R. Moonhor, pH-dependent structures of ferritin and apoferritin in solution: disassembly and reassembly, *Biomacromolecules*, 2011, **12**, 1629–1640.
14 H. Chen, S. Zhang, C. Xu and G. Zhao, Engineering protein interfaces yields ferritin disassembly and reassembly under benign experimental conditions, *Chem. Commun.*, 2016, **52**, 7402–7405.
15 S. Kim, J. O. Jeon, E. Jun, J. Jee, H. K. Jung, B. H. Lee, I. S. Kim and S. Kim, Designing Peptide Bunches on Nanocage for Bispecific or Superaffinity Targeting, *Biomacromolecules*, 2016, **17**, 1150–1159.
16 A. M. Sevcenco, M. Paravidino, J. S. Vrouwenvelder, H. T. Wolterbeek, M. C. van Loosdrecht and W. R. Hagen, Phosphate and arsenate removal efficiency by thermostable ferritinenzyme from Pyrococcus furiosus using radioisotopes, *Water Res.*, 2015, **76**, 181–186.
17 B. Sana, E. Johnson and S. Lim, The unique self-assembly/disassembly property of *Archaeoglobus fulgidus* ferritin and its implications on molecular release from the protein cage, *Biochim. Biophys. Acta*, 2015, **1850**, 2544–2551.
18 E. Johnson, D. Cascio, M. R. Sawaya, M. Gingery and I. Schröder, Crystal structures of a tetrahedral open pore ferritin from the hyperthermophilicarchaeon *Archaeoglobus fulgidus*, *Structure*, 2005, **13**, 637–648.
19 B. Sana, E. Johnson, P. Le Magueres, A. Criswell, D. Cascio and S. Lim, The role of nonconserved residues of *Archaeoglobus fulgidus* ferritin on its unique structure and biophysical properties, *J. Biol. Chem.*, 2013, **288**, 32663–32672.
20 Z. Heger, S. Skalickova, O. Zitka, V. Adam and R. Kizek, Apoferritin applications in nanomedicine, *Nanomedicine*, 2014, **9**, 2233–2245.
21 L. Li, C. J. Fang, J. C. Ryan, E. C. Niemi, J. A. Lebrón, P. J. Björkman, H. Arase, F. M. Torti, S. V. Torti, M. C. Nakamura and W. E. Seaman, Binding and uptake of H-ferritin are mediated by human transferrin receptor-1, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 3505–3510.
22 S. Levi, A. Luzzago, G. Cesareni, A. Cozzi, F. Franceschinelli, A. Albertini and P. Arosio, Mechanism of ferritin iron uptake: activity of the H-chain and deletion mapping of the ferro-oxidase site. A study of iron uptake and ferroxidase activity of human liver, recombinant H-chain ferritins, and of two H-chain deletion mutants, *J. Biol. Chem.*, 1988, **263**, 18086–18092.
23 C. Bernacchioni, V. Ghini, C. Pozzi, F. Di Pisa, E. C. Theil and P. Turano, Loop electrostatics modulates the intersubunit interactions in ferritin, *ACS Chem. Biol.*, 2014, **9**, 2517–2525.
24 L. Calisti, I. Benni, M. Cardoso Trabuco, P. Baiocco, B. Ruzicka, A. Boffi, E. Falvo, F. Malatesta and A. Bonamore, Probing bulky ligand entry in engineered archaeal ferritins, *Biochim. Biophys. Acta*, 2016, DOI: 10.1016, [Epub ahead of print].
25 W. Kabsch, Integration, scaling, space-group assignment and post-refinement, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 133–144.
26 P. Emsley and K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126–2132.
27 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, PROCHECK - a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.*, 1993, **26**, 283–29125.
28 V. B. Chen, W. B. Arendall 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallo-

 This journal is © The Royal Society of Chemistry 2017

graphy, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 12–21.
29 C. J. Russo and L. A. Passmore, Ultrastable gold substrates for electron cryomicroscopy, *Science*, 2014, **346**, 1377–1380.
30 J. Dubochet, M. Adrian, J. J. Chang, J. C. Homo, J. Lepault, A. W. McDowall and P. Schultz, Cryo-electron microscopy of vitrified specimens, *Q. Rev. Biophys.*, 1988, **21**, 129–228.
31 T. Wagenknecht, R. Grassucci and J. Frank, Electron microscopy and computer image averaging of ice-embedded large ribosomal subunits from *Escherichia coli*, *J. Mol. Biol.*, 1988, **199**, 137–147.
32 R. Danev, B. Buijsse, M. Khoshouei, J. M. Plitzko and W. Baumeister, Volta potential phase plate for in-focus phase contrast transmission electron microscopy, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15635–15640.
33 S. H. W. Scheres, Semi-automated selection of cryo-EM particles in RELION-1.3, *J. Struct. Biol.*, 2015, **189**, 114–122.
34 S. Chen, G. McMullan, A. R. Faruqi, G. N. Murshudov, J. M. Short, S. H. Scheres and R. Henderson, High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy, *Ultramicroscopy*, 2013, **135**, 24–35.
35 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt and E. C. Meng, Ferrin TE. UCSF Chimera–a visualization system for exploratory research and analysis, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
36 M. Morlando, S. Dini Modigliani, G. Torrelli, A. Rosa, V. Di Carlo, E. Caffarelli and I. Bozzoni, FUS stimulates microRNA biogenesis by facilitating co-transcriptional Drosha recruitment, *EMBO J.*, 2012, **31**, 4502.
37 X. Robert and P. Gouet, Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res.*, 2014, **42**, 320–324.

This journal is © The Royal Society of Chemistry 2017

Appendix III

Engineered ferritin for lanthanide binding.

Lorenzo Calisti[a], Matilde Cardoso Trabuco[b], Alberto Boffi[c,d], Claudia Testi[a], Linda Celeste Montemiglio[a], Amédée des Georges[e], Irene Benni[a], Andrea Ilari[d], BartłomiejTaciak[f,g], MaciejBiałasek[f], Tomasz Rygiel[g,h], Magdalena Król[f,g], Paola Baiocco[c], Alessandra Bonamore[a].

[a.] Department of Biochemical Sciences "Alessandro Rossi Fanelli", Sapienza University of Rome, Rome Italy.

[b.] Molirom srl, via Ravenna 8, 00161, Rome, Italy.

[c.] Center for Life Nano Science@Sapienza, Istituto Italiano di Tecnologia, V.le Regina Elena 291, Rome 00161, Italy

[d.] Institute of Molecular Biology and Pathology, National Research Council, Rome, Italy.

[e.] The City University of New York Advanced Science Research Center 85 St. Nicholas Terrace, New York, NY 10031.

[f.] Faculty of Veterinary Medicine, Warsaw University of Life Sciences, ul. Nowoursynowska 166 02-787, Warszawa, Poland.

[g.] Cellis Ltd., Gen. Zajaczka 28, 01-510 Warsaw, Poland

[h] Department of Immunology, Center for Biostructure Research, Medical University of Warsaw, Banacha 1a, 02-089 Warsaw, Poland.

*Corresponding author: paola.baiocco@iit.it

Keywords: Protein cage; Ferritin; cellular uptake; Terbium binding, protein crystallography; cryo electron microscopy.

Abstract

Mammalian ferritin H-homopolymers have been extensively studied to be used as nanocarriers for diverse applications in the targeted delivery of anticancer drugs and as imaging agents. In the present work, we have engineered a new construct in which a lanthanide binding tag (LBT) was genetically fused to the C-terminal end of mouse H ferritin, HFt. The HFt-LBT thus obtained possesses an extra Terbium binding site per subunit provided by six coordinating aminoacid side chains and a tryptophan residue in

its close proximity and is thus endowed with high Terbium ion affinity and strong FRET sensitization properties. The Terbium emission band at 544 nm for the HFt-LBT Tb(III) complex was detectable in the pM protein concentration range in solution, two order of magnitude higher with respect to control wild type mouse HFt. X-ray crystallography data at 2.9Å resolution and cryo-EM at 7 Å resolution demonstrated that HFt-LBT is correctly assembled as a 24-mer both in crystal and in solution. Tb(III) binding sites were identified in the protein threefold axes and in the ferroxidase sites for a total of about 56 Tb(III) atoms per 24-mer, comprising the 24 metal ions coordinated to the LBT. Moreover, HFt-LBT Tb(III) was demonstrated to be actively uptaken by selected tumor cell lines by confocal microscopy and FACS analysis of their FITC derivatives.

1. INTRODUCTION

Ferritin is a cage-like protein made of 24 subunits arranged in octahedral 432 symmetry with an outer diameter of roughly 12 nm and an inner diameter of 8 nm [1,2]. The symmetrical positioning of three or four subunits in the protein shell results in the formation of eight channels connecting the inner cavity to the outside and allows for the entry and exit of iron and other cations with a relatively broad selectivity [3]. Besides their physiological function, centered around intracellular iron uptake, the nanocage properties of ferritins have been exploited in a number of diverse biotechnological applications as drug delivery vectors [4], scaffolds for vaccine development [5] and tools for bioimaging [6]. In this context, ferritins have been proven particularly useful for the selective targeting to cell populations overexpressing the transferrin receptor (CD71), in particular iron avid tumor cells [7,8].

Among the many ingenious ferritin based constructs for bioimaging, only a few have been devoted to the creation of smart fluorescent probes and these include quantum dots [9] and gold nanoparticles [10, 11]. However, advanced optical imaging techniques need an expanded color palette of bright fluorescent probes for biological visualization in order to enable real-time cellular imaging with high spatial resolution for close up view into subcellular compartments and for providing key information on intracellular activities and macromolecular dynamics.

In this framework, fluorescent probes based on trivalent lanthanide ions are becoming widespread due to their unique photophysical properties [12-14]. Lanthanide f-orbitals with their high quantum yields are capable of efficiently radiating most of the absorbed energy, although their small absorption cross sections poses limits to their use. To improve the luminescent signals, small organic fluorophores that absorb in the UV region and transfer the absorbed light to the lanthanide atom are thus currently used in complex with the metal ion. Complexes of lanthanides are characterized by narrowband emission

spectra, large Stokes shift (150-300 nm), and excited state lifetimes within the range of micro to milliseconds. By exploiting the microsecond fluorescence of lanthanides, time-resolved spectroscopy allows for the elimination of the short living background signals whose lifetimes are usually not more than 10-15 ns and the enhancement of the sensitivity for recording the delayed signal [15]. Moreover, the inherently low extinction coefficient of lanthanide ions, due to the forbidden character of the electronic transitions, can be overcome by Förster resonance energy transfer (FRET) from an appropriately placed (within 5-6 Å distance) sensitizer-fluorophore onto the emitting level of the lanthanide, a phenomenon often referred to as "antenna effect" [16]. Among lanthanides, Tb(III) and Eu(III) are the most interesting due to their more intense microsecond fluorescence in the visible region [17, 18]. Among the various methods for the incorporation of lanthanide ions into biomolecules, a straightforward and generalizable approach has been proposed in recent years that integrates a lanthanide binding sequence as a protein co-expression tag via molecular biology strategies. On the basis of known properties of calcium binding loops, recent design and engineering studies have resulted in the development of short polypeptides comprising 20 or fewer encoded amino that are capable of tight and selective binding to lanthanides. These peptide sequences, identified as "lanthanide-binding tags" (LBT), show low-nanomolar affinities for the target ions and are selective for lanthanides over other common metal ions [19-22]. The probe nature of these protein tags has been demonstrated by their use for luminescence-based visualization on gels, as magnetic-field paramagnetic alignment agents in protein NMR experiments [23, 24] in fluorescence microscopy, [25, 26] and as partners in luminescence resonance energy transfer (LRET) studies [27].

Among the several proteins physiologically involved in metal binding that can be used for Terbium binding, ferritins may appear among the most suitable. In fact, native apoferritins have been reported to naturally bind Tb(III) within their iron binding sites [28]. In particular, it was demonstrated that mammalian apoferritins could bind more than one Tb(III) ion per subunit, corresponding respectively to ferroxidase site, threefold channel and nucleation centre [29, 30]. Moreover, upon excitation between 280-295 nm Tb(III) ferritin complex showed characteristic emission bands at 490 and 544 nm due to a FRET sensitization effect provided by aromatic aminoacids [30]. However, the distance between Terbium ions and aromatic moieties in native ferritin isoforms made the FRET efficiency very poor and suboptimal for any type of fluorescence/luminescence based measurement.

In the present paper, a LBT sequence has been genetically fused at the C-terminal end of heavy chain of mouse ferritin. The tag has been designed to be located inside the inner cavity such that the

lanthanide ion diffusing through the surface pores can eventually bind to the LBT sequence. The construct would thus act both as carrier targeted to CD71 receptors and as a FRET sensitizer.

2. METHODS

2.1 Protein design, expression and purification.

A synthetic gene encoding for mouse H chain ferritin (HFt) fused with a lanthanide binding peptide (LBT) was designed, synthesized, and optimized for *Escherichia coli* codon usage by Geneart (Geneart AG). LBT sequence YIDTNNDGWIEGDELLA [22] was added to the C-terminal of HFt generating HFt-LBT construct, that was subcloned into pET22-b vector (Novagene) between NdeI/XhoI restriction sites.

HFt-LBTwas overexpressed in *Escherichia coli* BL21 cells upon induction with 1 mM IPTG (Isopropyl-β-D-1-thiogalactopyranoside) at OD_{600}=0.6. Cells were harvested by centrifugation 16 hours post induction at 37°C.

Cells harvested from 1 L culture were resuspended in 20 mM HEPES buffer, pH 7.5, containing 200 mM NaCl, 1 mM TCEP (tris(2-carboxiethyl)phosphine), and a cOmplete TM Mini Protease Inhibitor Cocktail Tablet (Roche). Cells were disrupted by sonication and the soluble fraction was purified by heat treatment at 78°C for 10 minutes. Denatured *E. coli* proteins were removed by centrifugation at 15.000 rpm at 4°C for 1 hour. The soluble protein was further purified by ammonium sulfate precipitation. The precipitated fraction at 70% ammonium sulfate, was resuspended in 20 mM HEPES, 50 mM $MgCl_2$, pH 7.5 and dialyzed versus the same buffer. As final purification step, the protein was loaded onto a HiLoad 26/600 Superdex 200 pg column previously equilibrated in the same buffer using an ÄKTA-Prime system (GE Healthcare). Purified protein was concentrated to obtain the final protein preparation of 1 mg/mL and protein concentration was calculated by measuring the UV spectrum using an extinction coefficient of 32400 $M^{-1}cm^{-1}$. Protein yield was ~50 mg/l culture.

The expected molecular weight of 22662 Da was confirmed by MALDI-TOF Mass Spectrometry as reported in Supplementary Data (Figure S1).

2.2 Fluorescence spectroscopy

Fluorescence spectra and titrations were performed using FluoroMax 4 (Horiba) spectrofluorimeter with a Haake D8 refrigerated bath at 20 °C. Emission spectra were recorded between 450 and 560 nm, in order to include the luminescent maxima of Tb(III) (490 and 545 nm). The excitation wavelength was chosen at 295 nm to minimize the overlap of second order diffraction (570 nm) with the Tb(III) emission at 545 nm. Emission spectra were taken with excitation and emission band passes of 4

and 8 nm and corrected for the blank contribution and the instrument response at 295 nm in a quartz cell of 1 cm pathlength. Emission spectra were normalized to 1 at 545 nm.

Fluorescence measurements were performed using 1 μM apoHFt-LBT and apoHFt as a control in 100 mM MES buffer pH 6.4. A 50 mM $TbCl_3$ (Aldrich) stock solution was also prepared in MES buffer at pH 6.4. Fluorescence spectra were recorded upon 30 min incubation with excess of $TbCl_3$ in buffer solution followed by a washing step to remove the unbound metal. Titration data were obtained by addition of aliquots of $TbCl_3$ in buffer solution under continuous stirring.

2.3 Cryo-Electron microscopy

Holey-gold grids from Quantifoil R1.2/1.3 (Quantifoil Micro Tools GmbH) were prepared as described [31]; their surfaces were treated with plasma cleaning using a mixture of H_2 and O_2 and 3 μl of a solution containing 1 μM HFt-LBT Tb(III) complex was applied to the holey-gold grids. After 30 s waiting time, grids were blotted for 3 seconds at 100% humidity with filter paper and vitrified by rapidly plunging into liquid ethane at −180 °C [32, 33] with a Vitrobot Mark IV (FEI).

Data acquisition was performed using a FEI Titan Halo (FEI, Eindhoven) operating at 300 kV, while the specimen was maintained at liquid nitrogen temperatures. Datasets were collected with an automated data collection system [34] on a K2 Summit direct detector camera (Gatan, Pleasanton) operating in super-resolution mode, with a calibrated pixel size of 1.15 Å on the object scale and a magnification of 59 000×. Images were typically recorded with a defocus range between −0.7 and −3.0 μm and a dose of electrons on the specimen plane between 10 and 20 electrons/$Å^2$.

Data analysis was completely carried out using RELION 2.0 [35]. CTF correction was done using MotionCor2 [36]. Particles were picked and extracted from the original micrographs with the reference-based automated particle picking procedure implemented in RELION [37] and were 2D classified using 100 classes. 2D good classes were selected and then subjected to 3D classification with 8 classes (see Supplementary Figures S2 and S3), using as reference model the mouse H-ferritin structure (PDB code 3WNW). Resulting good classes were refined with the 3D Autorefine procedure in RELION.

The final 3D map resolution is calculated with the FSC = 0.143 criterion, based on the 'gold-standard' protocol [38], using soft masks with a 4 pixel soft edge, and has been corrected for the effects of the mask on the FSC curve using high-resolution noise substitution [39]. This final map was visualized using UCSF Chimera [40].

2.4 Crystallization and X-ray structure determination.

Crystals of apoHFt-LBT and HFt-LBT in complex with Tb(III) were obtained by mixing in a 2 µl hanging drop the purified protein at 15 mg/ml with a solution containing 1.8/2.0 M ammonium sulphate and 0.1 M Tris, pH 8.5, at 25° C within a week, cryo-protected by extensively washing the crystals in sodium malonate and flash-frozen in liquid nitrogen. Diffraction data were collected at XRD1 beamline at the Elettra Synchrotron, Trieste, Italy.

Data were processed with XDS [41] and scaled with Aimless [42] at a final resolution of 2.85 and 2.65 Å. The structures were solved by Molecular Replacement with Phaser [43, 44] using the structure of mouse H-chain modified ferritin (pdb code 3WNW) as search model. Model building and refinement were done using Coot [45] and Refmac5, respectively. The final model was analyzed with PROCHECK [46] and Molprobity [47]. Ramachandran Plot showed more than 98 % residues were in preferred regions and no outlier was observed in both structures. The validation of metal binding sites was performed using CheckMyMetal web server [48]. The final atomic coordinates and structure factors were deposited with the PDB Data Bank (www.rcsb.org) with accession code (5OBA and 5OBB, respectively). Complete data collection and refinement statistics are reported in Supplementary Data (Table S1).

2.5 Cell cultures and ferritins internalization

Human prostate cancer cell line DU-145 (ATCC® HTB81™), human colorectal cancer cell line HCT-116 (ATCC® CCL-247™), human breast cancer cell line MDA-MB-231 (ATCC® HTB-26™) and human ovarian cancer cell line SKOV-3 (ATCC® HTB-77™) were cultured in DMEM medium containing 10% FBS, 100 µg/ml streptomycin and 100 U/ml penicillin G in a humidified 37 °C incubator. The internalization assay was performed as follow: cells were detached using trypsin, then washed twice with PBS and resuspended in non-supplemented DMEM medium containing FITC-ferritin nanoparticles (HFt-LBT or mouse HFt as a control) at the final concentration 0.5 mg/ml for 1 hour in a humidified 37 °C incubator. After incubation cells were washed three times with PBS and subjected for confocal microscopy and flow cytometry analysis.

2.6 Confocal microscopy

Following internalization step described above, cells were seeded into 8-well Nunc™ Lab-Tek™ Chambered Coverglass with 200 µl DMEM medium containing 10% FBS, 100 µg/ml streptomycin and 100 U/ml penicillin G per well. Chambers with cells were then incubated on ice until microscopic visualization. Images were acquired using an inverted confocal microscope IX70 FV 500 (Olympus), with 488 nm laser, 20x objective lens and emission filter 505-560 nm. Image processing was performed

using ImageJ software (National Institutes of Health, www.imagej.nih.gov/ij/).

2.7 Flow cytometry analysis

Cells were incubated with FITC-ferritin nanoparticles as described previously then washed three times with PBS, and resuspended in FACS buffer (2% FBS, 1 mM EDTA). Internalization of ferritins before and after treatments was measured at the BD FACS Aria™ III equipped with a 488 nm laser. Cells were first gated by forward and side scatter area (FSC-A and SSC-A) plot and for singlets population (FSC-H and SSC-A), then detected in the channel for FITC expression (530/30nm filter) and side scatter parameter. The gate for the final detection was set according to the gate set on the control sample. Data were analyzed using BD FACS DIVA™ and Flow Josoftwares.

3. RESULTS

3.1 Fluorescence spectroscopy

Static emission spectra were recorded for both Terbium saturated HFt-LBT and wild type mouse HFt upon excitation at 295 nm. As reported in Figure 1, the intensity of the Terbium emission peak at 544 nm for HFt-LBT was at least two orders of magnitude higher with respect to HFt, indicating that the presence of the LBT effectively sensitizes the Terbium signal. The Terbium emission band at 544 nm for the HFt-LBT Tb(III) complex is detectable up to 4 pM protein concentration.

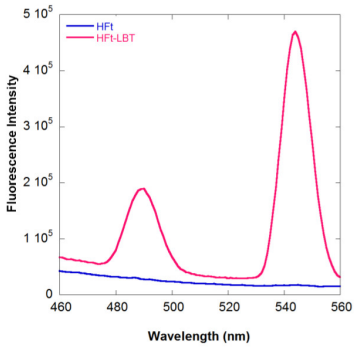


Figure 1: Fluorescence spectra of HFt-LBT Tb(III) complex.

Fluorescence spectra of HFt-LBT Tb(III) complex (red line) and wild type mouse HFtTb(III) complex (blue line) at the same protein concentration (1 μM). Spectra were recorded after 1 h incubation time with $TbCl_3$ in 0.1 M MES buffer pH 6.4.

Fluorescence titration analysis were thus carried out on HFt-LBT by adding free Tb(III) ions to the apoprotein (Figure 2 A and B).

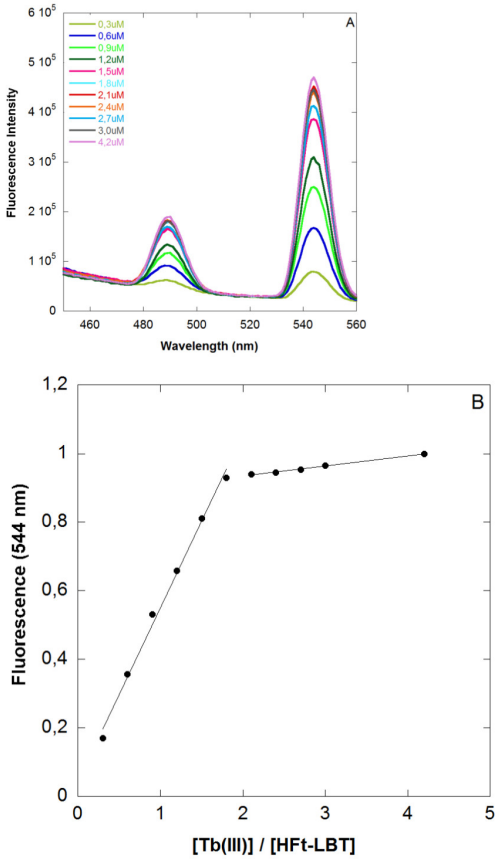


Figure 2: Fluorescence titration of HFt-LBT with Tb(III).

A) Fluorescence titration of HFt-LBT (1 µM) with incremental concentration of Tb(III) (0–4 equivalents) in 0.1 M MES buffer pH 6.4. Emission spectra were recorded in 1 cm pathlength cuvette

upon excitation at 295 nm. B) Fluorescence intensity of HFt-LBT Tb(III) complex as a function of the Tb(III)/HFt-LBT ratio. Fluorescent intensity was recorded at 545 nm and normalized to the emission maximum.

The titration endpoint was reached at 1.7 equivalent amount of Tb(III) per subunit. Extensive dialysis or size exclusion chromatography could not remove bound Tb(III). Titrations were conducted at pH 6.4 in MES buffer, in order to avoid precipitation of Terbium hydroxides, easily formed around neutrality. Analogous titrations carried out on wild type mouse HFt, demonstrated a negligible fluorescence contribution provided by Terbium bound to the metal binding sites of the native protein (*i.e.* ferroxidase site, threefold channel and nucleation centre). Thus, the presence of the tryptophan residue within the LBT, appears to account for most of the observed fluorescence signal. However, an additional contribution of the tryptophan residue as a FRET sensitizer is envisaged to enhance emission not only of the Tb(III) atom within the tag but also to the other Tb(III) atoms present in the ferroxidase site or in the threefold channels.

3.2 Cryo-EM microscopy

A plasma cleaned solution of HFt-LBT Tb(III), applied to holey-gold grids and vitrified, was used to obtain a complete 7.1 Å resolution map of the protein structure in solution (see Figure 3 and Figure S4).

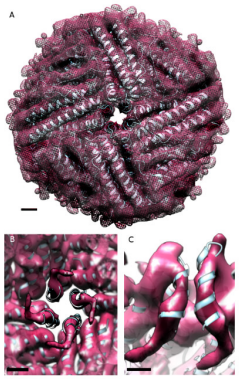


Figure 3: 3D postprocess final maps of HFt-LBT Tb(III) from Cryo EM analysis.

Data were obtained with RELION and visualized with UCSF Chimera. Light blue: X-ray diffraction data of the crystal Terbium structure. Pink: experimental Cryo-EM electronic density map.

Map resolution: 7.1 Angstroms. Scale bar=10 Angstroms. A: external view; B: internal view; C: focus on one internal C-terminal helix.

The final 3D map of HFt-LBT Tb(III) corresponds to an overall structure that conforms to the high resolution crystallographic data indicating perfect matching of the subunit assembly and helix axes. It is worth considering, from the analysis of the selected 2D classes, that significant densities appear within the internal cavity that can be attributed to the lanthanide binding tags (see Supplementary Data, Figure S2). Nevertheless, considerable heterogeneity of the observed signals does not allow even a partial reconstruction of at least a single conformation of the 24 LBTs. Only, a kinked fragment of the terminal tail can be modeled, which corresponds to a segment of 6-7 aminoacids.

3.3 X-ray crystal structure

The structures of apoHFt-LBT and HFt-LBT Tb(III) were determined by X-ray crystallography at a 2.85 Å and 2.65 Å resolution, respectively. They both crystallized in I222 space group with 24 identical subunits in the asymmetric unit (ASU) with a solvent content of 64,7%. The overall X-ray-structure confirms that the LBT does not affect the protein scaffold which corresponds to the native H-chain mouse ferritin (pdb code 3WNW) with a *rmsd* value of 0.1 Å. However, the LBT loop is not visible in both structures, due to the high flexibility of the C-terminus region. Temperature factors analysis shows a mean B factor of 57 $Å^2$ calculated for the main chain of the subunits in both structures and 78 $Å^2$ for the Terbium ions modelled in HFt-LBT Tb(III).

X-ray emission scans ranging from 4.0 to 8.0 KeV of the crystals were performed in order to confirm the presence of the Terbium in the samples. Characteristic L and M X-ray line energies of Terbium were clearly identified in the HFt-LBT Tb(III) crystals. Accordingly, the crystals of apoHFt-LBT lacked these emission energies and displayed the presence of residual iron atoms at the characteristic energies lines in the emission scan.

In HFt-LBT Tb(III) structure, 24 Terbium ions were positioned and successfully refined with a 75% occupancy in each ferroxidase site of the protein scaffold and 8 Terbium ions were positioned with 100 % occupancy in each three-fold axes. In the ferroxidase center, each Terbium ion is located in a trigonal planar coordination to OE1-Gln141 and to OE1 and OE2-Glu62, and to OE2-Glu27, in a range 2.6-3.2 Å distance (Figure 4A). In the 3-fold center, the Terbium ion is tetrahedrally coordinated to OE1-Glu134 of the three subunits, (at 2.2-2.4 Å distance) (Figure 4B). In the 4-fold channels were positioned and successfully refined one water molecule.

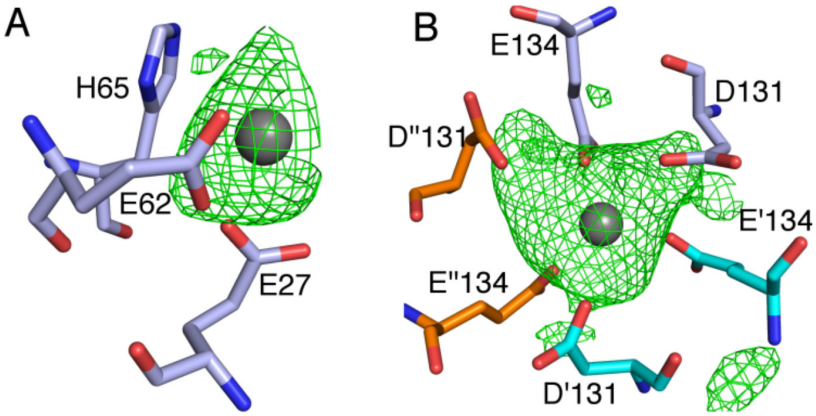


Figure 4: Terbium binding sites from X-ray crystallography.

The omit map contoured at 3 σ is shown as a green mesh A) In the ferroxidase center: a Tb(III), is shown as a grey sphere and B) in the 3-fold axes: the residues Glu131 and Glu134 of three different monomers are depicted as sticks in light blue, cyan and orange, respectively.

3.4 Ferritins uptake by tumor cell lines

After demonstration that HFt-LBT Tb(III) maintained the overall structure of the wild type mouse HFt, binding and internalization of the construct was analyzed in selected cancer cell lines. It is known that HFt is recognized and internalized by the CD71 (TfR1), which is overexpressed in many types of tumor cells but not in normal cells and healthy tissues [49]. Experiment on cells treated with the same amount (0.5 mg/ml) of HFt-LBT Tb(III) or HFt were undertaken in order to study uptake efficiency by cancer cells by flow cytometry and confocal microscopy. As a baseline for FITC fluorescence, control cells untreated with FITC-ferritins were used. In order to exclude any signal generated from outside particles sticking on the cell membrane due to unspecific binding or remaining from the washing steps, Trypan blue quenching was performed before FACS acquisition. In Supplementary Figures S5, the FACS analysis is summarized, shown as the percentage of cells internalizing the nanoparticles. These data highlighted that HFt-LBT Tb(III) nanoparticles are efficiently taken up by all cells of each cell line. All acquisition plots are shown in Figure S5.

In order to visualize HFt-LBT Tb(III) nanocages within cytoplasm, cells were observed under confocal microscopy. Confocal representative images of entire field of view of live cells incubated with HFt-LBT Tb(III) are shown in Figure 5A. Images confirmed the high extent of both HFt-LBT Tb(III) and wild type mouse HFt internalization and highlight similar cellular distribution in the cytoplasm and in the perinuclear space.

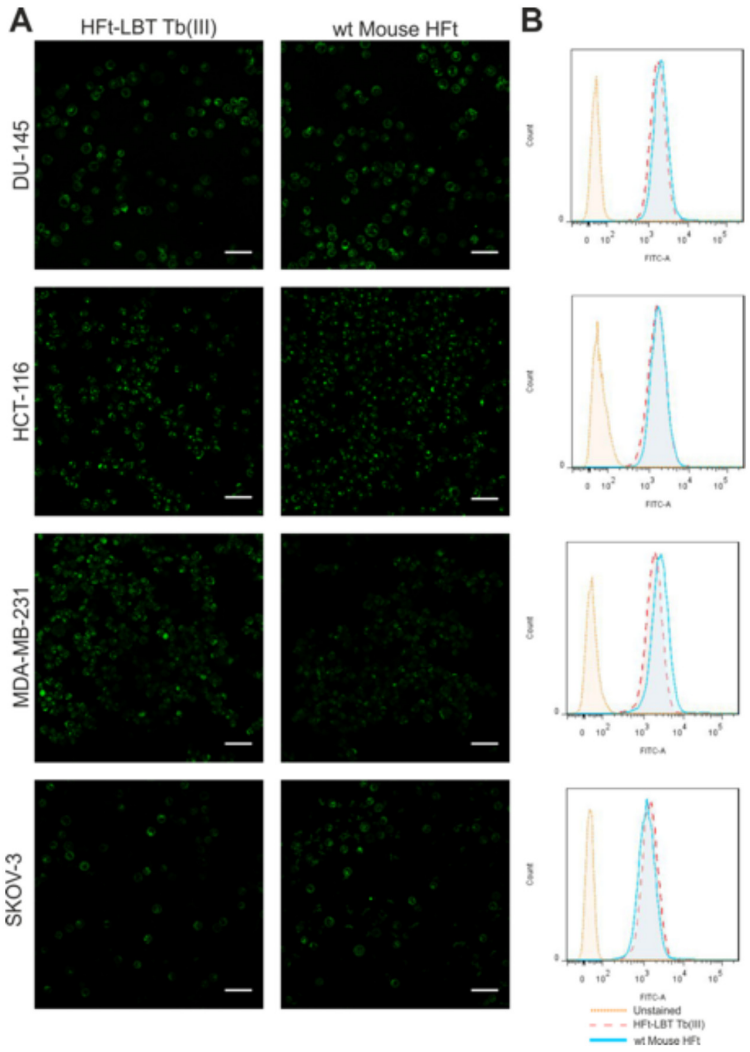

A
HFt-LBT Tb(III)
wt Mouse HFt
B
DU-145
HCT-116
MDA-MB-231
SKOV-3
Count
FITC-A
Unstained
HFt-LBT Tb(III)
wt Mouse HFt

Figure 5: Confocal microscopy images and flow cytometry analysis of HFt-LBT Tb(III) uptake by selected tumor cell lines.

DU-145, HCT-116, MDA-MB-231 and SKOV-3 cancer cells were incubated with either HFt-LBT Tb(III) or wt mouse HFt (0.5 mg/ml) for 60 min. A) Images acquired by confocal microscopy showing side by side comparison of cellular distribution of HFt-LBT Tb(III) and HFt conjugated with FITC. Scale bar = 50 μm. B) Flow cytometry analysis of HFt-LBT Tb(III) and HFt cellular uptake.

4. DISCUSSION

In the present work, an engineered ferritin construct, bearing a lanthanide binding tag on the C-terminal end of each subunit of mouse H ferritin, was designed in order to build a biomolecular nanosystem endowed with strong FRET sensitization properties. The peptide tag (LBT) was designed according to the findings of Martin *et al.*, [22] that demonstrated efficient lanthanide chelating properties of a peptide sequence derived from Ca^{2+} binding sites from Troponin C EF hand motif. The LBT has six metal-binding residues that form a coordination sphere around a lanthanide (III) ion and is endowed with high affinity and strong FRET effect from the encoded tryptophan residue to the Terbium or Europium ions. Spectroscopic and crystallographic studies showed that the inner coordination sphere of Tb(III) bound to this sequence was free from water molecules [21], a key feature for luminescence experiments, given the quenching effect of O-H vibration [50]. Accordingly, the HFt-LBT construct exhibited high affinity Tb(III) binding as demonstrated by fluorescence spectroscopy measurements in solution in comparison with the native protein. In these measurements, excitation of the tryptophan residue present in the center of the LBT provided an excellent FRET to the Terbium atom with a sustained narrow emission at the 544 nm line typical of Terbium excited state decay (see Figure 1). The thermodynamics of the Terbium binding process was however complex due to the well known presence of additional metal binding sites to the ferritin 24-mer [28-30]. In particular, the observed fluorescence signal accounted for 1.7 Tb(III) ions per subunit indicating that the presence of the tryptophan residue within the tag is able to act as an antenna system not only for the Tb(III) ion bound within the LBT tag but also for a number of extra Tb(III) atoms bound to ferritin binding site. Spurious effects due to other contributions are ruled out by control experiments on wild type mouse HFt, still able to bind Tb(III) within the canonical metal binding sites but endowed with negligible fluorescence in the Tb(III) emission regions.

Analysis of the three-dimensional structure of the HFt-LBT Tb(III) complex, by X-ray crystallography and cryo-EM demonstrates that the presence of the C-terminal tag does not affect the overall assembly of the protein and that the genetically fused tags point to the interior cavity. The peptide arm connecting the lanthanide binding loop to the C-terminal sequence is, as expected, flexible and does not allow for a complete resolution of the local structure [51]. Interestingly, cryo-EM reconstruction provided the identification of low resolution but definite patterns relative to the first few aminoacids of the tag, though the possible multiple orientations of the loop region precluded the observation of the Tb(III) complex. X-ray data further demonstrated that Terbium ion is efficiently complexed at the threefold axis channels (8 channels per 24-mer) by side chains of E131 and D134 aminoacids from each of the three adjacent subunits. Further Tb(III) binding occurs stoichiometrically at the ferroxidase site (24 sites per 24-mer), coordinated by carboxyl residues E27 and E62 with minor participation of H65. In contrast, LBT bound Tb(III) atoms were not visible by X-ray diffraction. In summary, HFt-LBT construct is capable of high affinity binding of 24 Tb(III) atoms at each tag and 20, out of 24 available (occupancy in X-ray structure corresponded to 75%), at the ferroxidase binding site. In addition, 8 fully occupied sites are observed at the entrance of the threefold channels for a total of 56 Terbium atoms within the whole 24-mer.

In the present investigation on confocal microscope, however, attempts to image live cells by direct excitation at 290-375 nm of the tryptophan residue in the lanthanide binding loop after ferritin uptake yielded very poor results due to the very high fluorescence background within the typical Terbium emission interval. We thus report measurements in which HFt-LBT Tb(III) or wild type mouse HFt have been labeled with common FITC. Results demonstrate that the HFt-LBT Tb(III) complex is very efficiently uptaken by all four human tumor cell lines selected. In particular, DU145 (from a central nervous system metastasis, of primary prostate adenocarcinoma origin), MDA-MB-231 (from invasive ductal carcinoma, Hepatoma cell line and SK-OV-3 (highly resistant ovarian cancer cell line) were demonstrated for the first time to display a high uptake of both wild type mouse HFt and HFt-LBT Tb(III). These cancer cell lines overexpress CD71 molecule and are subject of cancer therapy studies focused on this receptor (PMID: 20180585; PMID: 26491294; PMID: 23267137; PMID: 23752192).

Conclusions

In conclusion, we have shown that a ferritin nanocage can be engineered by addition of appropriate metal binding tags inside the cavity in order to provide specific additional metal sites in topologically selected positions. This approach is in alternative to the quasi random metal cluster

insertion into the ferritin cavity that has been commonly used by free diffusion of metal ions through the open pores on the surface of the macromolecule or by disassembly-reassembly of the 24-mer structure ("encapsulation" procedure) [8]. The importance of a guided allocation of metal sites inside the cavity is however essential to proceed with rationale positioning of antenna systems vs. lanthanide sites in order to assess the best geometry for efficient FRET. Rational design of such metal binding sites would foster most advanced applications such as the construction of up-converting nanoparticles or ultrabright fluorescent organic polymer for single molecule detection.

Acknowledgment

The authors are grateful to the ELETTRA XRD-1 beamline staff for their help and assistance during the data collection and fluorescence emission experiments. This study has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement nº 312284. The biological part of the study was supported by the grant from National Science Centre (NCN) in Poland no. UMO-2015/18/E/NZ6/00642. EU H2020 Project "X-Probe", Grant N° 637295, to A. Bonamore and M.C.T. is gratefully acknowledged. Flagship project "Nanomax" from MIUR to A. Boffi is also acknowledged.

References

[1] Arosio P, Elia L and Poli M 2017 Ferritin, cellular iron storage and regulation *IUBMB Life*. **69**(6) 414-422

[2] Theil EC 2013 Ferritin: the protein nanocage and iron biomineral in health and in disease *Inorg. Chem.* **52**(21) 12223-12233

[3] Calisti L, Benni I, Cardoso Trabuco M, Baiocco P, Ruzicka B, Boffi A, Falvo E, Malatesta F and Bonamore A2017 Probing bulky ligand entry in engineered archaeal ferritins *Biochim Biophys Acta.***1861**(2)450-456

[4] Zhen Z, Tang W, Chen H, Lin X, Todd T, Wang G, CowgerT, Chen X and Xie J 2013RGD-modified apoferritin nanoparticles for efficient drug delivery to tumors *ACS Nano* **7** 4830–4837

[5] Kanekiyo M, Wei C-J, Yassine HM, McTamney PM, Boyington JC, Whittle JR, Rao SS, Kong WP, Wang L and Nabel GJ 2013 Self-assembling influenza nanoparticle vaccines elicit broadly neutralizing H1N1 antibodies *Nature* **499** 102–106

[6] Valero E, Fiorini S, Tambalo S, Busquier H, Callejas-Fernández J, Marzola P, Gálvez N and Domínguez-Vera JM 2014 In vivo long-term magnetic resonance imaging activity of ferritin-based magnetic nanoparticles versus a standard contrast agent*J. Med. Chem.***57**(13)5686-5992

[7] Fracasso G, Falvo E, Colotti G, Fazi F, Ingegnere T, Amalfitano A, Doglietto GB, Alfieri S, Boffi A, Morea V, Conti G, Tremante E, Giacomini P, Arcovito A and Ceci P 2016 Selective delivery of doxorubicin by novel stimuli-sensitivenano-ferritins overcomes tumor refractoriness *J. Control Release* **239** 10-18

[8] Falvo E, Tremante E, Arcovito A, Papi M, Elad N, Boffi A, Morea V, Conti G, Toffoli G, Fracasso G, Giacomini P and Ceci P 2016 Improved Doxorubicin Encapsulation and Pharmacokinetics of Ferritin-Fusion Protein Nanocarriers Bearing Proline, Serine, and Alanine Elements *Biomacromolecules* **17**(2) 514-522

[9] Bain J and Staniland SS 2015 Bioinspired nanoreactors for the biomineralisation of metallic-based nanoparticles for nanomedicine *Phys. Chem. Chem. Phys.***17**(24)15508-15521

[10] Tian L, Dai Z, Liu X, Song B, Ye Z and Yuan J 2015Ratiometric Time-Gated Luminescence Probe for Nitric Oxide Based on an Apoferritin-Assembled Lanthanide Complex-Rhodamine Luminescence Resonance Energy Transfer System *Anal. Chem.* **87**(21) 10878-10885

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

[12] Handl HL and Gillies R 2005 Lanthanide-based luminescent assays for ligand-receptor interactions *J. Life Sci.***77** 361-371

[13] Hemmila I and Laitala VJ 2005 Progress in lanthanides as luminescent probes *Fluoresc.* **15** 529-542

[14] Bunzli J 2006 Benefiting from the unique properties of lanthanide ions *Acc. Chem. Res.* **39**, 53-61

[15] Rajapakse HE, Reddy DR, Mohandessi S, Butlin NG and Miller LW 2009 Luminescent terbium protein labels for time-resolved microscopy and screening *Angew. Chem. Int.* **48**, 4990-4992

[16] Gudgin Dickson EF, Pollak A and Diamandis EP 1995 Time-resolved detection of lanthanide luminescence for ultrasensitive bioanalytical assays *Photochem. Photobiol. B***27**(1)3-19

[17] Selvin PR 2002 Principles and biophysical applications of lanthanide-based probes *Annu. Rev. Biophys. Biomembr.* **31** 275-302

[18] Bunzli JC (2004) Luminescent lanthanide probes as diagnostic and therapeutic tools *Met. Ions. Biol. Syst.* **42** 39-75

[19] Franz KJ, Nitz M and Imperiali B 2003 Lanthanide-binding tags as versatile protein coexpression probes *Chembiochem.* **4**(4) 265-271

[20] Nitz M, Franz KJ, Maglathlin RL and Imperiali B 2003 A powerful combinatorial screen to identify high-affinity terbium(III)-binding peptides *Chembiochem.* **4**(4) 272-276

[21] NitzM, Sherawat, M, Franz KJ, Peisach E and Allen KN 2004 Structural origin of the high affinity of a chemically evolved lanthanide-binding peptide *Angew. Chem. Int. Ed.Engl.* **43**(28) 3682-3685

[22] Martin LJ, Hähnke MJ, Nitz M, Wöhnert J, Nicholas R, Silvaggi, Allen KR, Schwalbe H, Imperiali B 2007 Double-Lanthanide-Binding Tags: Design, Photophysical Properties, and NMR Applications *J. Am. Chem. Soc*. **129** (22) 7106–7113

[23] Wo¨hnert J, Franz KJ, Nitz MI, mperiali B and Schwalbe H 2003 Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings *J. Am.Chem. Soc.***125** 13338-13339

[24] Su XC, Huber T, Dixon NE and Otting G 2006 Site-specific labelling of proteins with a rigid lanthanide-binding tag *Chembiochem.* **7**1 599-1604

[25] Goda N, Tenno T, InomataK, Iwaya N and Sasaki Y 2007LBT/PTD dual tagged vector for purification, cellular protein delivery and visualization in living cells *Biochim. Biophys. Acta* **1773**(2) 141-614

[26] Shirakawa M and Hiroaki H 2007 LBT/PTD dual tagged vector for purification, cellular protein delivery and visualization in living cells *Biochim. Biophys. Acta* **1773**, 141-146

[27] Sculimbrene BR and Imperiali B 2006 Lanthanide-binding tags as luminescent probes for studying protein interactions *J. Am. Chem. Soc.* **128** 7346-7352

[28] Stefanini S, Chiancone E, Antonini E and Finazzi-Agro A 1983 Binding of Terbium to apoferritin: a fluorescence study *Arch. Biochem. Biophys.* **222**(2) 430-434

[29] Treffry A and Harrison PM 1984 Spectroscopic studies on the binding of iron, terbium, and zinc by apoferritin *J. Inorg. Biochem.***21**(1) 9-20

[30] Bou-Abdallah F, Arosio P, Levi S, Janus-Chandler C and Chasteen ND 2003 Defining metal ion inhibitor interactions with recombinant human H- and L-chain ferritins and site-directed variants: an isothermal titration calorimetry study *J. Biol. Inorg. Chem.* **8**(4) 489-497

[31] Russo C J and Passmore LA 2014Ultrastable gold substrates for electron cryomicroscopy *Science* **346** 1377–1380

[32] Dubochet J, Adrian M, Chang JJ, Homo JC, Lepault J, McDowall AW and Schultz P 1998 Cryo-electron microscopy of vitrified specimens *Q. Rev. Biophys.***21** 129–228

[33] Wagenknecht T, Grassucci R and Frank J. 1988 Electron microscopy and computer image averaging of ice-embedded large ribosomal subunits from Escherichia coli *J. Mol. Biol.* **199** 137–147

[34] Suloway C, Pulokas J, Fellmann D, Cheng A, Guerra F, Quispe J, Stagg S, Potter CS and Carragher B 2005 Automated molecular microscopy: the new Leginon system *Struct. Biol.***151** 41–60

[35] Scheres SH 2012 "RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination *J. Struct. Biol.* **180**(3) 519-530

[36] Zheng SQ, Palovcak E, Armache JP, Cheng Y and Agard DA 2016 Anisotropic Correction of Beam-induced Motion for Improved Single-particle Electron Cryo-microscopy *Nature Methods* **14** 331–332

[37] Scheres SHW 2015 Semi-automated selection of cryo-EM particles in RELION-1.3 *J. Struct. Biol.* **189** 114–122

[38] Rosenthal PB and Henderson R 2003 Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy *J. Mol. Biol.***333** 721–745

[39] Chen S, McMullan G, Faruqi AR, Murshudov GN, Short JM, Scheres SH and Henderson R 2013 High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryo-microscopy *Ultramicroscopy* **135** 24–35

[40] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM and Meng EC 2004 UCSF Chimera--a visualization system for exploratory research and analysis *J. Comput. Chem.***25**1605–1612

[41] Kabsch W 2010 Integration, scaling, space-group assignment and post-refinement *Acta Crystallogr. D Biol. Crystallogr.* **66**, 133-44

[42] Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A and Wilson KS 2011 Overview of the CCP4 suite and current developments *Acta Crystallogr. D Biol. Crystallogr.* **67**(4) 235-242

[43] Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC and Zwart PH 2010 PHENIX: a comprehensive Python-based system for macromolecular structure solution *Acta Cryst. D* **66** 213-221

[44] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC and Read RJ 2007 Phaser crystallographic software *J. Appl. Crystallogr.*: **40** 658-674

[45] Emsley P and Cowtan K 2004 Coot: model-building tools for molecular graphics *Acta Crystallogr. D Biol Crystallogr.* **60** 2126–2132

[46] Laskowski RA, MacArthur MW, Moss DS and Thornton JM 1993 PROCHECK - a program to check the stereochemical quality of protein structures J. *App. Cryst.***26** 283-29125

[47] Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS and Richardson DC 2010 MolProbity: all-atom structure validation for macromolecular crystallography *Acta Crystallogr. D Biol.Crystallogr.***66**(1)12-21

[48] Zheng H, Cooper DR, Porebski PJ, Shabalin IG, Handing KB and Minor W 2017 Check My Metal: a macromolecular metal-binding validation tool *Acta Crystallogr. D Struct. Biol.***73**(3) 223-233

[49] Habashy HO, Powe DG, Staka CM, Rakha EA, Ball G, Green AR, Aleskandarany M, Paish EC, Douglas Macmillan R, Nicholson RI, Ellis IO, Gee JM. Transferrin receptor (CD71) is a marker of poor prognosis in breast cancer and can predict response to tamoxifen. *Breast Cancer Res Treat*. 2010 Jan;119(2) 283-93

[50] Horrocks WD Jr and Sudnick DR 1979 Lanthanide ion probes of structure in biology. Laser-induced luminescence decay constants provide a direct measure of the number of metal-coordinated water molecules *J. Am. Chem. Soc.***101** 334-340

[51] Barthelmes D, Granz M, Barthelmes K, Allen KN, Imperiali B, Prisner T and Schwalbe H 2015 Encoded loop-lathanide-binding tags for log-range distance measurements in proteins by NMR and EPR spectroscopy *J. Biomol. NMR* **63** 275-282