Engineering apoferritin for targeted delivery of peptides or proteins for the treatment of breast cancer

By

Lenny Ferreira, B.Sc, M.Sc

A thesis submitted to the University of Nottingham for
the degree of Doctor of Philosophy

May 2022

Acknowledgements

Table of Content

2

Abbreviations

5-FAM	5-Carboxyfluorescein maleimide
AA	Amino acid
ACN	Acetonitrile
Afb	Affibody, $Z_{HER2:342}$ HER2 specific
*af*Ft	*Archaeoglobus fulgidus* ferritin
Apaf-1	Apoptosis-protease activating factor 1
apoFt	Hollow ferritin
AR	Androgen receptors
ATHase	Artificial transfer hydrogenase
BRCA	Breast Cancer gene
CPMV	Cowpea mosaic virus
Cy5.5	Cyanine-5.5
Cyt C	Cytochrome C
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DOX	Doxorubicin
EDTA	Ethylenediaminetetraacetic acid
ER+	Oestrogen receptor-positive
FBS	Foetal bovine serum
Ft	Ferritin
FACS	Fluorescence-activated cell sorting
GFP	Green fluorescent protein
H/L	Heavy/Light chain subunit of apoferritin
HB	Hypocrellin B
HBPN	Hydrophilic hyperbranched polyhydroxyl polymer nanoparticle
HER2	Human epidermal growth factor receptor 2
HER2+	Human epidermal growth factor receptor 2 positive
-HIS tag	Apoferritin without histidine-tag
Hsp	Heat shock protein
hrH	Horse heavy chain
huH 222	Human heavy chain apoferritin (with no cysteines) with mutations E61A, E62K, E64A, H65G and E67

huH	Human heavy chain apoferritin
huHC-	Human heavy chain apoferritin with no cysteines
huHC- K68C	Human heavy chain apoferritin (with no cysteines) with mutation K68C
huL	Human light chain apoferritin
IHC	Immunohistochemistry
IL4rP	Tumour-targeting IL4 receptor-binding peptide
IPTG	Isopropyl β- d-1-thiogalactopyranoside
IRPs	Iron regulatory proteins
LCNC	Long-circulating ferritin nanocages
MEM	Minimal essential medium
mjHsp	Methanococcus jannaschii small heat shock protein
NCs	Nanocarriers
NPs	Nanoparticles
PBS	0.1 M Phosphate-buffered saline, pH 7.5
PR+	Progesterone receptor-positive
PROM	Prominin-1 targeting peptide
pyFt	Pyrococcus furiosus ferritin
R5	Silaffin
RAD50	DNA repair protein RAD50
RNA	Ribonucleic acid
RPMI	Roswell park memorial institute medium
ROS	Reactive oxygen species
TFA	Trifluoroacetic acid
TfR-	Human heavy chain apoferritin (with no cysteines and ferroxidase mutation) with possible transferrin receptor binding mutations I70A, F71L and L72F
TfR1	Transferrin receptor 1
TRAIL	Tumour necrosis factor (TNF)-related apoptosis-inducing Ligand
VLPs	Virus-like proteins
WB	Western blot
$ZnF_{16}Pc$	Zinc hexadecafluorophthalocyanine
Δ160	Human heavy chain apoferritin with a stop codon at amino acid position 160
ΔDE	Human heavy chain apoferritin without HIS-tag and AVI-tag and with a stop codon at amino acid position 160

Amino acids

Ala (A)	Alanine
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (D)	Aspartic acid
Cys (C)	Cysteine
Glu (E)	Glutamic acid
Gln (Q)	Glutamine
Gly (G)	Glycine
His (H)	Histidine
Ile (I)	Isoleucine
Leu (L)	Leucine
Lys (K)	Lysine
Met (M)	Methionine
Phe (F)	Phenylalanine
Pro (P)	Proline
Ser (S)	Serine
Thr (T)	Threonine
Tyr (Y)	Tryptophan
Val (V)	Valine

Abstract

Ferritin is a universal intracellular protein that stores iron and accounts for the majority of iron storage in the body. Ferritin is produced by almost all living organisms, including archaea, bacteria, algae, higher plants, and animals. Human ferritin is comprised of 24 subunits, which can self-assemble into a protein cage above pH 4.5. Human apoferritin, which is ferritin void of its iron core, has many characteristics that make it favourable as a drug delivery system: its biocompatibility: its biodegradability: and its capacity for spatially defined multivalent ligand display. The major focus of this PhD study is targeted delivery of human heavy chain apoferritin to breast cancer cells. The study initially focused on removing the ferroxidase ability of human heavy chain apoferritin (human heavy chain apoferritin 222). This is because highly reactive hydroxy radicals are produced during ferroxidation. Given that there is a risk that hydroxy radicals could affect encapsulated payloads, it was essential that this function be removed. Following removal of the ferroxidase residues, circular dichroism and native polyacrylamide gel electrophoresis showed no major changes to both the secondary and tertiary structure compared to its parent human heavy chain apoferritin. Kinetic studies of human heavy chain apoferritin 222 showed a slower iron core formation time than its parent protein. Following this, potential residues on human heavy chain apoferritin that could be involved in transferrin receptor 1 receptor binding were identified. Human heavy chain apoferritin has the innate ability to target transferrin receptor 1, a membrane protein expressed on most normal cells and overexpressed in many cancers, including breast cancers. However, this innate ability can cause problems when human heavy chain apoferritin is primed with other ligands, e.g., affibodies, for targeting alternative biomarkers such as the human epidermal growth receptor 2. One of the possible issues when attaching new targeting ligands is where secondary non-specific uptake by transferrin receptor 1 is undesirable. Therefore potential residues along the BC loop that could be involved in transferrin receptor 1 binding were removed (giving the TfR- mutant). A mutant cysteine residue was also introduced in the interior of human heavy chain apoferritin for conjugation of fluorescent probes. Maleimide conjugation of fluorescent probes was achieved using 8 M urea to disassemble the apoferritin cage and make the cysteine accessible. The tagged apoferritin was then evaluated by confocal microscopy, and results indicated that TfR- was not taken up by SKBR3, a high transferrin receptor 1 expressing breast cancer cell line. Reduced uptake by TfR- confirmed that mutations along the BC loop (residues 80-82) considerably reduce uptake of human heavy chain apoferritin via transferrin

receptor 1 receptor. This finding was quantitatively confirmed using flow cytometry. To evaluate the effect of a new targeting ligand on human heavy chain apoferritin on transferrin receptor 1 binding, human epidermal growth factor receptor 2 targeting ligand (affibody) was engineered onto the N-terminus of the apoferritin subunits. Native polyacrylamide gel electrophoresis analysis confirmed that the targeting ligand did not affect the tertiary structure or assembly of human heavy chain apoferritin. Homology modelling studies of affibody-human heavy chain apoferritin indicated only partial obstruction of the transferrin receptor 1 receptor binding site by the affibody ligand. Flow cytometry studies using breast cancer cell lines BT474 (HER2+), SKBR3 (HER2+), MDA-MB-231 (HER2 normal) and MDA-MB-468 (HER2-) showed that the addition of affibodies onto human heavy chain apoferritin does not obstruct transferrin receptor 1 binding completely.

The second part of this PhD study focused on delivering therapeutic proteins using human heavy chain apoferritin. Human heavy chain apoferritin's ability to encapsulate various anti-cancer drugs stems from its characteristic 8 nm diameter hollow core and self-assembly properties. *In vitro*, apoferritin has been shown to disassemble at very acidic pH (2-3) or in the presence of 8 M urea and reassembles above pH 4.0 or after removing the denaturant. At the time this research was conducted, there were no reports describing protein encapsulation in human heavy chain apoferritin. Therefore, initial studies evaluated human heavy chain apoferritin's capacity to encapsulate therapeutic proteins such as bovine heart cytochrome C. An encapsulation value of ~3 cytochrome C molecules per apoferritin cage was calculated using the urea- and pH-based encapsulation methods. The next step involved engineering human heavy chain apoferritin to disassemble at moderate pH (4-7). The last 23 amino acids of human heavy chain apoferritin were removed (giving the ΔDE mutant), and native polyacrylamide gel electrophoresis and dynamic light scattering confirmed ΔDE had a disassembly point at pH 4. Computational modelling using PyMol showed a drastic change in electrostatic charge and channel diameter around the four-fold channel of ΔDE compared to human heavy chain apoferritin. Encapsulation studies using a pH-based method confirmed that ΔDE can successfully encapsulate cytochrome C at pH 4. MTT assays also showed that delivery of ΔDE encapsulated cytochrome C at 406 ± 42 nM reduced SKBR3 cell growth by 50%; in contrast, MRC-5 cell proliferation was not significantly affected, suggesting that differential cellular targeting is possible

1.0 Research overview

In the field of breast cancer treatment, the human body and cancer mass pose many obstacles that hinder the delivery and efficacy of chemotherapeutics, including renal filtration of compounds and accumulation of cancer-promoting mutations. As such, there is a need to develop delivery vehicles that will not only aid the delivery of a chemotherapeutic but also increase its pharmacokinetic and pharmacodynamic outcome. The delivery platform investigated in the current PhD project is human heavy chain apoferritin (ApoFt), a hollow 12 nm protein cage. ApoFt's unique characteristics, including its ability to be internalised by cells and biocompatibility, will be expanded upon in section 1.2. These characteristics have allowed it to be selected over other platforms, such as a synthetic polymeric-based delivery platform. However, even though apoFt possesses beneficial properties for drug delivery, the growing complexity and difficulty of targeting cancers such as breast cancer has meant that there is a need to modify and functionalise apoFt further to improve its targeting ability and efficacy. Therefore, it is essential to understand the unique characteristics of similar protein delivery systems, including small heat shock proteins, as it will aid apoferritin development. The unique features of other protein-based delivery systems will be expanded upon in section 1.2

Once apoFt's structural nature has been understood and its characteristics adapted for the intended purpose, apoFt can be applicable for various uses, including photodynamic therapy and vaccine development, as outlined in section 1.3. In the field of drug delivery, apoFt has been used as a vehicle to deliver therapeutic compounds, with some treatments requiring new targeting ligands on apoFt's surface. In this regard, it is important to understand and study different therapeutic fields as it will give insight into how the structural nature of the targeting ligand affects apoFt's ability to interact with its endogenous receptor. The nature and purpose of these targeting ligands will be expanded upon in section 1.5. This knowledge will be very important for the current work, which involves studying the size and distribution of a human epidermal growth factor receptor 2 (HER2) targeting ligand on human heavy chain apoFt and how this affects cellular internalisation.

In addition to using apoFt as a targeting vehicle, apoFt's 8 nm core, hydrophilic channels, and pH disassembly (pH 2)/reassembly (pH 7.5) properties have also paved its way as a delivery platform for therapeutic compounds, including chemotherapeutic and photodynamic agents. However, as with any delivery system, apoFt's unique encapsulating properties have limited the range of compounds it can encapsulate, as outlined in section 1.3. To overcome this limitation, the current work will focus on creating a new apoFt variant that could encapsulate

compounds or proteins under less acidic pH disassembly conditions. However, given a protein's large size and complex nature, there is a lack of experimental data on apoFt-based peptides/proteins' encapsulation, as outlined in section 1.4. Nevertheless, as therapeutic peptides/proteins are likely to have fewer non-specific target effects than compounds, it is important to expand apoFt's ability to encapsulate therapeutic peptides/proteins such as cytochrome C.

1.1 Breast cancers

Breast cancer is the second most common cancer worldwide and the leading cause of cancer deaths in women aged 20 to 59 years.[1] The incidence in most countries is projected to rise due to the increasing number of breast cancer-associated risk factors, including late age of first pregnancy, shorter or no periods of breastfeeding and late menopause. Other contributing factors include obesity, smoking, alcohol consumption, lack of physical activity and postmenopausal hormone therapy with oestrogen and progesterone.[2] Breast cancers, like most cancers, are heterogenic and have been classified into four major subtypes: luminal subtype A and B, basal-like and HER2-positive.[3] Each subtype is characterised by various clinical and pathological factors, including hormone receptor status, as seen in Table 1.1. Routinely, in clinical practice, three biomarkers are used to assess these cancer subtypes: the oestrogen receptor (ER), the progesterone receptor (PR) and HER2: ErbB2. However, the genetic and molecular profile changes as cancers progress and evolve, leading to further tumour heterogeneity and different biological characteristics.[4] Therefore, these markers are an approximation, and levels of other markers, such as the androgen receptors (AR) need to be considered to obtain an accurate profile of the cancer subtype.[5] The role of the AR is significant, given that in recent years HER2's role as the primary indicator for breast cancer progression has come under scrutiny following many clinical trials. Analysis of the N9831 adjuvant trastuzumab (Herceptin) trial (a monoclonal antibody used to treat HER2+ breast cancer, which will be discussed in section 1.4) noted that although treatment of HER2+ breast cancer patients with trastuzumab led to reduced cancer recurrence, there was not a linear dose-response between HER2 gene copy/protein levels and trastuzumab dose.[6, 7, 8] The outcome of the trial indicated that increased HER2+ expression in breast cancers did not necessarily correlate with downstream gene transcription levels.[8] One reason for this outcome could be that cross-talk between different receptor pathways, including the AR, could also stimulate the cells. Although its exact role is unclear, some evidence has shown that the AR could play a more

complementary role with HER2 in progressing breast cancer, and dual targeting of both these receptors might lead to better patient survival post chemotherapy.[10, 11]

	Triple negative ER-, PR-, HER2-	HER2+	Luminal B	Luminal A
% of breast cancers	15-20	10-15	20	40
Receptor expression	HER2			ER+/PR+
Prognosis	Poor			Good

Table 1. 1 Characteristics of breast cancer subtypes. ER+ (Oestrogen receptor positive), PR+ (Progesterone receptor positive) and HER2+ (Human epidermal growth factor receptor 2 positive). Modified from Breast cancer, McMaster Pathophysiology Review.[264]

1.1.1 Delivery of chemotherapeutics to breast cancer cells and overcoming resistance

Standard breast cancer therapy primarily includes surgery with radiotherapy and chemotherapy. Chemotherapy provides the additional benefit of being non-invasive and less stressful on neighbouring healthy cells than radiotherapy and surgery, increasing the number of chemotherapeutic-based clinical trials. However, many potential drug candidates fail to demonstrate efficacy in clinical trials. Numerous factors play a part in this, including the complexity of the human body, which is not replicable in a laboratory setting; the heterogeneity of cancer cells leading to drug resistance; the reduced specificity against cancer cells and poor drug biodistribution. To circumvent drug resistance, some trials have tested the effects of dosing multiple cytotoxic chemotherapeutics.[11, 12, 13] This type of treatment has the added benefit of synergistically targeting multiple molecular and cellular functions (pertinent to tumourigenesis), thereby increasing cell death. However, the delivery of chemotherapeutic drugs has many pharmacokinetic and pharmacodynamic limitations. These include poor aqueous stability and toxicity to healthy cells leading to several adverse side effects, including nausea, loss of hair and appetite through the death of healthy rapidly-dividing cells. To overcome these challenges, nanotechnology has been utilised for the delivery of chemotherapeutics.

1.1.2 Drug delivery systems

Nanotechnology in the field of drug delivery involves the use of nano-sized particles (NP), which are 1-500 nm in one dimension. When used as a platform to deliver therapeutic and diagnostic compounds, NPs can be referred to as nanocarriers (NC).[14] NCs have at least one dimension in the size range between 1 and 500 nm, and they typically offer a large surface area to volume ratio. Promising outcomes during *in vivo* studies and clinical trials have led to a large pool of NCs being developed for cancer therapy. These include liposomes, polymeric NCs, silica NCs and protein nanocages. One of the key advantages of using NCs as drug delivery systems over chemotherapeutics-only treatments is their innate ability to passively accumulate in tumour tissue, thereby reducing off-target effects, e.g., targeting non-cancerous healthy cells. This effect is aided by the immature porous nature of vasculature and the lack of lymphatic drainage surrounding tumour cells, an effect known as the enhanced permeability and retention effect.[15] As such, a smaller dose can be administered, whilst achieving a higher therapeutic index, given that more of the NC-entrapped compound would accumulate around cancer cells vs healthy cells.[16, 17, 18] Although the nano dimensions of NCs have paved their way for drug delivery, there are limitations to a NC's size (5-100 nm), particularly during *in vivo* experimentation. These limitations prevent fast renal filtration (vehicles <5 nm) and rapid elimination by the spleen (vehicles >100 nm).[19] In addition to the size, the composition and shape of NCs can also be fine-tuned to exert additional pharmaceutically beneficial properties such as sustained drug release, increased systematic drug half-life, targeted delivery, multi-drug delivery and increased drug efficacy. These added benefits have led to some NCs advancing into clinical trials. Examples include the use of liposomes for the intra-tumoural delivery of a putative tumour suppressor gene *FUS1* to lung cancers and the use of a polymeric micelle NCs for the delivery of paclitaxel to ovarian cancers.[20, 21] However, different types of NCs exert different benefits over their counterparts such as protein-based NCs and polymeric NCs, and therefore it is important to know the beneficial characteristics of the system of choice, but also the drawbacks different NCs might have both *in vitro* and *in vivo.*

1.1.2.1 Polymeric vs protein-based drug delivery systems

One type of NC that has gained interest over the years is polymeric-based NCs. Polymeric-

based NCs are colloidal systems in which therapeutic agents can be dissolved, encapsulated, or entrapped within a polymer matrix.[22] Depending on the method used to form the polymeric NCs, they can either form nanocapsules or a porous nanosphere. Nanocapsules are systems that have an aqueous/oil core in which a polymer matrix surrounds the therapeutic cage.[23] In comparison, nanospheres are continuous polymeric matrices in which the therapeutic agent is dispersed.[23] Polymeric NCs can also be produced from different nanomaterials, including chitosan, polylactide, poly(D, L-lactide) and co-polymer poly(lactide-co-glycolide). [24, 25, 26] As polymeric NCs can be formed from several different materials, they have the ability to deliver water-soluble and lipid-soluble drugs, as well as to be fine-tuned to possess different charge, size, and shape properties.[27] In addition, polymeric NCs can be modified to release drugs in a sustained manner over a period of time. For example, NCs formed from poly(lactide-co-glycolide) can be designed to undergo biodegradation at varying rates depending on the ratio of lactide/glycolide. Using this characteristic of poly(lactide-co-glycolide), Kim *et al.,* 2012 were able to treat lung cancers in mice. In the study, doxorubicin (DOX) was loaded into polymeric poly(lactide-co-glycolide) and delivered to C57BL/6 mice with B16F10 lung cancer cells. The study showed that DOX-loaded poly(lactide-co-glycolide) gradually released DOX over two weeks, reducing tumour mass.[28] However, there are limitations to polymer NCs including non-heterogeneous distribution of particle size, increased risk of particle aggregation and toxicity, non-biocompatibility, immunogenicity and lack of degradation.[29]

Compared to polymeric NCs, some protein-based drug NCs such as apoFt and small heat shock proteins, are naturally present in mammals. As these protein NCs originate from mammals, they have a reduced risk of being immunogenic, are biodegradable, biocompatible, homogenous, and show high thermal stability.[30] A recent study conducted in mice evaluating the immunogenicity of horse spleen ferritin (Ft) concluded that horse spleen Ft did not elicit a higher immune response than the control group (injected with saline), indicated by similar white blood and platelet counts.[31] The low/no response is likely owed to the high sequence identity (91%) between horse spleen Ft and mouse Ft. In the case of mice or humans, the risk of an immunogenic response would be even lower, given that human heavy chain Ft and mice heavy chain Ft share a sequence identity of 93%. In addition to low immunogenicity, other attributes that make protein-based NCs attractive include their highly defined repetitive structures, which result in the regular distribution of moieties.

As outlined in Table 1.2, both polymeric and protein-based delivery systems have beneficial characteristics that make them useful for delivering therapeutics. However, as the ultimate aim of this project is to deliver therapeutic compounds/proteins into animals and then humans,

human heavy chain apoferritin (huH) was chosen. This is because huH is native to humans and would elicit a low immunogenic response compared to polymeric-based systems. In addition, huH, like all protein-based NCs, forms a symmetrical cage, a characteristic that is important for the current PhD project, which involves quantitative uptake studies on a HER2 targeting ligand. The display of a targeting/ligand in a symmetrical manner would naturally produce a fixed nanocage to ligand ratio and, therefore, would remove data variability caused by batch-to-batch differences that would typically occur for polymeric systems. This characteristic of protein NCs would allow an accurate conclusion on the uptake effect of the HER2 targeting ligand. As such, the project focused on protein-based NC, huH.

	Advantages	Disadvantages
Polymeric-based drug delivery systems	• Sustained drug release • tuneable	• Immunogenic response • Aggregation and toxicity *in vivo*
Protein-based drug delivery systems	• Biocompatibility • Biodegradability	• High cost of production • Fixed size - unable to increase drug load

Table 1. 2 A comparison between polymeric and protein-based drug delivery systems.

1.2 Protein-based nanocarriers

ApoFt is innately equipped with key characteristics to aid its normal function as an iron storage protein, including its ferroxidase core and the ability to bind and be internalised by transferrin receptor 1 (TfR1) expressing cells. Both apoFt's iron-storage characteristics and its structural features, which have driven apoFt as a NC for targeted cancer therapy, will be expanded in section 1.2.1. In addition to analysing apoFt's inherent characteristics, it is also important to compare apoFt with other protein-based NCs. Understanding and analysing the advantageous characteristics of other protein-based NCs and the modifications applied to these systems will give an insight into how apoFt can be improved as a NC. The protein-based delivery systems that have gained significant interest include small-heat shock proteins and virus-like proteins such as cowpea mosaic virus (CPMV). The de novo design of protein cages is also worth mentioning, as it has recently become a new research focus.[32–35] A size comparison of the apoFt against protein-NCs, including CPMV, can be seen in Figure 1.1.
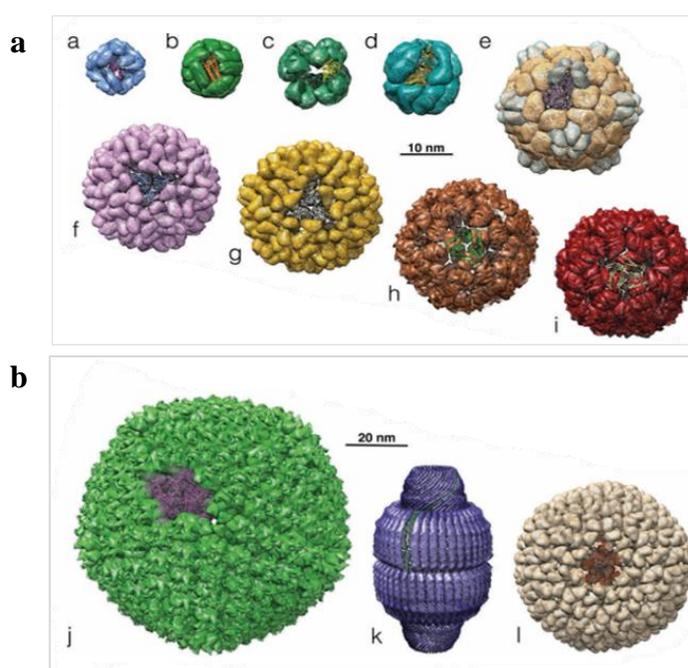
Figure 1. 1 Size comparison nature-derived protein cages. **a**, (a) Small heat-shock protein (400 kDa) [32], (b) Apoferritin (480 kDa) [265], (c) Pyruvate dehydrogenase multienzyme complex (600 kDa) [266], (d) Thermosome (1061 kDa) [267] , (e) Cowpea mosaic virus (3940 kDa) [268], (f) Brome mosaic virus(4600 kDa) [269], (g) Cowpea chlorotic mottle virus(4600 kDa) [270], (h) Bacterio phage lambda (14.2kDa per 180 subunit) [271] and (i) Bacteriophage MS2 (13kDa per 180 subunit) [272]. **b**, (j) Human adenovirus (156000 kDa) [273]. (k) Vault particle (240 kDa) [274] and (l) Bacteriophage P22 (320 kDa) [256, 275]. Size bar for **a**, (10 nm) and **b**, (20 nm).

1.2.1 Ferritin

1.2.1.1 Iron and ferritin

In most organisms, iron metabolism is vital for many biological processes, including DNA biosynthesis, oxygen transport, and anabolic metabolism.[36, 37, 38] However, the human body must regulate Fe^{2+} concentration within a defined range to prevent deviation from homeostatic levels. Given that Fe^{2+} is redox-active, elevated levels of Fe^{2+} can lead to the generation of reactive oxygen species (ROS), affecting signalling pathways, including the death receptor (extrinsic apoptosis) pathway.[39] However, cells have evolved to use specialised proteins and systems for regulating/transporting Fe^{2+} including the iron transport and storage proteins, transferrin and Ft. Transferrin is the primary protein involved in iron transport, with a healthy individual having 2.0 - 3.6 mg/ mL of transferrin in their blood.[40] The high levels of transferrin can be attributed to the fact that each transferrin molecule can only carry two Fe^{3+}.[41] In comparison, Ft, which is present in the serum at much lower levels, in the ng/ mL range with the majority being in the cell cytoplasm, can store up to 4500 Fe^{3+}. Ft that is sequestered in the blood is known as serum Ft. Serum Ft is primarily secreted by hepatocytes and lymphoid cells and is normally glycosylated.[42, 43] The levels of serum Ft vary with age and gender, but as a clinical reference, this level lies within the ranges of 30 - 300 ng/ mL and 10 - 200 ng/ mL for healthy men and women, respectively.[44] The level of serum Ft is relatively low and is generally considered a by-product of damaged cells rather than a means of iron transport.[45, 46] As such, high serum Ft levels are associated with several diseases, such as inflammation and metabolic syndrome.[47, 48] The low levels of Ft in serum are also owed to Ft's primary function, as an iron store protein within cells, converting Fe^{2+} into an insoluble and safe Fe^{3+} hydrate form, Fe(O)OH, within its core.[49]

1.2.1.2 Ferritin family

The Ft family is found ubiquitously in all organisms, reflecting its central role in iron storage. However, environmental pressures and variable energy consumption have led to distinct evolutionary Fts, with low sequence similarities (< 50%) between prokaryotes, archaea and eukaryotes. This divergence has led to structurally and characteristically different Fts, including DNA-binding proteins from starved cells (DPS), *Archaeoglobus fulgidus* and huH. DPS is one of the smallest Fts and comprises only 12 subunits, which form a 9 nm diameter cage with a 2,

3-symmetry axis. Its primary function is to protect chromosomes from oxidative damage, during carbon or nitrogen starvation and irradiation, by inducing chromosomal condensation.[50] Recent mutagenesis work on DPS has shown that a F47E substitution along the AB loop enables DPS's 2-, 3- symmetry axis to adopt a 2-, 3-, 4- symmetry, similar to the structure found in mammalian Ft (Figure 1.2).[51] This has not only elucidated the evolutionary mutation responsible for the divergence of Ft members but also highlights the potential importance of the AB loop in changing the structural characteristics of Fts such as huH.
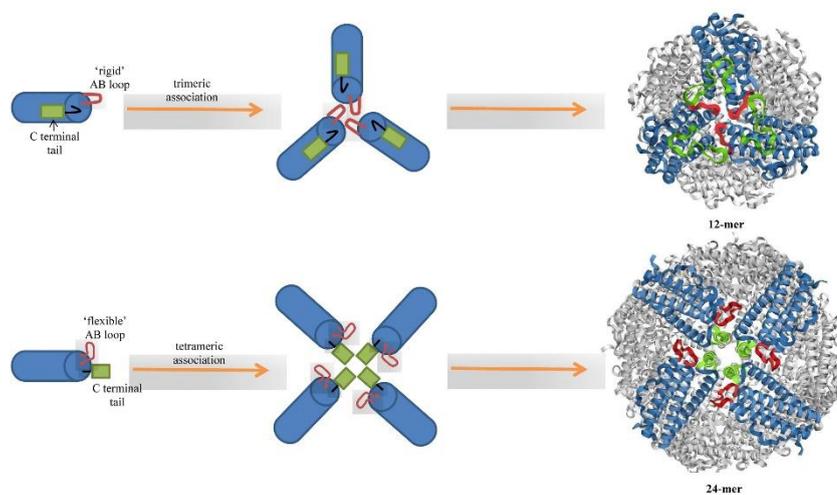
Figure 1.2 Effects of the F47E mutation on DPS' trimeric interface. The F47E mutation results in the AB loop to adopt a flexible structure and allows the C-terminal tail to protrude, resulting in a 24-mer cage.[51]

Although DPS exhibits interesting innate structural characteristics, it does not have valuable NC properties that could be translated into huH, such as high thermal resilience. Therefore, the focus must be shifted towards the archaea Ft family, including *Pyrococcus furiosus* Ft (*py*Ft) and *Archaeoglobus fulgidus* Ft (*af*Ft), which possess unique characteristics. Although from the archaea branch, and as such not ideal for human-based therapeutics as they can elicit an immune response, the characteristics of archaea Fts can still be adapted to mammalian-based Ft. *Py*Ft is a 24-mer protein with a 2-, 3-, 4- point group symmetry. It can withstand up to 100 °C for 24 hrs, an evolutionary trait caused by *py*Ft's anaerobic and hyperthermophilic environment. Another beneficial trait is its ability to nucleate silver and phosphates; a characteristic owed to the different amino acids (AA) present at its nucleation site.[20, 21] This feature, if incorporated

into huH, would expand the range of compounds that huH could deliver. Similar to *Py*Ft, *Af*Ft has also evolved from a thermophilic environment and has the classical 24-mer subunit assembly but self-assembles with a 2-, 3- symmetry. The absence of a 4- symmetry point leads to four large tetrahedral channels (45 Å wide) forming during assembly. In addition to its four large tetrahedral channels, *af*Ft also has the ability to disassemble/reassemble with the aid of metal ions such as magnesium or iron ions or in the presence of salts.[54, 55] However, some problems could arise when incorporating these characteristics into huH including the presence of both the metal ion and the desired chemotherapeutic within huH's core, leading to speculations about the drug's observed therapeutic efficacy. Another problem that may arise is the incorporation of the metal ion within the protein's interface, affecting the protein structure/function. Nevertheless, to expand huH's capabilities, it would be beneficial to incorporate some of the properties that bacterial and archaeal Fts possess, such as their high thermal resilience. However, it should be stated that due to some of the differences between archaea Ft and huH, such as the 2-, 3-, 4- symmetry axis, incorporating archaea's traits might affect the structural features of huH such as the 3-4 Å wide hydrophobic and hydrophilic channels (Figure 1.3).

3-4 Å
Wide

Core diameter 8 nm

Cage diameter 12 nm

C-terminus

N-terminus
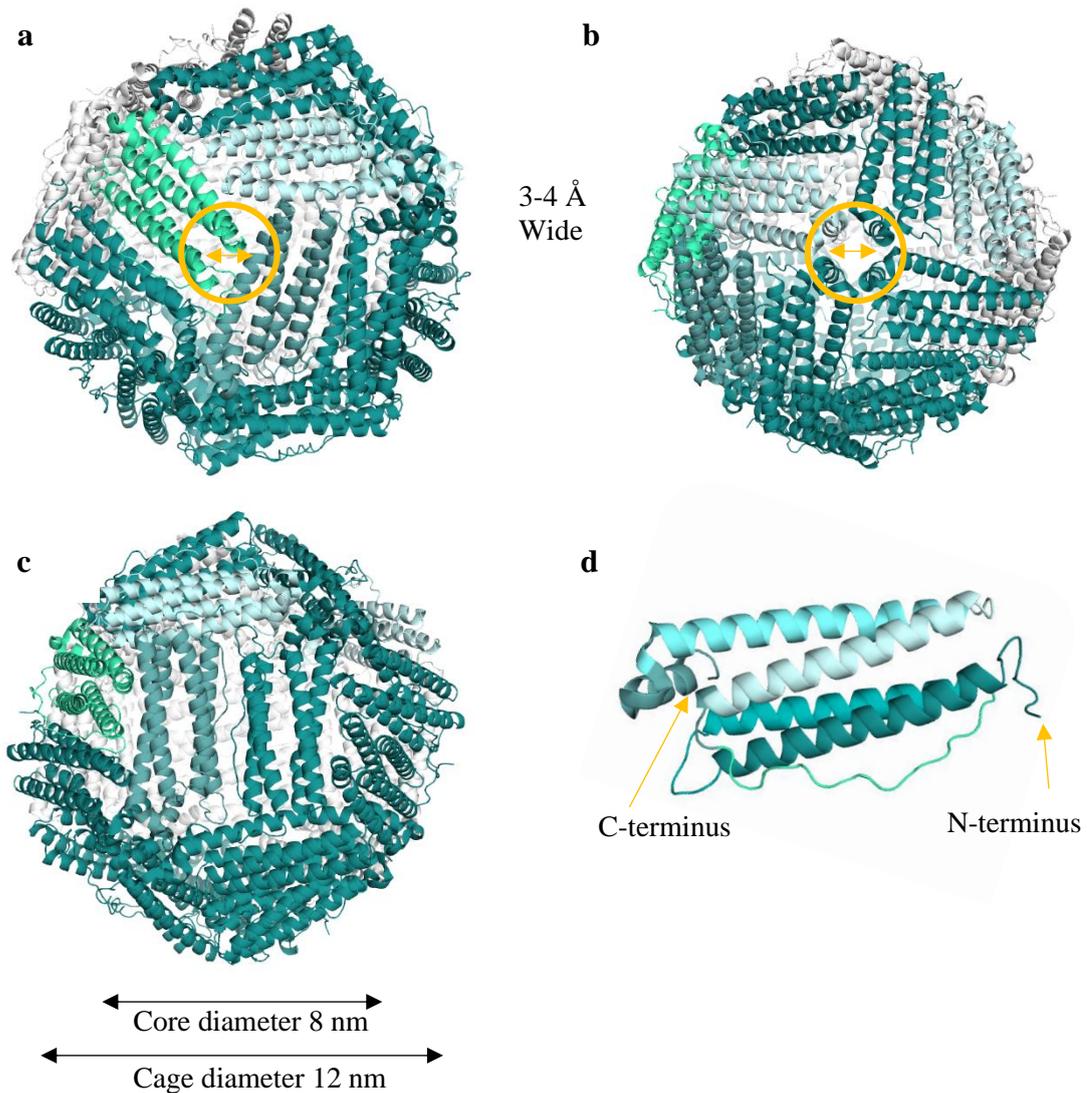
Figure 1. 3 Quaternary structure of huH cage. **a**, 3-fold interface, **b**, 4-fold interface, **c**, 2-fold interface and **d**, huH subunit. PDB 2FHA.

Similarly, to archaeal Ft, mammalian Ft comprises 24 subunits that form a 12 nm octahedral 2-, 3-, 4- symmetry cage, with an 8 nm hollow core (Figure 1.3). The arrangement of the subunits results in Ft having six hydrophobic and eight hydrophilic channels that are 3-4 Å wide. The subunit exists as two homologues, heavy chain (H; 21 kDa) Ft and light chain (L; 19 kDa), with a 55% identity, and are able to fold into a similar secondary and tertiary structure as seen in Figure 1.4. Mammalian Ft forms a cage with any combination of H and L subunits, as both subunits have different properties and attributes that contribute to iron storage. Ft that is composed of only 24 human heavy chain Ft (huH) has many advantages as a therapeutic nanocage delivery platform due to its ability to catalyse iron, cellular TfR1 targeting capabilities

and ability to encapsulate small molecules.[56] As a result huH has been used to encapsulate and deliver small molecules and compounds, including magnetic resonance imaging agents, genes, anticancer drugs, fluorescent dyes and other compounds for cancer therapy (Table 1.3).

```
huH   MTTASTSQVRQNYHQDSEAAINRQINLELYASYVYLSMSYYFDRDDVALK
          +SQ+RQNY  D EAA+N  +NL L ASY YLS+ +YFDRDDVAL+
huL       SSQIRQNYSTDVEAAVNSLVNLYLQASYTYLSLGFYFDRDDVALE

huH   NFAKYFLHQSHEEREHAEKLMKLQNQRGGRIFLQDIKKPDCDDWESGLN
         + +F    + E+RE E+L+K+QNQRGGR   QDIKKP  D+W
huL   +GVSHFFRELAEEKREGYERLLKMQNQRGGRALFQDIKKPAEDEWGKTP

huH    AMECALHLEKNVNQSLLELHKLATDKNDPHLCDFIETHYLNEQVKAIKEL
       AM+ A+ LEK +NQ+LL+LH L + + DPHL CDF+ETH+L+E+VK IK+
huL    DAMKAAMALEKKLNQALLDLHALGSARTDPHLCDFLETHFLDEEVKLIKK

huH    GDHVTNLRKMGAPESGLAEYLFDKHTLGDSDNES
       GDH+TNL ++G PE+GL EYLF++ TL
huL    GDHLTNLHRLGGPEAGLGEYLFERLTL
```

Figure 1.4 Amino acid alignment of huH and human light chain ferritin (huL). huH: Green, huL: Blue, alignment: Black, Gaps: Absence or non-conserved amino acid substitution and (+): Conserved amino acid substitution

Encapsulated material	Type of ferritin	Application	Ref
Fe$_3$O$_4$	huH	MRI contrast agent	57, 58, 59
	huH		60
^{64}Cu	huH		61
Gd^{3+}	hrH		62
	apoFt		63
siRNA	huH	Gene silencing	64
Cisplatin	hrH	Chemotherapy	65, 66
	huH		67
Doxorubicin	huH		68, 69, 70
Curcumin	huH		71
ZnF16Pc	apoFt	Photodynamic Therapy	72
GFP	*af*Ft	-	73

Table 1. 3 Application of apoFt in treatment and diagnosis of cancer. huH: Human heavy chain Ft, hrH: Horse heavy chain Ft, apoFt: Hollow ferritin, GFP: Green fluorescent protein and *af*Ft: *Archaeoglobus fulgidus Ft*

1.2.1.3 Ferroxidase activity

As mentioned earlier, human Ft can incorporate up to 4500 Fe^{3+} within its cage. This innate trait is owed to the H subunit, which contains a ferroxidase region that catalyses Fe^{2+} to Fe^{3+}. In comparison, the L subunit is not involved in ferroxidase activity and has only been shown to aid iron core stabilisation.[74, 75]

Most of the AAs involved in Fe^{2+} oxidation are present in the dinuclear ferroxidase region, found along helix B (Figure 1.5a). These residues are highly conserved among different species and consist of binding sites A and B, with coordinating residues E27, E61, E62, D65, E107 and Q141 (Figure 1.5b).[76, 77] Mutations of these critical residues to their equivalent L subunit AA residues (A61, K62, A64, G65 and A67) lead to a loss in ferroxidase activity and iron nucleation.[78] During the two-step ferroxidation process, oxygen is consumed, which leads to a momentary production of hydrogen peroxide and a peroxo intermediate, as seen in Equation (1) in Scheme 1.1. Once Fe(O)OH ions are produced, the final reaction begins and functions

in an autocatalytic manner. This process is outlined in Equation (2) and is known as the mineralisation process. Finally the hydrogen peroxide produced in Equation (1) is used up during the detoxification step, Equation (3). Oxidation of Fe^{2+}, using O_2 or H_2O, leads to rapid sequestration and mineralisation of up to 4500 Fe^{3+}/nanocage, producing a stable Fe(O)OH mineral core.[79] In addition to its ferroxidase ability, the H subunit also has the ability to be internalised via TfR1. This is a complementary ability to its ferroxidase ability, enabling human Ft to move iron between neighbouring cells.
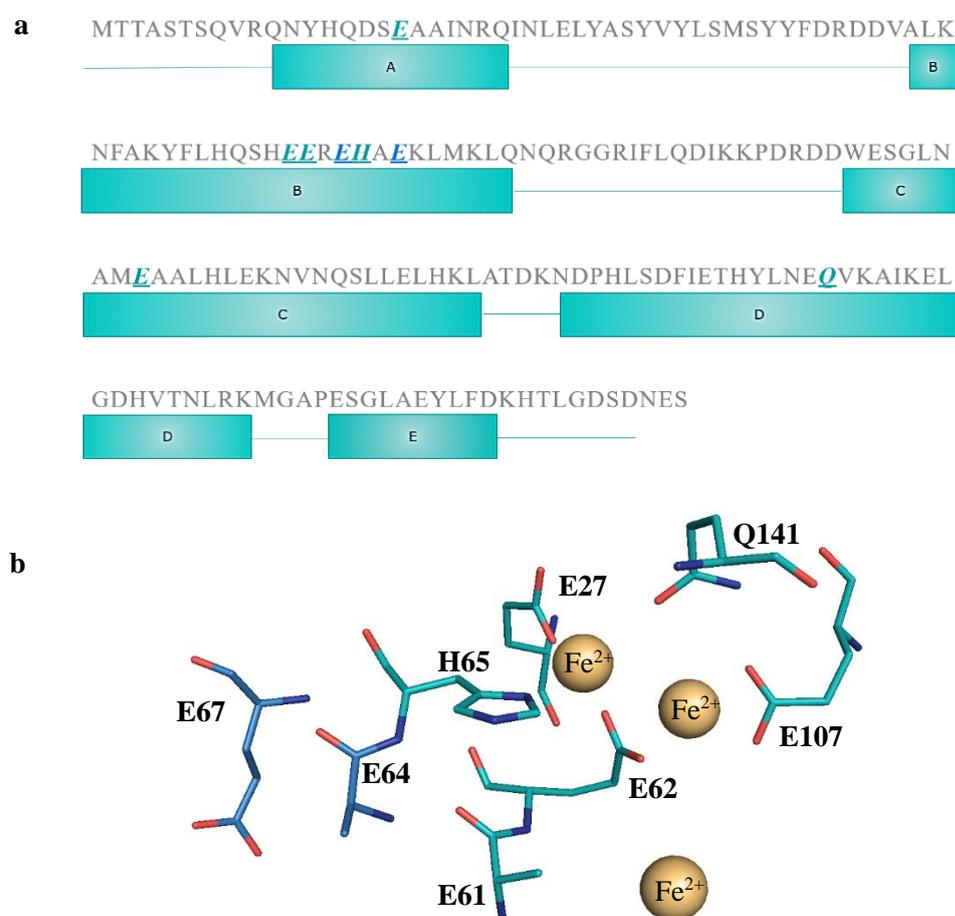
Figure 1.5 Ferroxidase and putative iron nucleation centre of huH. **a,** Schematic representation of huH primary structure. Residues involved in ferroxidase activity are italicised and coloured, and **b,** coordination environment around Fe(II). Carbon atoms of the residues involved in ferroxidation, and putative nucleation are shown in green and light blue, respectively. PDB: 2FHA

$$2Fe(II) + O_2 + 4H_2O \rightarrow 2Fe(O)OH \ (Core) + H_2O_2 + 4H^+ \tag{1}$$

$$2Fe(II) + O_2 + 6H_2O \rightarrow 4Fe(O)OH \ (Core) + 8H^+ \tag{2}$$

$$2Fe(II) + H_2O_2 + 2H_2O \rightarrow 2Fe(O)OH \ (Core) + 4H^+ \tag{3}$$
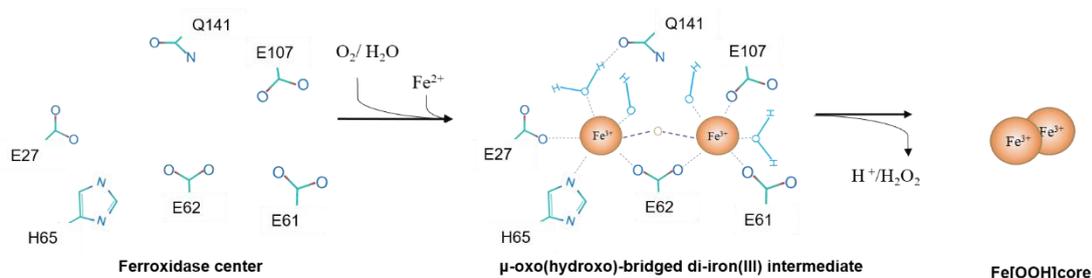
Scheme 1.1 Three step process showing the iron oxidation process in the ferroxidase centre. Scheme adapted from reference [276]

1.2.1.4 Human heavy chain ferritin's internalisation

As mentioned earlier, huH has the innate ability to be internalised by TfR1/cluster of differentiation 71 (CD71). However, research by Chen *et al.,* 2015 stated that T-cell immunoglobulin and mucin domain protein-2 (Tim-2) could also facilitate huH uptake.[80] However, the research also concluded that Tim-2 is only expressed on specific cells, including liver, kidney, spleen, B-cells and T-cells. In comparison, TfR1, a homodimeric type II transmembrane protein, is present on all cells, except for mature erythrocytes and some terminally differentiated cells.[81] The expression levels of TfR1 vary among cells, with immature erythroid cells, placental tissues and cancer cells, including breast, colon, lung and thyroid cancers displaying a higher density of TfR1 on their surfaces.[116, 117, 84, 85, 86, 87] However, it should be noted that this is not the case for all cancer types, such as progressive benign kidney tumours, where TfR1 is not overexpressed.[88, 89] Considering its upregulation in many cancers, as shown in Figure 1.6, multiple NC-based treatments have exploited TfR1 as a targetable biomarker for cancer treatment, including the use of a monoclonal anti-TfR1 antibody (mAb A24) to block Ft binding, resulting in apoptosis in tumour T lymphocytes.[90]

In the case of iron uptake, both huH and transferrin are able to utilise TfR1 as a means to deliver iron into cells. Experimental studies examining transferrin/huH and TfR1 binding showed that

mutations in the conserved tripeptide arginine-glycine-aspartic acid (RGD) motif of TfR1 still efficiently permitted the uptake of huH but not transferrin, indicating that huH and transferrin have distinctly different TfR1 contact points.[91] Further examination of huH uptake in numerous human cell lines, including activated lymphocytes and blocking studies with an anti-TfR1 antibody, showed that the majority of huH uptake in these cells can be attributed to TfR1.[92] Although the exact region where huH binds to TfR1 had still not been elucidated De Turris *et al*., 2017 discovered that engineering archaeal apoFt, with nine residues of huH, IFLQDIKKP (from residue 70 to 79) located between helix B and helix C permitted internalisation of the new chimeric archaea apoFt via human TfR1 expressing cells (Figure 1.7).[93] Following this, single-particle cryo-electron microscopy data revealed key regions on huH that were involved in TfR1-huH binding: the external BC loop and the N-terminus of the A-helix. The key residues identified along these regions include Q14, D15, E17-A19, N21, R22, N25 (A-helix), E116, K119, D123 (the C-terminus of the C helix) and R79, F81, Q83, K86, K87 (on the external BC-loop).[94]
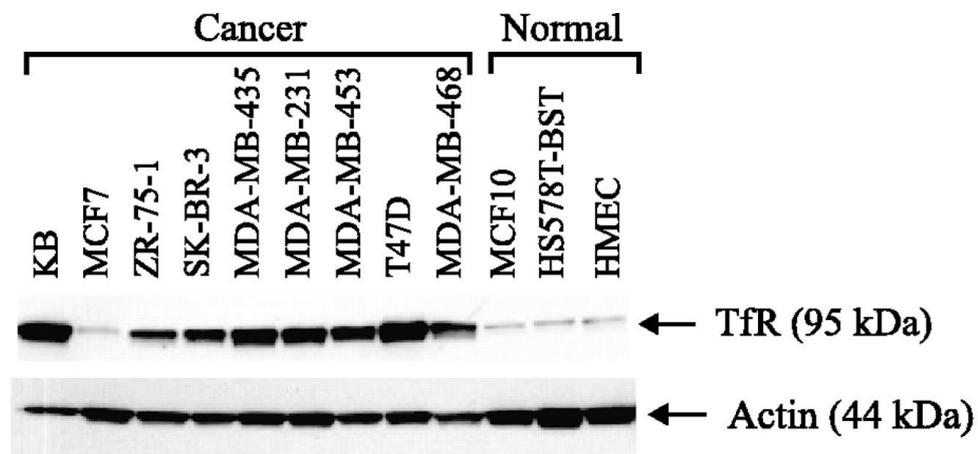
Figure 1.6 Western blot analyses of TfR1 (95kDa) expression in normal and cancer cell lines. KB: Human nasopharyngeal epidermoid carcinoma, MCF7/ZR-75-1/SK-BR-3/ MDA-MB-435/MDA-MB-231/MDA-MB-453/T47D/MDA-MB-468: Human breast carcinoma, MCF10/HS578T-BST: Normal mammary immortalised cell lines and HMEC: normal mammary primary cells. Edited 21/2/2018 [243]
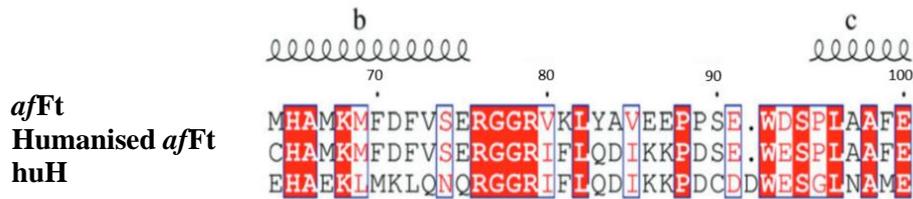
Figure 1.7 Shows a sequence alignment of the BC loop from Archaeal Ft and huH. The alignment shows similarities between the archaeal and human Ft, whilst narrowing the region human heavy chain Ft binds to TfR1. *af*Ft: Archeal Ft, Humanised *af*Ft: humanised archaeal Ft and hHCft: human heavy chain Ft. Adapted from reference.[93]

Once bound to TfR1, huH undergoes clathrin-coated pit-mediated endocytosis.[95] Following uptake, huH is compartmentalised within the early endosome, which later matures in the late endosome. huH is then trafficked into the autophagosomes to finally reach the lysosome, where it is degraded or is diverted to the nucleus for genomic protection (Figure 1.8). The huH degradation process, termed ferritinophagy, depends on many factors, including the high serum levels which have been shown increase levels of intracellular huH in the lysosome.[96]
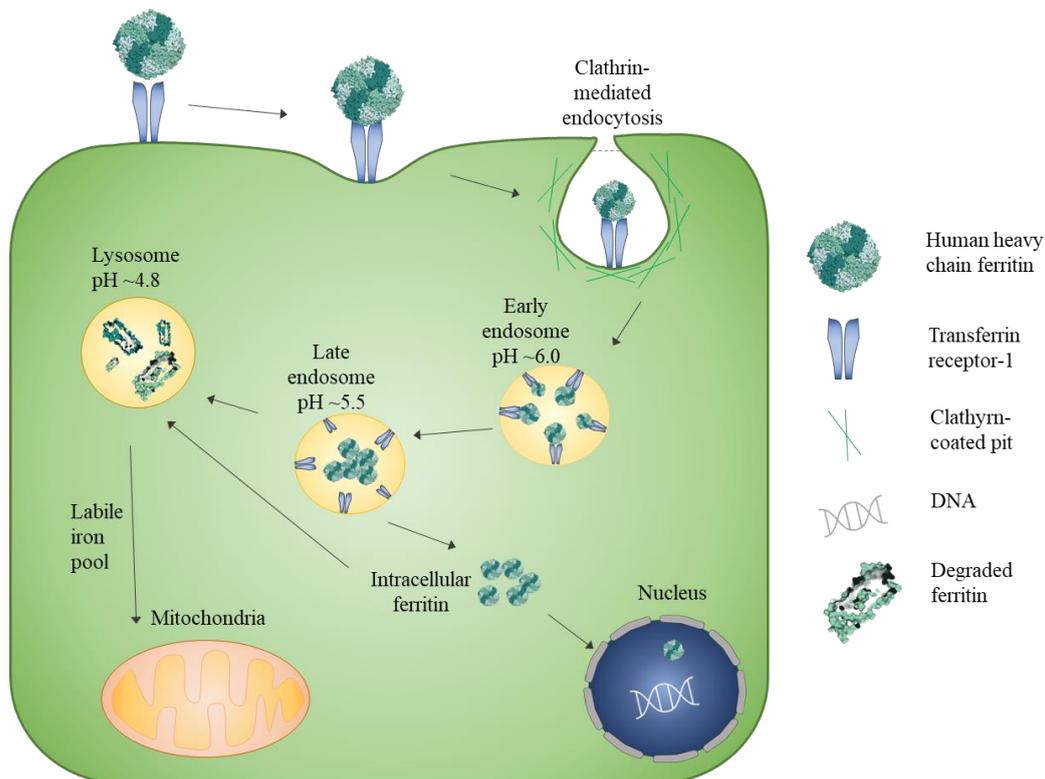
Figure 1.8 The possible endocytic pathways following Ft internalisation. Ft binds to TfR1 and is endocytosed via clathrin coated pit-mediated endocytosis. The budded vesicle fuses with an endosome to form the early endosome. The early endosome transforms into the late endosome, where the environment is more acidic, causing Ft to dissociate from TfR1. Ft can then either be degraded in the lysosome or be released into the cytosol for nucleus localisation.

1.2.1.5 Three-fold and four-fold channels

Prior to and during ferroxidase activity, iron is transported into the Ft core via channels present at huH's interface. The channels naturally occur during Ft's 24-mer cage assembly and formation of the 2-, 3-, 4- interface, leading to 8 three-fold and 6 four-fold channels naturally forming. In human Ft, the three-fold channels are present at the three-fold axis and are mainly lined with hydrophilic residues from the C-helix of each subunit.[97] These hydrophilic residues comprise of three-negatively charged residues, D131, E134, and D139, which are conserved in vertebrates and plant Ft but not in invertebrate Ft.[98] Douglas and Ripoll calculated the electrostatic potential around the three-fold channels and showed that the presence of an electrostatic gradient from the entry point of the three-fold channels towards Ft's core.[99] This supported experimental data, which showed that the three-fold channels are the site of iron entry for ferroxidase activity.[100, 101] In comparison, the four-fold channels are formed from the E-helix (four adjacent Ft subunits) and consist of primarily hydrophobic residues. Unlike the

three-fold channels, the electrostatic potential around the four-fold channels is directed away from the cavity, hindering cation entry. This is because the hydrophobic residues lining the four-fold channels do not generate an attractive force to aid ion entry into the cavity. As such, the ions need to overcome a high energy barrier at the gate to permeate the channel.[102] However, it has been noted that the four-fold channels have a similar composition to the gramicidin A channel, which has been shown to conduct protons.[103] Therefore, it can be hypothesised that Ft's four-fold channel could have a similar function and be involved in H^+ movement during ferroxidation.

1.2.1.6 Engineering ferritin's interfaces

As well as aiding the formation of the three-fold and four-fold channels in Ft, the 2-, 3-, 4-interface also plays a crucial part in stabilising huH. In addition, huH's unique interface also enables huH to have high thermal resilience and be able to undergo pH-dependent assembly/disassembly at pH 7 and pH 3, respectively. However, characteristics such as small channel size and extreme pH disassembly conditions have limited the number and type of bioactive compounds that huH could encapsulated, therefore hindering huH's advance in nanomedicine. Early genetic mutations on the C-terminal sequence, located near the 4-fold interface of huH, highlighted key residues that would beneficially change the characteristic nature of huH, e.g. higher pH disassembly point.[104, 105] Following this, Chen *et al.,* 2016 introduced an amber codon close to the C-terminus, at residue 160, resulting in the deletion of the final 23 AAs.[106] The deleted segment comprised the flexible loop between the D and E helix (DE turn) and the E-helix, both of which are involved in forming the four-fold channels (Figure 1.9). The new fabricated non-native apoFt showed a significantly lower molecular weight of 17 kDa and a decreased overall negative charge of the core but was seen to dissociate at a higher pH (pH 4.0), while still reassembling at neutral pH. An identical construct was also produced by Ahn *et al.,* 2018.[107] The authors used Cryo-EM to show that the new mutant had a channel diameter of 18 Å and an expansion of the cage by 2 Å outward.[107] However, the variant produced by Ahn *et al.,* was shown to disassemble at pH 6.0, higher than that stated by Chen *et al.,* 2016. One possible explanation for the difference in pH disassembly points seen could be a consequence of NaCl in the final protein buffer. In the case of Ahn *et al.,* 2018, protein disassembly was performed in Tris and 150 mM NaCl. In comparison, Chen *et al.,* 2016, only used Tris buffer. Given that NaCl influences the movement of ionisable residues, the simple presence of NaCl could have changed the disassembly point. Nevertheless, considering these

new unique characteristics, the newly-engineered apoFt mutant showed promise and could be exploited for encapsulation and delivery of pH-sensitive bioactive compounds as well as larger compounds.
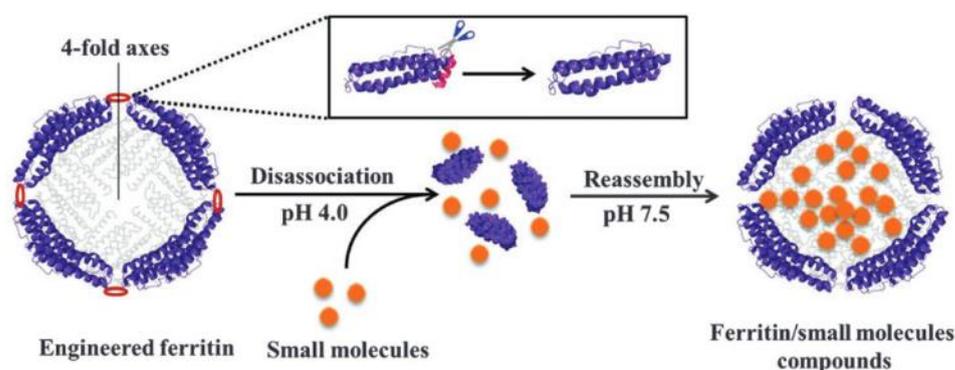
Figure 1.9 Schematic representation of the ΔDE mutant. The new mutant was shown to disassemble at pH 4.0.[257]

While the C4 interface has been shown to have little effect on the overall protein shell structure, studies on the C3-C4 interface indicate that changes to this region can alter the symmetry of the cage. Zhang *et al.,* 2016 demonstrated that deletion of six non-essential subunit-subunit interface residues located along helix D, 139NEQVKA144 produced a 17 nm diameter non-native apoFt.[108] The resulting complex was a heteropolymer with 48 subunits composed of two different subunits (same AA sequence, but different folding configurations) derived from an identical polypeptide chain. It was also noted that the change in structure altered the three-fold interaction along which the eight hydrophilic channels are present: increasing the channel size from 3-5 Å to 25 Å. However, the non-native apoFt was unstable in solution, due to the electrostatic repulsive forces that occur between subunit interfaces. Parallel to this, Zhang *et al.,* 2016 showed that insertion of a heptapeptide, into the same region along helix D, produced a 10 nm by 8 nm 16-mer cage with a new pseudo-D2 interface.[109] The new pseudo-D2 interface produced four holes, each with an area of 6.5 Å but obscured the C-terminus of two subunits. TEM images of the 16-mer cage showed dissociation at pH 3.0 and reassembly at pH ~7.5. Although both constructs demonstrated the important role of the C3-C4 interface in the geometry of apoFt, they are not promising in terms of advancing apoFt in nanomedicine, as they both lead to structural instability.

1.2.2 Small heat shock proteins

Ft being one of the smallest protein nanocages in nature, contains notable NC properties including TfR1-targeting capabilities, 14 channels and pH-dependent disassembly for encapsulation. However, it is essential to examine other protein NCs, as it will aid the structural development of the Ft cage.

As noted in section 1.2.1, Ft is composed of 24 α-helical subunits resulting in a 12 nm cage with high thermal resilience. Given this beneficial characteristic, other protein nanocages, including the heat shock protein from the hyperthermophilic archaea *Methanococcus jannaschii* family (*mj*Hsp) have also adopted a similar structure. *mj*Hsps have a very similar architecture to Ft, being composed of 24 monomers (Figure 1.10) that self-assemble into a globular structure with an outer diameter of 12 nm and an inner core diameter of 6.5 nm, featuring eight trigonal 3 nm channels and six square 1.7 nm channels.[110] However, each monomer of *mj*Hsp is mainly composed of beta-strands as opposed to Ft's alpha-helical-rich subunits. The main function of Hsps are to aid the synthesis and proper folding of proteins during the presence of a stressful stimulus. Due to their important role in protein folding, Hsps have evolved to be resilient against thermal denaturation, with some members, like those from the hyperthermophilic archaea *Methanococcus jannaschii* family (*mj*Hsp), withstanding temperatures of up to 70 °C.[111] As such *mj*Hsp has been evaluated to deliver many anticancer drugs including DOX and 3-phosphoinositide dependent protein kinase 1 inhibitor OSU 03012 (Figure 1.11).[112, 113, 114] Although Hsps can withstand extreme conditions, their relatively small internal capacity to hold compounds and the presence of large channels of 3 nm, make Hsps less favourable than other protein NCs for delivery.
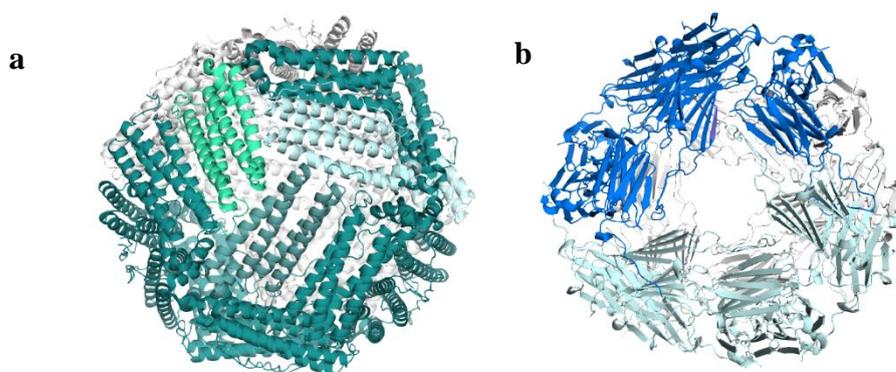
Figure 1. 10 Secondary structures of huH and mjHsp. **a**, Three-fold channel view of Ft (Left, PDB: 2FHA) and **b**, Triangular channel of mjHsp (Right, PDB: 1SHS).
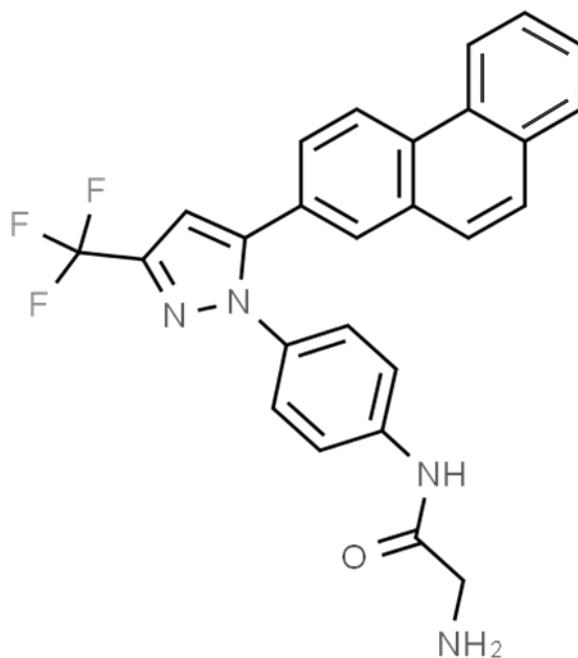
Figure 1.11 Chemical structure of OSU 03012; a
PDK1 inhibitor AR-12. F: Fluorines, N: Nitrogen, O:
Oxygen.

1.2.3 Virus-like proteins

Both Ft and *mj*Hsp have a diameter of 12 nm and as such are on the lower end of the protein nanocage scale. In comparison, virus-like proteins (VLPs), which are virus capsid-forming proteins, once void of genetic material, are among the largest types of protein nanocages (ranging from 10 to 1000 nm). In addition, they have the innate ability to target and be taken up by cells. Among the VLPs, only CPMV shows similar characteristics to human Ft. This includes CPMV's ability to self-assemble into a 60-mer from two types of subunits: small (24 kDa) and large (41 kDa) subunit. In addition, CPMV can also withstand a wide range of pHs from pH 3-8 and temperatures reaching 70 °C.[11, 12] Two key benefits of using CMPV as a NC are its positively charged core and its ability to interact and enter endothelial cells *via* the cytoskeletal protein vimentin. Vimentin is an intermediate filament protein that is mainly expressed in mesenchymal cells, including endothelial cells. Tumours originating from epithelial cells that show upregulation of vimentin usually have metastatic properties such as migratory capabilities.[116, 117] Chao Wang *et al.,* 2019 showed that the presence of CPMV near

ovarian cancers led to immune-mediated anti-tumour effects.[118] Together, these findings promoted the use of CPMV as a drug delivery system.

As well as using the innate biology of VLPS to target cancers, VLPs have also been modified to make them better therapeutic cancer vaccines. One such approach was by using the SpyTag/SpyCatcher conjugation system, which utilises an engineered surface fibronectin-binding protein of *Streptococcus pyogenes* (SpyCatcher) to recognise and covalently bind a thirteen AA peptide (SpyTag), as schematically shown in Figure 1.12. In such a way, Palladini *et al.,* 2018, genetically fused a HER2-antigen to the SpyCatcher sequence. This allowed the attachment of a HER2 antigen to AP205, a bacteriophage VLP cage, which had the SpyTag sequence genetically inserted. The resulting complex showed an increase in anti-HER-2 IgG response and halted mammary carcinoma progression.[119]
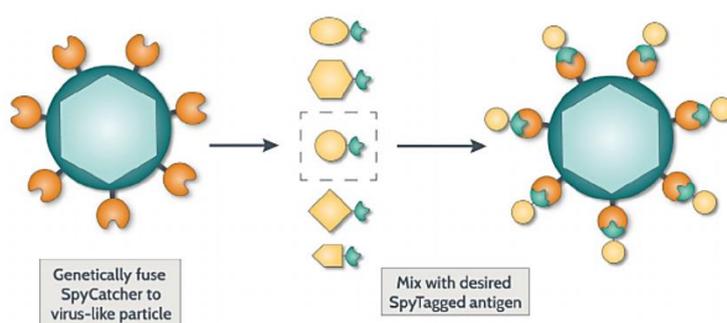
Figure 1.12 The SpyTag/SpyCatcher system. The SpyCatcher sequence is genetically inserted onto the exterior-facing region of VLPs and expressed. The SpyTag is also genetically inserted upstream of the ligands/antigens. Both components are mixed, resulting in a ligand/antigen presenting VLP. Modified from [119].

1.2.4 *De novo* **proteins**

As with any nature-derived protein NCs, several challenges arise, including reduced cellular uptake, low half-life, compound release tunability, and particle toxicity. The *de novo* design of protein NCs provides the opportunity to tailor nanocages with beneficial traits such as size and environmental responsive disassembly. However, this requires extensive knowledge of peptide bond length, protein-protein interaction, and protein folding, all of which need to be inputted into a computer-designed structure and modelled, leading to extensive processing times. Nevertheless, Cristie-David and Marsh, 2019 were able to design and construct tetrahedral,

octahedral, and icosahedral nanocages using a simple parallel coiled coils method (Figure 1.13).[120] They were also able to achieve metal-dependent assembly/disassembly of a nanocage through the introduction of a histidine residue at the assembly interface. Using a similar trimeric building block method and RosettaDesign calculations, a software used to identify the AAs used in protein interfaces, King and colleagues, 2014 were able to design *de novo* nanocages of different structures.[35] The expressed proteins showed similar architecture to those predicted by computational design models and mimicked the highly organised architecture of protein nanocages with high atomic accuracy. As such, the use of *de novo* design methods would allow fast and programmable novel features to many protein NCs such as Ft, overcoming drawbacks seen in many engineered protein nanocages.
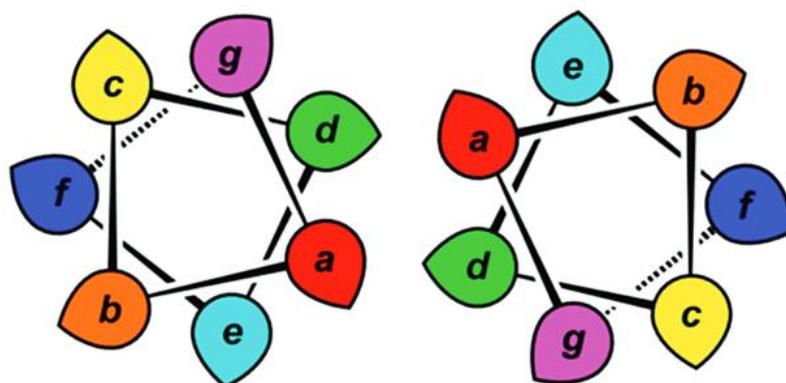
Figure 1.13 Dimeric coil-coil assembly. A left-handed supercoil with 3.5 residues per turn with a heptad repeat configuration. a and b are hydrophobic residues. c, d and f are background residues not involved in the interface. e and g are polar residues involved in ionic- or hydrogen- bonding. Modified from [277].

1.3 Use of apoferritin for medical purposes

As outlined in section 1.2.1, apoFt exhibits a range of beneficial characteristics, including its pH-dependent disassembly/assembly property and its innate ability to target TfR1-overexpressing cancer cells. As such, apoFt's characteristics have made it applicable for a range of medical purposes, including delivery of chemotherapeutics, photodynamic imaging therapy and use in vaccines.

1.3.1 Delivery of chemotherapeutics using apoferritin

Chemotherapeutic agents such as cisplatin and anthracyclines, including DOX, epirubicin and idarubicin, have been used to treat a wide range of cancers (Figure 1.14). However, a major disadvantage encountered with cisplatin and anthracyclines is their dose-dependent toxicity towards non-cancerous cells and poor accessibility to tumour sites, leading to significant collateral death in healthy cells. This problem can be partially resolved by encapsulating chemotherapeutics into NCs such as apoFt.[77, 78] The concept of encapsulating molecules and ionic species such as Zn^{2+} and Ni^{2+} within apoFt stemmed from the observation that apoFt could dissociate at pH 2.0 and spontaneously reassemble the 24-mer cage upon an increase to physiological pH: thereby sequestering and entrapping solvents and small inorganic molecules.[123, 124] However, a recent small-angle X-ray scattering study on apoFt pH disassembly and reassembly, showed that this process can cause irreversible apoFt cage damage and as such affect apoFt's stability.[125] To circumvent this possible problem, encapsulation can also be achieved by passive diffusion, also known as the 'nanoreactor' route, via the 3-4 Å channels of the intact apoFt cage (Figure 1.15).[126] However, the latter route is restricted to small positively charged molecules or ions that can sterically and electrostatically pass through the channels.[127] Although it has not been reported how cations other than iron can enter through the three-fold channel, it has been observed that iron transfer is mediated by conserved glutamic acid residues along the channel.[128] Following passage through the three-fold channel, the encapsulated positively charged ions or the protonated amine regions of a compound such as epirubicin and idarubicin can electrostatically interact with clusters of glutamic acid residues present in the interior of apoFt's cage. This electrostatic attraction also prevents diffusion of some of the compounds out of the cage, thereby increasing the encapsulation efficiency of the compound.[129, 272] Using this entrapment method, DOX has been encapsulated within apoFt, with the entrapped DOX showing a 10-fold higher intra-tumoural drug concentration than free DOX, thereby demonstrating the high capability of apoFt as a drug NC (Figure 1.16).[69] In addition, delivery of apoFt-encapsulated DOX has also been shown to improve the maximum tolerated dose of DOX by four-fold when compared to the free drug. In comparison to DOX, cisplatin's aqua complex cis-$[PtCl(NH_3)_2(H_2O)]+$ size is much smaller and, therefore, can easily diffuse through apoFt's three-fold channel (Figure 1.15a). Once encapsulated, cisplatin can bind to residues such as His132 on the D helix, slowing down its release.[122]

	R1	R2	R3	R4	R5
Doxorubicin	OCH_3	H	OH	NH_2	OH
Epirubicin	OCH_3	OH	H	NH_2	OH
Idarubicin	H	H	OH	NH_2	H
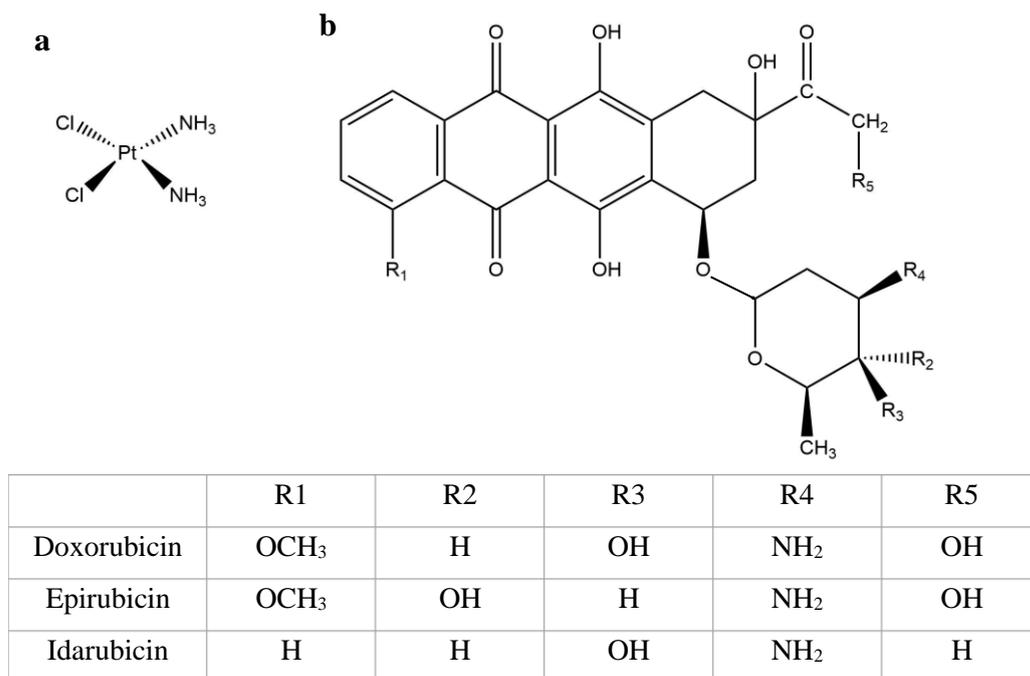
Figure 1. 14 Chemical structure of **a**, Cisplatin and **b**, Anthracyclines. R_n groups: Possible sites of modifications.
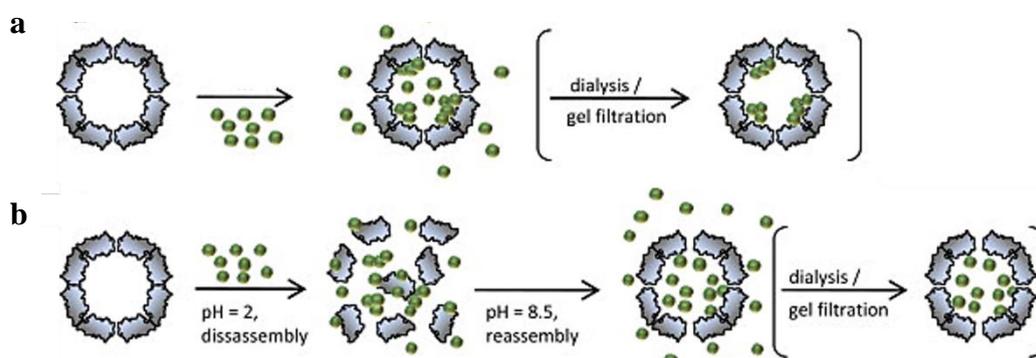
Figure 1.15 Methods of encapsulating compounds in apoFt. **a**, Encapsulation via the 'nanoreactor' route. Compounds passively diffuse into the apoFt core via the channels. Dialysis and gel filtration are then used to remove excess compound. **b**, Encapsulation via the pH-based disassembly/assembly. At < pH 3.0, disassembles the apoFt cage into subunits. Addition of compounds and then an increase to > pH 8.0, results in apoFt cage formation and encapsulation of the compound. Dialysis and gel filtration are then used to remove excess compound. Modified from reference.[133]
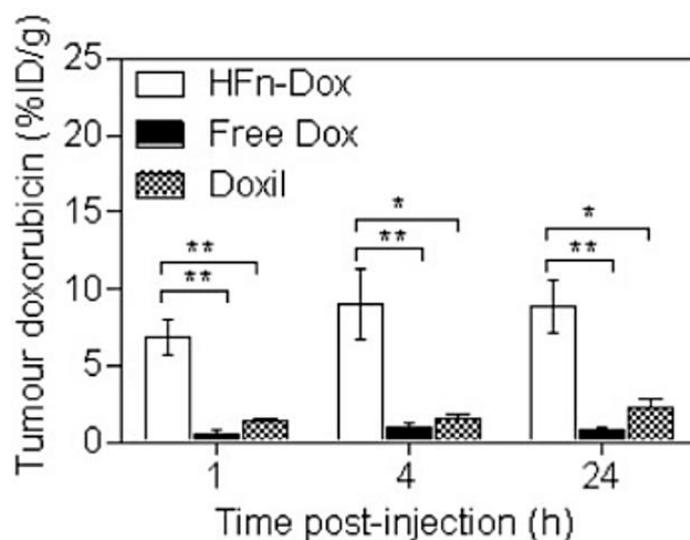
Figure 1. 16 The concentrations of DOX in tumour at 1, 4, and 24 hrs post-injection times. H-Ft-Doxorubicin (HFn-Dox), non-encapsulated (Free DOX) and nontargeted liposome (Doxil). Bars represent means ± SD, Student t-test.[278]

Encapsulation of compounds larger than cisplatin, including benzothiazoles and zinc-based compounds such as zinc hexadecafluorophthalocyanine ($ZnF_{16}Pc$) (Figure 1.17a) can also be achieved via a simple gradient diffusion (nanoreactor route) through apoFt's channels.[130–132] However, in the case of benzothiazoles such as 2-(3,4-dimethoxyphenyl)-5-fluorobenzothiazole (GW610), a potent experimental anticancer compound (Figure 1.17b) its poor water solubility makes it unfavourable for encapsulation. To overcome this problem, AAs exhibiting a range of charge and polarity were conjugated to GW610 with basic or non-polar AAs. The resulting prodrugs demonstrated enhanced encapsulation efficiency compared to the parent GW610 at pH 7.4. By nature, the linear structure of benzothiazole and its derivatives (Figure 1.18) make them more favourable for encapsulation compared to other drug compounds, including DOX, which require urea or hydrostatic pressure to aid encapsulation via the three-fold channel.[133, 134] However, small and linear compounds are also more susceptible to 'leakage' from the three-fold channel than larger compounds, resulting in a faster drug release rate. For example, in the case of benzothiazole GW610, 50% drug release was observed in 4 hrs, significantly higher than DOX, which only showed 20% release after 60 hrs, at pH 7.4.[133, 135] Large compounds including $ZnF_{16}Pc$, a photosensitizer investigated for use in image-guided photodynamic therapy, have also been encapsulated within apoFt via the nano-

reactor route. However, given that $ZnF_{16}Pc$ has a diameter >10 Å, it raises a question as to how the compound was able to diffuse through apoFt's 3-4 Å channels rather than adhering to the external surface. The author reasoned that the presence of DMSO in the drug solution could have contributed to channel widening, but this hypothesis was never proven. Permeation studies on huL, indicate that uncharged molecules larger than maltose (13 Å) but smaller than maltotetraose (20 Å) in diameter can diffuse through apoFt's 3-4 Å wide channels.[136] Therefore, it is reasonable to speculate that apoFt's channels are larger and more flexible than those calculated from *in silico* models of apoFt.
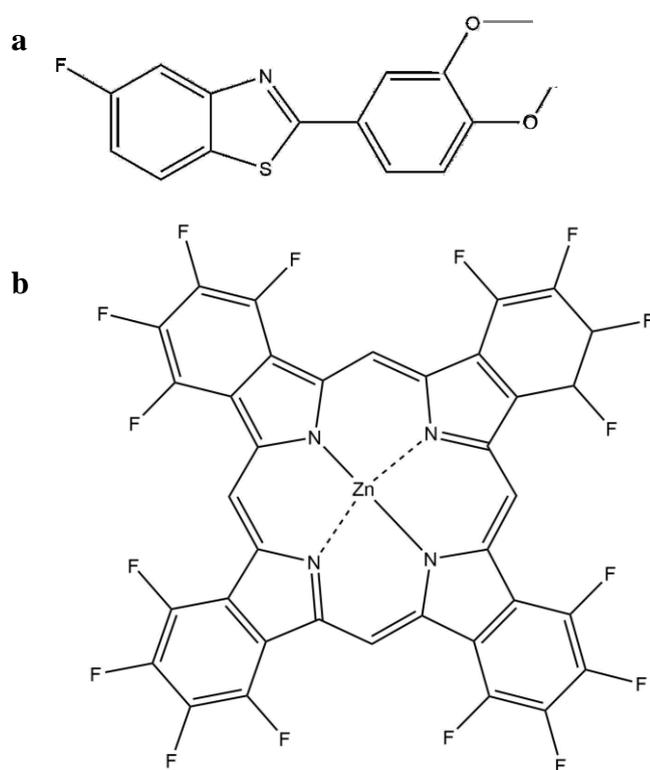
Figure 1. 17 Structure of **a,** Benzothiazole GW610 and **b**, $ZnF_{16}Pc$
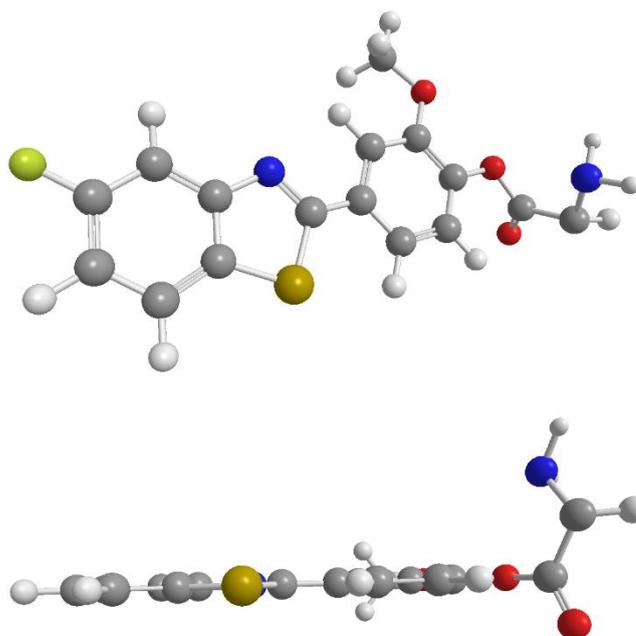
Figure 1. 18 Ball and stick view of benzothiazole derivatives. Fluorine: Green, Sulphur: Yellow, Nitrogen: Blue, Oxygen: Red, White: Hydrogen, Carbon: Grey.

1.3.2 Delivery of photodynamic agents using apoferritin

As well as a delivery system for cancer therapeutics, apoFt can also be used in molecular imaging. One way Ft is naturally suited for this role is by its innate ability to store iron, a heavy atom, and as such, leading to its use in electron cryotomograms or MRIs.[137, 138] In addition, apoFt is naturally found in every organism, and it is, therefore, possible to monitor Ft in both normal and diseased cells. However, given that Ft levels vary significantly between different tissues and even between individuals, it would be impossible to compare or verify the link between Ft levels and a specific stage of a disease. Nevertheless, apoFt's three-fold channel has permitted it to encapsulate imaging-based compounds, including zinc hexadecafluorophthalocyanine, a photosensitiser for phototoxicity in glioblastomas, and gadolinium complexes for MRI imaging.[132] Using the disassembly/reassembly route, hypocrellin B (HB), a photosensitiser naturally found in *Hypocrella bambusea,* was loaded into apoFt. Generation of ROS rby HB was achieved by exposure to 630 nm LED light, resulting in triple-negative breast cancer cells, MDA-MB-231 to have reduced viability over time (exposure time of 64 seconds) when compared to the free compound.[139] However, even though this type of therapy shows promise, there are some problems that may arise, including the narrow time frame of ROS production. As ROS production can only be stimulated by wavelengths at 630 nm, the tiosming of stimulation needs to be accurate, as early stimulation could lead the HB-

apoFt to produce ROS outside the tumour, which could be detrimental to normal cells and tissues. However, if HB-apoFt is irradiated too late, e.g., after endocytosis, apoFt would be degraded by lysosomes, leading to the release of HB, and reduced phototoxicity compared to HB-Ft.[139]

1.3.3 Exterior modification of apoferritin

In addition to encapsulating compounds within its core, the exterior of apoFt can also be modified either genetically or via chemical conjugation. In the case of apoFt, the protrusion of the N-terminus to the outer surface and the presence of lysine and cysteine residues on the exterior of the cage provides the opportunity to display 24+ therapeutic molecules either by chemical modification or genetic engineering. In addition, due to the configuration of subunits, in a 2, 3, 4 symmetrical manner, any proteins that are engineered on the N-terminus will form a 4-bundle 'cluster' around the exterior of the four-fold channel. This cluster configuration increases the binding potential of the engineered protein/peptide for their respective receptor/target and, therefore the avidity of the therapeutic. This characteristic of apoFt has been utilised for immune-dependent therapies to stimulate a higher immune response while using a lower dosing concentration.

When large proteins or peptides are involved, such as those for immune-dependent therapy, the target of interest can be engineered onto the N-terminus, the flexible loop region between helix B or helix C and, more recently, the C-terminus. Surface exposure of proteins engineered onto the C-terminus, which is normally located inside the protein core, results from apoFt's insufficient cavity volume, ~2.1 x 10^3 Å, for the containment of large peptides. The lack of cavity space causes the C-terminus to flip out, exposing the target peptide. In doing so, the C-terminus engineered targeting protein/peptide has 24 copies that are mainly distributed around the 3-fold axis.[140] This configuration can also be achieved using a truncated version of huH, whereby the E helix is shortened (sFt). Using this apoFt variant, Kim *et al.,* 2016, were able to dually functionalise apoFt's exterior, with the N-terminus presenting a pro-apoptotic peptide and the C-terminus presenting a GFP molecule (Figure 1.19).[141] The study also confirmed that the presence of GFP molecules on the exterior of sFt prevented TfR1 mediated endocytosis. However, the addition of the 'CGKRK' on the pro-apoptotic peptide enabled uptake of sFt-GFP via the p32 receptor. The presence of the GFP molecules probably did not block p32-CGKRK binding as the alpha-amphipathic pro-apoptotic peptide would have protruded

outwards via the N-terminus flexible chain, preventing any GFP obstruction. Nevertheless, the dual-payload was able to successfully enter MDA-MB-231 cells and promote apoptosis.
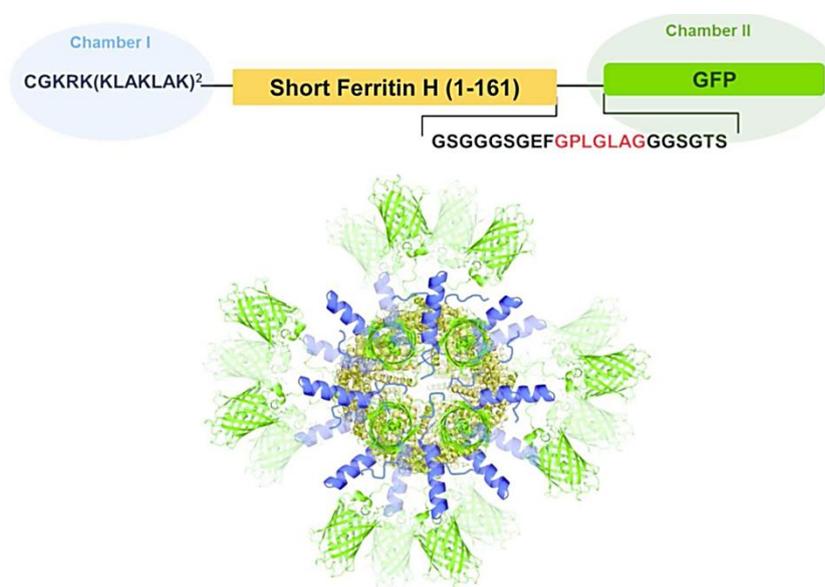
Figure 1. 19 Dual presentation of GFP and a pro-apoptotic peptide on short Ft. MMP2 cleavage site, GPLAGLAG (Red). Modified from [141]

ApoFt can also be chemically modified to increase apoFt's capacity to delivery therapeutic compounds. One such case is the conjugation of cyclodextrin to apoFt, called cyclo-Ft for the delivery of insoluble compounds. Cyclodextrins were either covalently attached to the exterior of apoFt via copper(I)-catalysed azide/alkyne cycloaddition or were encapsulated via the pH disassembly method with the aid of iron ions. Conjugation was achieved by introducing a cysteine mutation on the exterior of Ft, allowing the successful attachment of cyclodextrin via a propargyl maleimide. Successful internalisation of cyclo-apoFt and reduction in cholesterol levels via attachment of cycldextrins to cholesterol, were seen in macrophage foam (lipid-laden) cells, with a higher reduction seen with externally functionalised cyclo-apoFt than that of cyclodextrins that were encapsulated. Conjugation of cyclodextrin to apoFt did not seem to hinder binding and uptake of apoFt via TfR1, even though the site of conjugation was located on the BC loop region, a region thought to be involved in TfR1-huH binding. One possible reason for this was the size of β-cyclodextrins and the relative flexibility of the BC loop. Even

though the primary aim was to reduce intracellular cholesterol in macrophages and therefore prevent atherosclerosis development, there is potential for therapeutic delivery of insoluble chemotherapeutic agents, including paclitaxel and therefore, the strategy is worthy of consideration as a viable functionalisation alternative for the protein endocytosis, enhancing its drug delivery ability.[142, 143]

As outlined, an external facing load such as GFP or cyclodextrins, can be displayed on apoFt's exterior either by engineering it onto apoFt N-/C- terminus or chemical conjugation. Using these two routes, many therapeutic peptides or proteins have been engineered on apoFt's exterior for the development of vaccines and cancer-targeted therapeutics.

1.3.3.1 Vaccines

As noted earlier, apoFt's exterior can be systemically functionalised with therapeutic proteins, either via protein engineering or chemical modifications. This has promoted apoFt to be used in an array of fields, including as a VLP for antigen display in the development of vaccines.[145, 146] High symmetry and self-assembling characteristics, as well as its ability to display antigens in a multivalent organised manner, has made apoFt an important scaffold for vaccine development. Although many other protein nanocages such as virus particles, have these innate abilities, they have the disadvantage of eliciting their own immune response, while species-specific apoFt can be manufactured to overcome this problem. Using what Ivelin *et al.*, 2018 referred to as a two-component system, apoFt was able to present two antigens on its surface, derived from the HIV-1 envelope and influenza haemagglutinin.[144] This was achieved by mixing two different apoFt subunits, that were genetically engineered to harbour the desired antigens on their N-termini. The resulting cage allowed dual antigen display of both an influenza antigen and an HIV antigen, and as a result stimulated an antibody response against both the respective antigens. An equimolar ratio of each Ft displaying both antigens was achieved by using insect *Trichoplusiani* Ft which naturally reassembles with 12 copies of both L and H chain.[145] Another apoFt-based vaccine was also investigated against enterovirus 71, a virus that causes hand, foot and mouth disease.[146] Here researchers functionalised segments of the enterovirus 71 virus on apoFt's exterior in three different locations: N-terminal region, C-terminal region, and the BC-loop region, with no modification altering apoFt's ability to assemble. Although all three epitope regions induced an immunisation response, the epitope located along the loop region was the only one to produce an antibody neutralisation

response.[146] This could be a consequence of several factors, including the accessibility of antibodies to the loop region or the distance between epitopes on the cage, which could hinder or expose the epitopes region. This study highlighted the importance of ligand/antigen length and the spatial distribution of antigens on a 3-D cage, both of which could affect the strength of the immune response.

1.3.3.2 Peptides/proteins for cancer therapy

In addition to vaccine development, apoFt's characteristic of displaying 24 peptides/proteins on its exterior has also paved its way for cancer therapy. Examples of peptides that have been engineered onto apoFt include prominin-1 (PROM) peptide (8 AA: TISWPPRC) and silaffin (R5) peptide (15 AA: SSKKSGSYSGSKGSK), used for targeted cancer stem cell therapy and for sustained drug release, respectively.[95, 96] In the case of PROM, a simple prominin-1 targeting peptide was engineered to the exterior of apoFt for the delivery of irinotecan, a topoisomerase I inhibitor. In comparison, the R5 peptide was not used as a targeting peptide but was engineered onto apoFt to increase its drug delivery properties. Genetic fusion of R5 to apoFt resulted in a silica-enveloped apoFt, which was used to deliver DOX in a sustained and pH-dependent manner.

In the case of PROM and R5, both ligands were small peptides, reducing any obstruction for TfR1 to bind to apoFt. However, as the size of engineered protein/peptide increases, the relationship between apoFt and the engineered ligand becomes more complex. Many things must be considered, including steric hindrance of neighbouring peptide/protein, the flexibility of the displayed peptide/protein and stability of apoFt cage. These considerations were taken into account when Lee *et al.,* 2017 attached an epidermal growth factor receptor (EGFR/HER1) targeting affibody (a small 52 AA molecule discussed in section 1.4) to apoFt.[149] To increase the half-life of apoFt, varying lengths of XTENs, which are unstructured proteins composed of hydrophilic residues and are known to extend the *in vivo* half-life, were engineered onto the C-terminus of apoFt, termed long circulating apoFt nanocages (LCNC). 3D modelling of LCNC highlighted cluster formation of 'clouds' around apoFt and pharmacokinetic analysis of LCNCs showed increased circulation time for nanocages with longer length XTEN peptides including LCNC144 and LCNC288. Using both 3D modelling and pharmacokinetic profiles, it was determined that longer length XTEN peptides formed a larger hydration state, sufficient to cover the surface of apoFt, thus increasing its residency time. It was also shown that LCNC144 reduced binding to macrophages compared to wild type apoFt, decreasing the

reticuloendothelial system's clearance rate. Once it was verified that LCNC144 improved apoFt circulation residency time, an EGFR-targeting affibody peptide was engineered to the C-terminus of LCNC144 (Figure 1.20). *In vivo* studies in mice showed a faster accumulation rate of affibody-LCNC144 in EGFR-overexpressing 4T1 breast cancer tumours compared to apoFt and affibody-apoFt. However, it was also observed that the affibody-LCNC144 had a high accumulation rate in the liver, possibly because of an increase in apoFt's overall size.
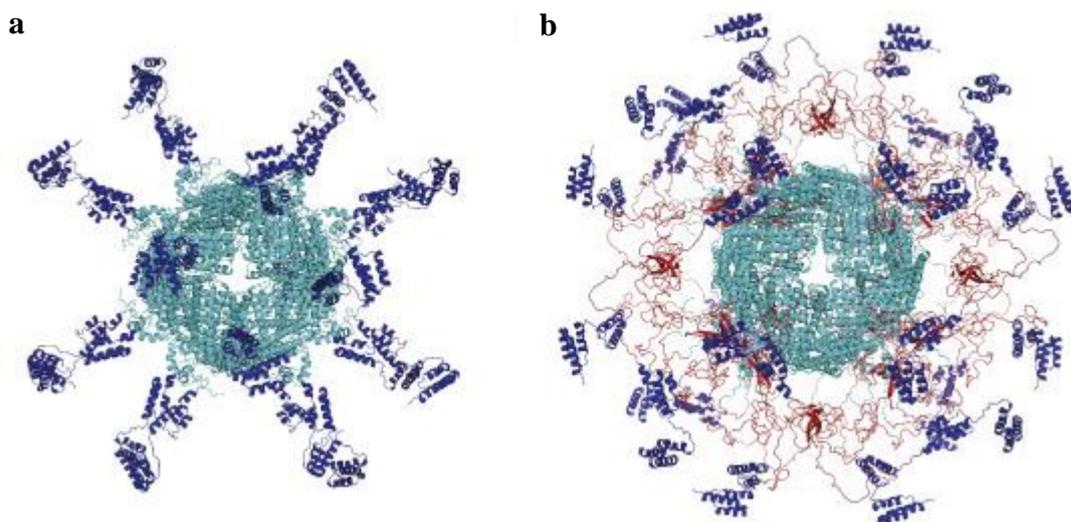
Figure 1.20 3D model of **a**, affibody-LCNC144 and **b**, LCNC144. Light blue: Ferritin, Red: long chain XTEN 144 peptide and dark blue: EGFR-targeting affibodies.[92]

With respect to the affibody-LCNC144 construct, both the affibody and the XTEN were genetically attached to the C-terminus of apoFt. To increase the therapeutic capabilities of apoFt as a NC for cancer treatment, both the N-terminus and C-terminus can be utilised as demonstrated in section 1.3.3.1, for displaying antigens against enterovirus 71. Utilising a similar design, tumour necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) and a tumour-targeting IL4 receptor-binding peptide (IL4rP) were fused to a short apoFt (lacking N-terminus AA residues 1-15, and helix five located at the C-terminus).[150] Initially, to increase the efficacious effect of the N-terminus engineered TRAIL ligand, a triple-helical domain was added to the C-terminal of the TRAIL protein. However, the trimer TRAIL-apoFt still showed a low efficacious effect similar to the single TRAIL monomer, indicating improper formation of the trimeric TRAIL protein. Therefore, a flexible linker was attached between the TRAIL ligand and apoFt, resulting in a10-fold higher cytotoxic effect compared to monomer TRAIL. IL4rP, with a MMP cleavage site for nonspecific binding, was also engineered onto the C-

terminus of apoFt, resulting in the final complex termed TRAIL-apoFt ^{IL4rP} (Figure 1.21). Although the resulting complex had a large diameter of 22.4 nm, the complex was still able to form a cage. *In vivo* studies of TRAIL-apoFt ^{IL4rP} indicated that the presence of IL4rP increased tumour localisation, with a 4-fold higher uptake compared to TRAIL-ATNC and a 6-fold increase uptake compared to TRAIL alone. In addition, binding and uptake studies also showed that TRAIL-apoFt ^{IL4rP} was not sequestered by apoFt's physiological receptor, TfR1. This indicated that both the TRAIL and IL4rP can block apoFt-TRf1 interaction. Both *in vivo* studies in breast cancer xenograft and orthotopic pancreatic cancer model, showed increased cytotoxic effects of TRAIL-ATNC^{IL4rP}, with breast cancer tumours decreasing by 64.2% and TRAIL only having a 20.2% growth inhibition.
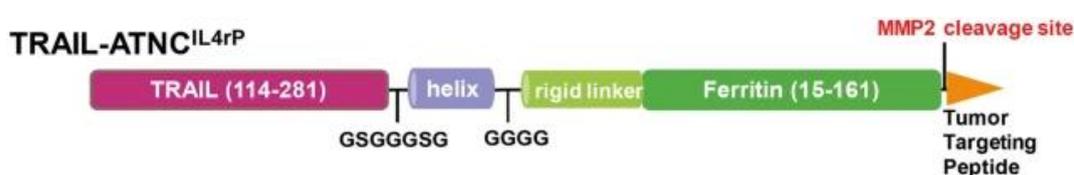
Figure 1. 21 Schematic diagram showing the TRAIL-ATNC^{IL4rP} construct. Figure has been modified to show only the schematic image of TRAIL-ATNC^{IL4rP} from reference.[98]

In the case of LCNCs and TRAIL-apoFt ^{IL4Rp} both therapeutic ligands are displayed on the exterior of apoFt. However, these types of constructs only utilise a part of apoFt's capability as a NC, given that they do not focus on encapsulating therapeutics within apoFt. As outlined in section 1.3.1, apoFt's nano core, with a diameter of 8 nm, can be used to encapsulate many therapeutic agents including therapeutic proteins/peptides. Because research focusing on apoFt's capabilities to encapsulate therapeutic proteins/peptides is not extensive, there is still the opportunity to develop apoFt as a NC for therapeutic proteins.

1.4 Therapeutic peptides or proteins

1.4.1 Encapsulation of peptides or proteins

In recent years, peptide therapeutics have gained momentum in cancer therapy because of their high selectivity, high potency and reduced off-target side-effects compared to small molecules.[151, 152] However, peptide drug development has been hampered due to many factors

including rapid elimination through the renal filtration system and proteolytic cleavage, which reduces their *in vivo* half-life.[153] To overcome some of these drawbacks, peptides have been encapsulated within NCs including liposomes. [154, 155] Drug delivery systems such as liposome and viral nanocages whose size ranges from 26-500 nm eclipse apoFt in terms of carrying capactiy, which has a more limited cavity diameter of 8 nm, and as such narrowing the type of therapeutic agents that can be encapsulated within apoFt.[156] Nevertheless, the successful delivery of a single insulin receptor-siRNA molecule (20–24 nucleotides) has been achieved using apoFt. Delivery of this complex into primary T-cells was shown to reduce insulin receptor levels, owing to the silencing effect of siRNA.[64] Encapsulation of siRNA was achieved via the disassembly/reassembly route and successful encapsulation was promoted by a change in charge of apoFt's core during disassembly. This change occurred was as a result of a drop in pH (<pH 5.0) resulting in the negative interior becoming increasingly positively charged and promoting the encapsulation of a single negatively charged RNA molecule. However, with regards to peptides, because of their small size, a higher number of therapeutic ligands can be encapsulated.

The idea of entrapping proteins within the apoFt cage has to date not been explored beyond the two examples detailed below. Until recently only an artificial transfer hydrogenase (ATHase) (MW: 16.5 kDa and diameter: 4.5 nm × 5.5 nm) and a green fluorescent protein (GFP) variant (MW:32 kDa and diameter: 4.2 nm x 2.4 nm) have been encapsulated within horse spleen Ft (hoFt) and archaea (*af*Ft), respectively.[157, 158] In the case of ATHase encapsulation, hoFt required pH disassembly and assembly of the cage.[158] In this rare case, ATHase was able to stably reform within the cage and as such reduce cyclic imines. However, most enzymes are not able to withstand such extreme pH conditions e.g., pH 3 during apoFt disassembly, with most proteins undergoing irreversible denaturation. In comparison, *Af*Ft has the added benefit of having 45 Å wide triangular channels, which facilitated GFP uptake. An initial theoretical encapsulation efficiency was calculated based on a 70% packing density, leading to the presumption that five GFP molecules could be entrapped within the core. This encapsulation efficiency was achieved via the nanoreactor route within the cage, with loading and stabilisation of GFP molecules aided by *af*Ft's negatively charged luminal cavity and GFP's positively charged surface. Looking at other protein nanocages that have encapsulated target proteins/peptides, it is possible to encapsulate proteins/peptides within protein NCs without compromising the encapsulated protein's structural integrity. For example, engineered fluorescent proteins such as enhanced GFP (EGFP) (MW:32.7 kDa and a diameter of 4.2 nm x 2.4 nm) have also been encapsulated within virus-like protein cages such as bacteriophage p22

(Figure 1.22).[159] p22 has an exterior diameter of 64 nm and an internal ellipsoid *cavity* diameter of 54 nm.[160] Fusion of EGFP to the bacteriophage's virus cage enabled the incorporation of EGFP during heterologous expression of p22's protein coat. The method to encapsulate fused proteins during heterologous expression has since been applied successfully to encapsulate more functional proteins such as alcohol dehydrogenase and cellulase B (MW: 58 kDa), a heteromeric β-glycosidase enzyme.[161, 162] However, encapsulation of larger proteins within virus-based NCs such as p22 is only possible due to p22's large cavity and pores, characteristics that apoFt does not have. Therefore, in the case of apoFt, a lower size limit must be identified to genetically engineer a protein at the C-terminus of apoFt for encapsulation. This is because once the number or size of the encapsulated protein surpasses the volume of the cavity, apoFt 'flips' the protein to the exterior of the cage. Such a phenomenon has been seen with a short apoFt lacking the fifth helix which had GFP engineered on its C-terminus (detailed in section 1.3.3.2).[163]
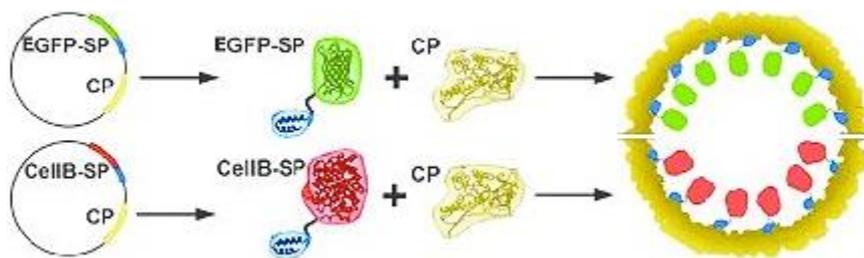
Figure 1.22 Formation of encapsulated CelB and EGFP in p22. Co-expression of P22 coat protein (CP, yellow) and N-terminal truncated scaffold protein (SP, blue) engineered to EGFP (green) or CelB monomers (red) results in encapsulation of desired protein. Modified from.[38, 39]

1.4.2 Cytochrome C and cancers

Given that human apoFt only has a small internal cavity of 8 nm, the number of suitable therapeutic protein candidates for encapsulation is limited. This selection is further filtered to therapeutic protein candidates that are able to withstand the extreme pH encountered during apoFt disassembly and assembly. Although limited, it not impossible to find a protein that carries these traits, such as cytochrome C (Cyt C). Cyt C is a small ~12 kDa haem containing protein, located between the outer and inner membrane of the mitochondria. Its primary

function is to transfer electrons between complex III (Cyt C oxidoreductase) and complex IV (cytochrome oxidase) in the electron transport chain.[164] Because of its crucial role in producing cellular energy, the loss of Cyt C has been shown to cause embryonic death in mice.[165] Given its essential role in the energy transport chain, Cyt C is highly conserved among animals with 53% identical residues among man, pig, rabbit and chicken.[166] With such high homology, most Cyt C homologues have a preserved tertiary structure: Four α-helices surrounding a haem group (Figure 1.23).[167]
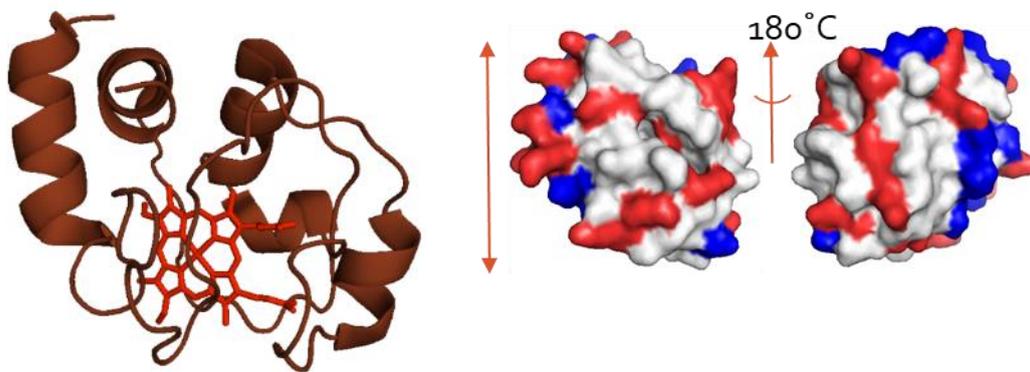
Figure 1.23 Cyt C from bovine heart. **a**, Cyt C polypeptide chain (dark red) surrounds it haem group (red) and **b,** Surface charge of Cyt C. Positively charge residues (red) and negatively charged residues (blue). PDB: 2B4Z

In addition to its role in energy production, Cyt C also plays another role in programmed cell death execution, specifically apoptosis. When an apoptotic stimulus is present, such as an increase in ROS, Cyt C is released from the mitochondria and mediates the oligomerisation of an adaptor molecule, apoptosis-protease activating factor 1 (Apaf-1). Apaf-1 then binds and activates caspase-9, which results in caspase-9 proteolytically self-processing into an active form. The activated caspase 9 then cleaves caspase 3/7 into their active forms, leading to apoptosis (Figure 1.24).[168] Due to its role in activating the pro-apoptotic pathway and its high stability in the presence of higher urea concentration (8 M), guanidinium hydrochloride and extreme pH (pH 3), Cyt C has been investigated for cancer treatment.[169, 170] Initial studies with Cyt C have shown that cancer cells, particularly breast cancers, are hypersensitive to Cyt C.[171] This hypersensitivity is caused by overexpression of putative human leukocyte antigen-associated protein-I (PHAPI), a tumour suppressor protein, which has been shown to enhance Apaf-1 oligomerisation, leading to increased caspase-9 activation in breast cancer cells.[171] Cyt

C-hypersensitivity is also seen in other cancer cells, including lung cancers. Manoj-Saxena *et al.,* 2018 showed that targeted delivery of Cyt C (conjugated to TfR), activated caspase-3 and caspase-9 in A549 lung cancer cells.[172] A similar result was achieved when Cyt C was encapsulated within a hydrophilic hyperbranched polyhydroxyl polymer nanoparticle (HBPN).[173] In such a case, folic acid, a ligand that binds to folate receptors, was conjugated to HBPN, allowing targeted treatment of A549 cells. Significant apoptosis of A549 cells was demonstrated compared to low-level folate receptor-expressing cells. Overall, these studies indicate a use for Cyt C in cancer therapy particularly those cancers that over-express the protein PHAPI. However, it should be noted that Cyt C like many proteins is susceptible to protease degradation, and therefore delivery via a NC would be the most suitable route, to ensure Cyt C's structural integrity before entry into the cell.
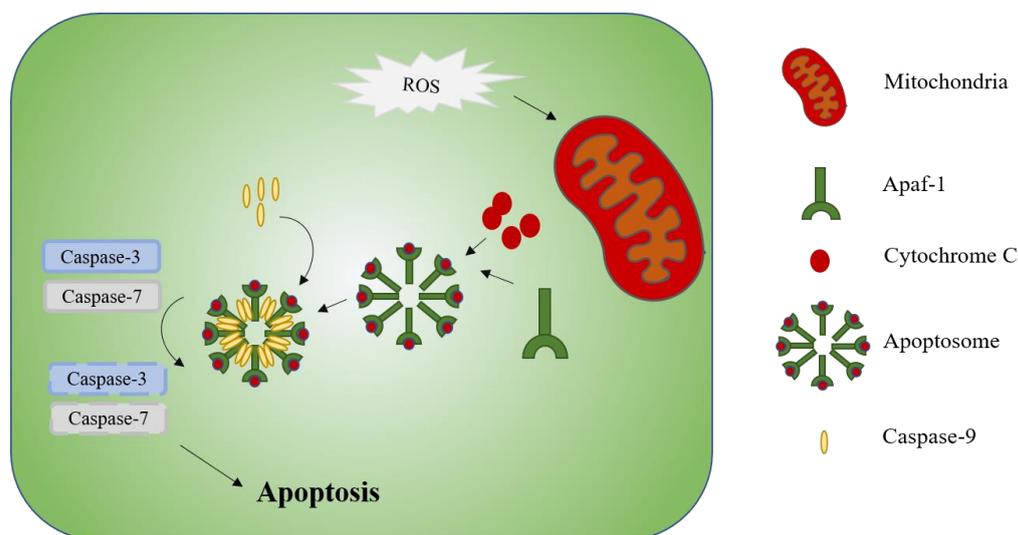
Figure 1.24 Cytochrome C's role in apoptosis. Increased presence of reactive oxygen species (ROS), leads to release of cytochrome C (Cyt C) from the mitochrondria. Cyt C binds to apoptotic protease activating factor 1 (Apaf-1), causing oligomerisation of Apaf-1, to form the apoptosome. This enables caspase-9 to be recruited, self-cleave and dimerise into an activated form. Caspase 9 then promotes capsase-3 and -7 to undergo autolytic cleavage. Activation of both caspase proteins causes apoptosis.

1.5 Human epidermal growth factor

1.5.1 Human epidermal growth factor receptor 2 (HER2/ErB2)

As well using therapeutic proteins such as Cyt C to target cellular proteins, external targets such as over-expressed receptors on cancers, can also be exploited for the treatment of breast cancers. However, most of the treatments in clinical trials are saturated with either inhibitors, monoclonal antibodies, or a combination of both. Monoclonal antibodies have two functions, for immunotherapy (e.g., enlisting a patient's own immune system to kill target cells) and to target specific cells. In breast cancer therapy, monoclonal antibodies have been targeted towards HER2. However, as with any therapeutic that targets HER2 or other over-expressing cell surface receptors, it is important to understand tumour heterogeneity given the varied response of different breast cancer cell types to treatments such as monoclonal antibodies.

As indicated earlier (section 1.1), breast cancers may be classified into four major subtypes depending on the overexpression or absence of three receptors: HER2, PR and ER.[174] HER2 belongs to the human epidermal growth factor receptor (EGFR) family which consists of (EGFR/ErbB1/HER1), ErbB2 (HER2), ErbB3 (HER3), and ERbB4 (HER4). All four receptors have a cysteine-rich extracellular ligand-binding site, a transmembrane region, and an intracellular domain which has tyrosine kinase catalytic activity (Figure 1.25).[175] Upon ligand binding, these receptors can form homo- or hetero-dimers leading to autophosphorylation of the tyrosine residues. Dimerisation of different EGFRs has been shown to recruit different cellular proteins, leading to differential signalling and diverse biological effects.[176] Unlike other members of the EGFR family, HER2 does not change conformation between its active and inactive form and therefore exists in its constitutively active form, with no known EGF-like ligands.[177] Therefore, its signalling is activated by dimerisation with other EGF receptors. The main signalling pathways that are activated by the EGFR family are the phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K), mitogen-activated protein kinase (MAPK) and protein kinase C (PKC) (Figure 1. 26).[178, 179]
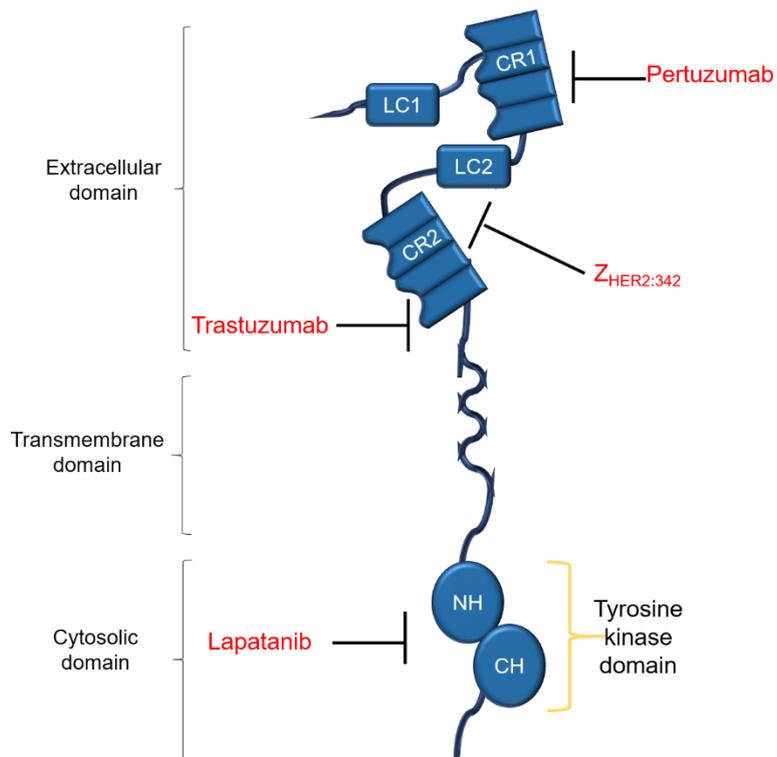
Figure 1.25 A schematic outline of the HER2 receptor. HER2 is comprised of three regions: an extracellular domain, a transmembrane domain, and the cytosolic domain. The extracellular domain is composed of four domains: Two leucine-rich domains (LC1 and LC2) and two cysteine-rich domains (CR1 and CR2). The cytosolic domain is composed of the amino (NH) and carboxyl (CH) lobes, which form the tyrosine kinase domain. HER2-targeted therapies which are close to the domains they target are in red.
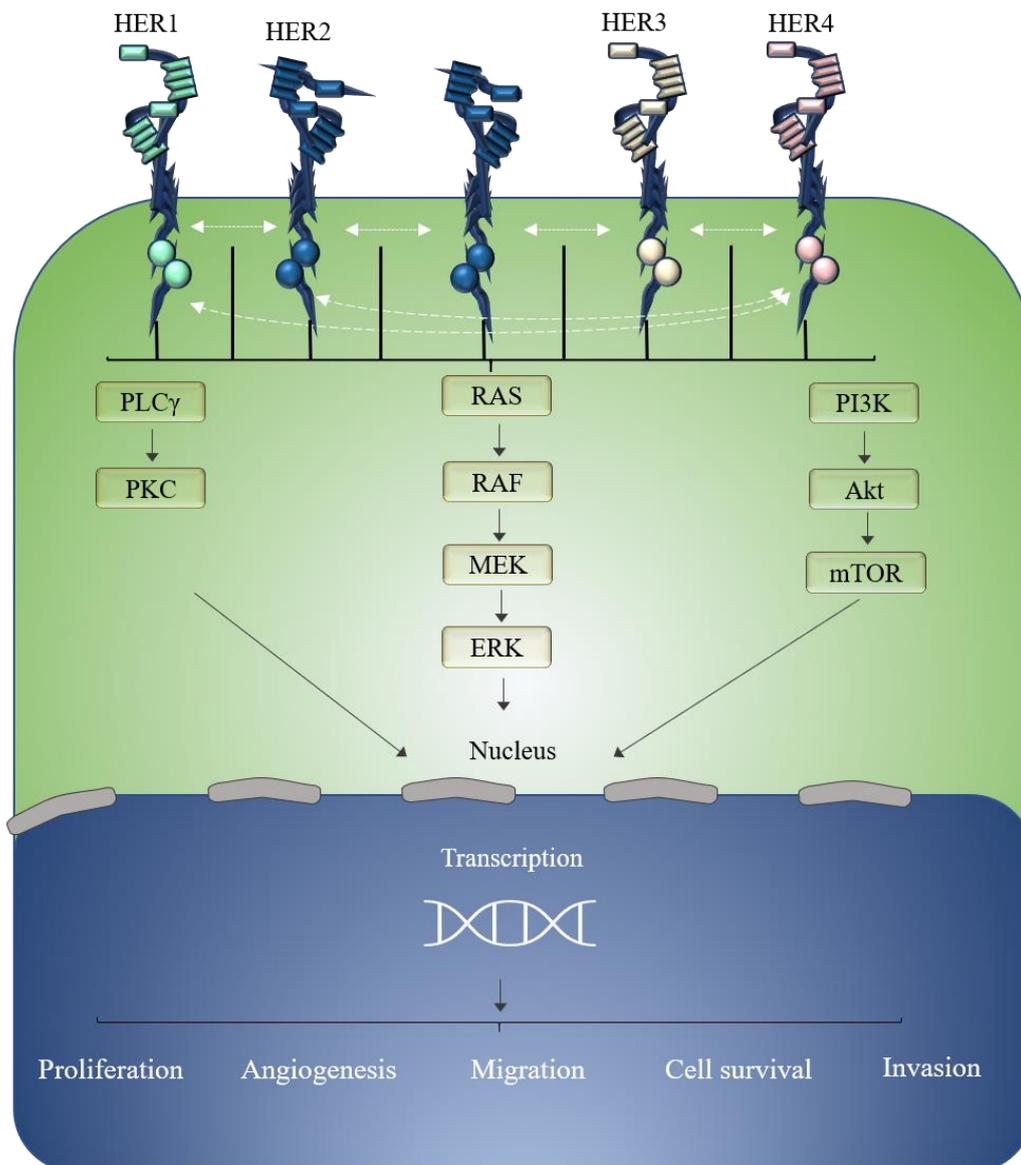
Figure 1.26 Schematic representation of EGFR family dimerisation and the different signal transduction pathways. EGFRs homo- or heterodimerise, activating various different downstream signalling pathways. Receptor dimerisation leads to transcription of one or more genes involved in proliferation, angiogenesis, migration, cell survival or invasion in cancer cells.

As HER2 is overexpressed in ~1/3 of invasive breast cancers, strategies have been developed to target HER2 and inhibit downstream signalling including humanised recombinant monoclonal antibodies, small molecules and affibodies such as trastuzumab (Herceptin) and pertuzumab, lapatinib and $Z_{HER2:342}$, respectively.[180] Trastuzumab, a recombinant humanised

monoclonal antibody, was the first directed therapy against HER2. It targets the cysteine-rich domain 2 (CR2) of HER2 and its main mechanism of action is through antibody-dependent cellular cytotoxicity (ADCC) by engaging with the Fc receptor of immune cells.[181] Trastuzumab has also been shown to inhibit HER2/HER3 ligand-independent dimerisation in HER2-overexpressing cells (Figure 1.27a,b). The downstream effects of this include inhibition on the MAPK and PI3K/Akt signalling pathway, leading to cell-cycle arrest and apoptosis, thereby reducing tumour growth and survival.[182] Studies have shown that trastuzumab also has anti-angiogenic effects, causing a reduction in micro vessel density and size in PT1590 (esophageal adenocarcinoma)-xenograft primary tumours and human mammary carcinomas.[183, 184] Although trastuzumab has greatly aided therapy against HER2+ cells, only ~30% of patients (women with HER2-overexpressing metastatic breast cancer) respond to trastuzumab therapy, indicating intrinsic resistance to the drug.[185] In addition, it has been shown that even patients who respond well, still experience disease progression after ~12 months, indicating a secondary mechanism of resistance.[186] Potential mechanisms for resistance include changes in the HER2 signalling pathway including mutations in PI3K, which result in constitutive AKT phosphorylation or mutations the tumour suppressor gene, PTEN. As this acquired resistance does not affect trastuzumab binding, it allows the HER2 receptor to be used as a marker for cytotoxic therapies including delivery of an antibody–drug conjugate, trastuzumab emtansine (T-DM1, Kadcyla®). T-DMI is composed of trastuzumab, which allows targeted delivery and uptake via HER2 expressing cells, linked to a highly potent natural-product-derived microtubule destabilising agent, with a drug to antibody ratio of 3.5:1.[187, 188] Once inside the cell, the DM1 component binds to tubulin, promotes microtubule depolymerisation and inhibits microtubule polymerisation. This leads to cell-cycle arrest in metaphase and apoptosis in breast and gastric cancer cells.[189, 187] However, this type of cytotoxic therapy relies on trastuzumab binding to the HER2 receptor and might not work where structural changes to HER2 prevent trastuzumab from binding. In this case, various other drugs have been investigated to overcome trastuzumab resistance including lapatinib, an EGFR/ErbB2/HER2 tyrosine kinase inhibitor, gefitinib, an EGFR/HER1 tyrosine kinase inhibitor and affibodies such as $Z_{HER2:342}$.[49, 50]
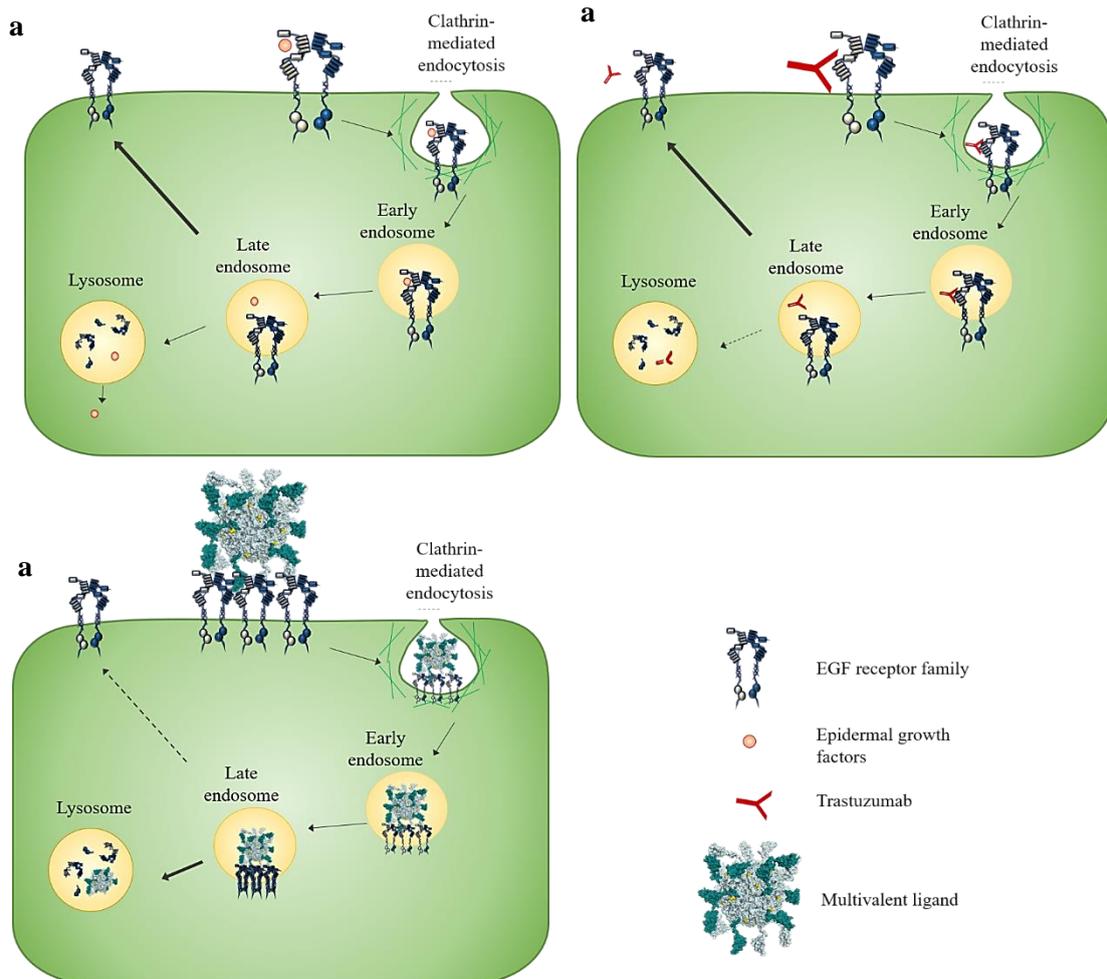
Figure 1.27 Different internalisation routes activated by different ligands. **a**, Ligand: Natural binding HER2/HER3 ligands. Upon ligand binding, the HER2/HER3 complex is activated and undergoes clathrin-mediated endocytosis. The ligand/receptor complex is then transferred from the early endosome to the late endosome. Under the acidic conditions (pH 4-5) of the late endosome, the ligand detaches, allowing the receptors to be predominantly recycled back to the cell surface membrane or degraded in the lysosome. **b**, Ligand: trastuzumab. The endocytosis route for trastuzumab and the HER2/HER3 receptor is similar to the natural ligand, whereby the receptor recycling route is dominant. **c**, Ligand: multivalent ligand. Binding of a multivalent ligand to multiple receptors induces HER2/HER3 receptor oligomerisation. The oligomerisation of receptors leads to degradation route being the dominant internalisation pathway. Bold arrows (major route) and dotted arrows (minor route).

1.5.2 Affibodies

To overcome drug-resistance including intrinsic and secondary tratuzumab resistance, new therapeutics are being investigated. One of these therapies are affibodies, a new class of ligands formed from the Z-domain in the immunoglobulin-binding region of *Staphylococcus aureus* protein A. They consist of 58 AAs (5.7 kDa) that fold into a three-helix bundle (Figure 1.28). Mutations to 13 solvent-accessible surface residues on helix one and two allow a large library of affibody constructs to be generated with diverse targets, including EGFR/HER1, HER2 and HER3.[192–195] They have also attracted considerable attention due to their high solubility, thermal stability (T_m: the temperature where the ratio of unfolded to folded proteins is equal) (T_m 65-75 °C) and pM to µM affinities, making them promising candidates for medical applications including imaging, receptor quantification and as therapeutic agents.[196–200] One of the first affibodies used for imaging was ^{111}In-labelled $Z_{HER2:342}$ for the detection of HER2 expression in ovarian cancers ^{111}In-labelled $Z_{HER2:342}$ demonstrated good receptor specificity, with HER2 binding being blocked by non-labelled $Z_{HER2:342}$ and good tumour uptake, with only marginal amounts being detected in other tissues including the heart and liver.[201] However, high amounts of ^{111}In-labelled $Z_{HER2:342}$ were also seen in the kidneys 1-48 hrs after injection. This accumulation is expected as proteins with sizes between ~2-50 KDa are usually filtered by the kidneys.[202] To overcome accumulation in the kidneys, proteins such as serum albumin have been fused to $Z_{HER2:342}$, resulting in higher biodistribution in the tumours compared to the kidneys.[203]
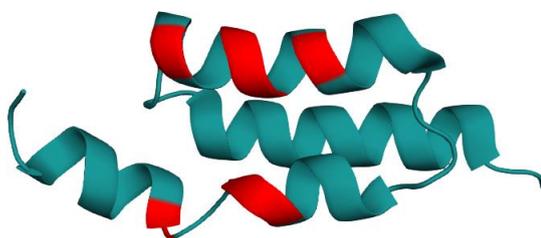
Figure 1.28 Structure of $Z_{HER2:342}$. Red segments highlight the notable contact areas of $Z_{HER2:342}$ t with HER2: Arginine 10, Tyrosine 13, Tryptophan 14, Arginine 28, Isoleucine 31, Arginine 32 and Tyrosine 35.[191] PDB 2KZI

To further increase their therapeutic effect, affibodies have also been engineered to other NCs including apoFt. Schardt *et al.,* 2017 engineered a multivalent HER3 Z_{05413} affibody by fusing EGFR-targeting affibodies onto the N-terminus of huH, thereby increasing the valency number. This resulted in increased efficacy via affibody-mediated HER3 downregulation in several cancer cell lines including breast cancer BT474 cells.[204] One possible explanation for the increased efficacy is because multivalent binding has been shown to induce rapid HER2 internalisation and degradation of multivalent drugs compared to trastuzumab alone (Figure 1.27c).[205] The EGFR affibody-apoFt construct also demonstrated high biodistribution in the tumour and liver.[206] However, the high retention of both huH and huH-EGFR-targeting affibodies in the liver could be due to the high presence of TfR1 receptors on hepatocytes or elimination of apoFt from the body via the liver.[66, 67] Given that affibodies offer such beneficial characteristics, they are at the fore-front of research-based therapeutic imaging. However, it should be noted that given their origins from bacteria, their ability to elicit a high immune response may be possible.

Aims and objectives of the research

The general aim of this PhD project was to further develop huH as a nanocarrier for the targeted treatment of breast cancers. To achieve this, the project was divided into two objectives: To engineer huH for targeted delivery towards HER2-overexpressing cancer cells and to engineer huH to deliver a wider array of therapeutic pay loads, including apoptotic proteins.

The initial step of the first objective was to ensure that huH was suitable for encapsulating and delivering payloads. This was achieved by identifying and mutating the residues involved in ferroxidation. The primary reason for removing huH's ferroxidase activity was to reduce the possibility of the production of hydroxyl species, an intermediate species that is produced during ferroxidation. Due to their highly oxidative nature, the presence of hydroxyl species within the core could negatively affect any future organic payloads. Following this, the focus moved to huH's targeting capabilities. Although huH had been shown to bind to TfR1, a receptor expressed in normal cells and overexpressed in cancer cells, including breast cancers, the amino acids on huH involved in TfR1 binding have not been identified. Therefore, potential residues on the exterior of huH based on sequence comparison between huL and huH were identified and removed, to produce the TfR- mutant. In addition, the histidine purification tag (HIS-tag) on huH was also evaluated for its effect on uptake. The primary aim for investigating the HIS-tag was to confirm whether the presence of the positively charged polyhistidine tag would hinder or promote huH uptake. However, results from the HIS-tag uptake studies indicated no negative or positive effect on apoFt uptake in SKBR3 cells. A single cysteine mutant was introduced for fluorescent labelling and *in vitro* tracking and quantification of huH and TfR- in breast cancer cells. The effects of both mutations were investigated in SKBR3 cells using confocal microscopy. In addition to removing the TfR1 binding site, the effect of dual TfR1/HER2 uptake by a huH-HER2 engineered moiety, Afb, was also investigated using *in silico* analysis. Although the addition of an Afb would enable huH uptake via HER2, which is overexpressed in many breast cancers, there would also be a risk of secondary uptake via the TfR1. Therefore, it was essential to investigate if the Afbs on the surface of huH would be large enough to block TfR1 binding. Homology modelling predicted that the presence of multiple Afbs on the surface of huH could prevent TfR1 binding. This hypothesis was evaluated by flow cytometry using breast cancer cells SKBR3, MDA-MB-231, MDA-MDA-468 and BT474, expressing a range TfR1/HER2.

The second objective of this study was to evaluate the capacity of huH to encapsulate and deliver peptides/proteins to breast cancer cells. To date, there have not been any therapeutic

proteins that have been encapsulated within huH. This is possibly due to huH's small cavity, 8 nm, and its extreme disassembly/re-assembly conditions. However, given that therapeutic proteins have a high target specificity compared to compounds, it was important to identify a protein that could be encapsulated within huH's small core and be cytotoxic towards breast cancer cells. Cyt C, a mitochondrial protein, was selected as it can induce apoptosis in breast cancer cells. Initial experiments investigated huH's ability to encapsulate Cyt C using the urea- and the pH- disassembly methods. Due to the acidic condition (pH 2-3) of the pH-based encapsulation method which could affect the structure and characteristics of the encapsulated protein, HuH was engineered to disassemble at a higher pH (ΔDE). Native PAGE and dynamic light scattering were then used to identify the new variant's pH disassembly point. Once the new pH disassembly point was identified, ΔDE's capacity to encapsulate Cyt C was evaluated using native PAGE and UV-vis.-spectroscopy. Further *in vitro* experiments were thwarted due to the COVID-19 national lockdown. Therefore, basic protein modelling was undertaken to investigate the structural changes on the 3- and 4- fold channel of ΔDE. Finally, ΔDE's capacity to deliver Cyt C to TfR1 expressing cells and Cyt C's ability to inhibit cell proliferation were evaluated using MTT assays.

2.0 Materials and Methods

All reagents and materials were purchased from MilliporeSigma owned by Merck (Previously known as Sigma Aldrich). They were used as received without further purification unless stated otherwise.

2.1 Buffers

All buffers and solutions were prepared in the laboratory. The compositions of all the buffers used are summarised in Table 2.1.

Buffer	Composition
Conjugation buffer	8 M Urea, 20 mM Tris pH 7.2, 5-carboxyfluorescein maleimide (5-FAM) (10 times the molar concentration of individual apoFt subunits in the sample).
LB	5 g of Luria-Bertani broth was dissolved in 200 mL of de-ionised water. The solution was then autoclaved to sterilise it.
Mass spectrometry	
Elution solution	Saturated solution of sinapic acid dissolved in 0.1% v/v trifluoroacetic acid (TFA) with 50:50 v/v acetonitrile (ACN): H_2O
Wash solution /Equilibration solution:	0.1% v/v TFA in H_2O
Wetting solution:	50:50 v/v ACN: H_2O
In vitro **studies**	
Antibody buffer:	1% v/v FBS in PBS containing a final concentration of 66 mM anti-TfR1 antibody (transferrin receptor 1 monoclonal antibody (T56/14), PE) or anti-HER2 antibody (Human epidermal growth factor 2 monoclonal antibody (9G1D4B1), PE).

Blocking buffer:	1% v/v Foetal bovine serum (FBS) in PBS.
Fixation solution:	3.7% v/v Formaldehyde in phosphate-buffered saline (PBS)
MTT solution:	2 mg/ mL of MTT in dH_2O
Trypsin buffer:	10x solution of Trypsin- Ethylenediaminetetraacetic acid (EDTA) solution was diluted in PBS in a 1:10 ratio.
Protein Purification	
IMAC binding buffer	20 mM Tris pH 8.0, 10 mM imidazole (+ 8 M urea for insoluble fractions)
Immobilised metal affinity chromatography (IMAC) elution buffer	20 mM Tris pH 8.0, 1M imidazole (+ 8 M urea for insoluble fractions)
Ion exchange buffer	20 mM Tris pH 7.2, 150 mM NaCl
Ion exchange elution buffer:	20 mM Tris pH 7.2, 1000 mM NaCl
SDS loading buffer:	50 mM Tris pH 6.8, 100 mM DTT, 2% w/v SDS, 10% v/v glycerol, 0.1% w/v bromophenol blue.
1x SDS running buffer	25 mM Tris, 250 mM glycine, 0.1% w/v SDS
1x Native loading buffer	62.5 mM Tris-HCl, pH 6.8, 50% v/v glycerol, 0.1% w/v bromophenol blue.

Table 2.1 The composition of buffers and solutions used in their respective experiments.

2.2 Bacterial strains

Details of bacterial strains used to express the protein of interest are given in Table 2.2. The cells lines in Table 2.2 were used to either store or express the protein of interest. BL21(DE3) *E.coli* cells were used to express proteins. This cell type carries genes that enable protein expression of recombinant protein using the bacteriophage T7 promoter-based expression systems. These include F- (Does not carry the F plasmid, for recombination events), ompT (mutation in protein protease 7), gal (galactose non-utilising), lon (mutation in lon protease), λ(DE3) (encodes the T7 RNA polymerase). DH5α *E.coli* cells were used as a means to stably clone and store protein expression vectors. Genes encoded in this cell line include endA1 (mutation in endonuclease I), gyrA96 (mutation in DNA gyrase), recA1 (mutation in proteins involved *E.coli* recombination and fhuA2 (resistant to phage T1).

Cell Line	Genotype	Application
BL21(DE3)	F⁻, ompT, gal, lon and λ(DE3)	Protein expression
DH5α	endA1, gyrA96, recA1 and fhuA2	Storage

Table 2.2. Bacterial strains and their genotype

To allow expression of a desired protein, the gene encoding the protein was cloned into vectors containing the T7 promoter (Table 2.3). All genes were codon optimised for expression in BL21(DE3) *E.coli.* Gene sequences for each protein are given in SFigure 3.2.

Vector	Gene	Promoter and antibiotic resistance	Source
PJexpress414	huHC- (Human heavy chain cysteine free) SFigure 3.2.	T7 and ampicillin	DNA2.0 (CA,USA)
	huH 222 (Human heavy chain cysteine free with ferroxidase		

	mutation)
	huH 222 TfR- (Human heavy chain cysteine free with ferroxidase mutation and transferrin receptor 1 mutations)
	ΔDE huH 222 TfR- (Human heavy chain cysteine free with the last 23 AA removed, ferroxidase mutation and transferrin receptor 1 mutations)

Table 2.3 Characteristics of plasmid vector and the gene inserted.

2.3 Plasmid vector

2.4 Plasmid quantification

LB (10 mL), supplemented with ampicillin at a final concentration of 100 μg/ mL, was inoculated with a DH5α *E.coli* containing the respective plasmid vector (Table 2.3). Overnight 10 mL *E.coli* cultures were harvested at 1200 *g* for 20 mins, and the supernatant was discarded. A Wizard® Plus SV Miniprep DNA Purification kit (Promega) was then used to extract the plasmid according to the manufacturer's protocol and eluted in a final volume of 50 μL nuclease-free water. Plasmid concentration was quantified using NanoDrop® A-1000 and stored at -20 °C.

2.5 DNA Mutagenesis

Site-directed mutagenesis was performed using a New England Biolabs (NEB) Q5® Site-directed mutagenesis kit, according to the manufacturer's protocol. This method allowed rapid, site-specific mutagenesis of double-stranded plasmid DNA. Mutagenic primers were created

using NEBaseChanger® software available at https://nebasechanger.neb.com/ and purchased from Merck (Formerly known as Sigma-Aldrich). The primers are listed in Table 2.4. Briefly, 15.5 μL of Q5 hot start high-fidelity 2x master mix, 1.5 μL of 10 μM forward primer, 1.5 μL of 10 μM reverse primer, 1 μL of template DNA (15 ng/ μL) and 9.0 μL of nuclease-free water were loaded and mixed into a thin-walled PCR tube. The standard PCR reaction cycling conditions are described in Table 2.5.

Primer name	Sequence 5' to 3'
pJEXP414_Fwd	TAATACGACTCACTATAGGGG
pJEXP414_Rev	CTCAAGACCCGTTTAGAGGC
Ferroxidase mutant primers	
1st huH 222_Fwd	CGCGCGCACGCCGAGAAACTGAT
1st huH 222_Rev	TTTTGCATGGGACTGGTGCAGGAA
2nd huH 222_Fwd	CCGCGAAACTGATGAAACTGCAGAACCAG
2nd huH 222_Rev	CGCCCGCGCGTTTTGCATGGGA
Amber mutant 160	
Δ 160_Fwd	GCGTAAGATGGGTTAACCGGAGTCCG
Δ160_REV	AGGTTCGTAACATGGTCGC
Transferrin receptor mutant	
1st TfR-_Fwd	TGTTGCAAGACATTAAGAAACCGGATCGC
1st TfR-_Rev	GCGCACGACCGCCACGCTGGTT
2nd TfR-_Fwd	CAAGACATTAAGAAACCGGATCGCG
2nd TfR-_Rev	AAACAGCGCACGACCGCC
K68C mutant	
huHC- K68C_Fwd	ATGCAATGTGAACCAAAGCTTGCTGG
huHC- K68C_Rev	TCCAAGTGCAGGGCAGCTTCCA
TfR- K68C_Fwd	GGGCGCCGCGTGCCTGATGAAACTG
TfR- K68C_Rev	GCGCGTTTTGCATGGGAC

Table 2.4 Primers for mutagenesis as designed by NEB base changer

Step	Temp	Time
Initial denaturation	98 °C	30 seconds
25 cycles	98 °C	10 seconds
	50-72 °C *	30 seconds
	72 °C	30 seconds /kb
Final extension	72 °C	2 mins
Hold	4 °C	

Table 2.5 PCR cycling condition according to New England Biolab Q5® Site-Directed mutagenesis. * Used T_a of mutagenic primers designed in NEB base changer as shown in **Table 2.4.**

Following PCR, the DNA product was carried forward for phosphorylation by kinases, DNA intramolecular circularisation by ligases and removal of the original template by DpnI. This process was performed by an enzyme mix (kinase, Ligase and DpnI) KLD (New England Biolab), and the reaction was performed according to the manufacturer's protocol. Briefly, 1 µL of the PCR product, 5 µL of the 2x KLD reaction buffer, 1 µL 10x KLD enzyme mix and 3 µL of nuclease-free water were mixed and incubated for 30 mins at RT. Following this, transformation was performed using 5 µL of the final KLD mixture and 50 µL of chemically competent DH5α *E.coli* cells, according to section 2.6.

2.6 Transformation for chemically competent *E. coli*

Plasmids were transformed into *E.coli* cells according to the protocol described in Sambrook *et al.*, 2001.[69] Briefly, aliquots of either BL21 (DE3) *E.coli* or DH5α *E.coli* (Table 2.2) were thawed on ice, and 5 µL of 50-100 ng of DNA plasmid were added to 50 µL of cell suspension, mixed gently and left on ice for 30 mins. Cells were then heat shocked at 42 °C for 30 secs and then quickly placed on ice for 5 mins. Next, cells were added to 950 µL of preheated super optimal broth medium (New England Biolab) (37 °C) in 15 mL falcon tubes and grown at 37 °C for 1 hr with 180 rpm shaking to allow cell recovery. Finally, the cell suspension (70 µL) was spread on a 37 °C heated LB agar ampicillin selection plate and incubated at 37 °C for 12-16 hrs.

2.7 Protein purification

2.7.1 Small scale expression

LB (10 mL), supplemented with ampicillin at a final concentration of 100 µg/ mL, were inoculated with BL21 (DE3) *E.coli* cells containing the respective plasmid vector (Table 2.3) and grown overnight at 37 °C, shaking at 180 rpm. Overnight cultures (500 µL) were added to 250 mL flasks containing 50 mL of LB with ampicillin at a final working concentration of 100 µg/ mL. Cultures were grown at 37 °C and shaken at 180 rpm until the cells reached OD_{600} 0.6-0.8. After OD_{600} 0.6-0.8 was reached, the LB was divided into two 125 mL flasks with 2 x 25 mL cultures. Flask one was induced with a final concentration of 1 mM isopropyl β- D-1-thiogalactopyranoside (IPTG) to promote protein expression. The second flask was cultured in non-inducing conditions as a control. Both flasks were cultured at 37 °C for 3 hrs or overnight at 20 °C. Following expression, 1 mL sample suspension was measured at OD_{600} to obtain the cell density and determine the cell suspension volume required for SDS-page analysis. For example, if 1 mL of cell suspension produced a readout of OD_{600} 2.4, then using the following equation $1000/2.4 = 417$, 417 µL of cell suspension would be taken forward for analysis. Cell suspensions were pelleted at 13,000 *g* for 1 min, and the supernatant was discarded. Pellets were resuspended in 200 µL of cell lytic buffer containing 1 µL of lysozyme (50 mg/ mL), 0.5 µL benzonase (Sigma), and 3 µL of 1 M $MgCl_2$. Induced cell pellets were further centrifuged at 13,000 *g* for 10 mins to separate the soluble and insoluble fractions, and the supernatants (soluble) were extracted into a new tube. The cell pellets (insoluble) were resuspended in 200 µL of 20 mM Tris pH 8.0. The samples were then analysed by SDS-PAGE for protein overexpression, according to section 2.7.2.1.

2.7.2 Polyacrylamide gel electrophoresis

2.7.2.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Mini (0.7 mm thick gels were cast using Biorad's Mini-PROTEAN spacer plates) 15% resolving and 6% stacking gels were prepared according to the quantities stated in Table 2.6. The gel was assembled in Biorad™ electrophoresis cells according to the manufacturer's instructions. Protein samples were mixed with 4x SDS loading buffer (Table 2.1) in a 1:4 v/v ratio, vortexed briefly and denatured at 95 °C for 5 mins. Subsequently, samples were briefly

centrifuged, and 6 µL of each sample was loaded onto the gel, as well as 3 µL of pre-stained protein ladder (Thermo Scientific™ PageRuler™). The system was buffered with a 1x SDS running buffer (Table 2.1). Electrophoresis was initially run for 10 mins at 180 V, followed by 45 mins at 200 V to allow protein separation. After electrophoresis, the gel was washed in deionised water 3x for 5 mins at RT to remove the presence of SDS. Next, the gel was stained with activated GelCode™ Blue Safe protein stain for 1 hr at RT. Following this, the gel was de-stained in deionised water 3x for 5 min at RT and then overnight at RT. The de-stained gel was visualised using the Bio-Rad gel imaging system.

	Volume (mL)	
	15% w/v resolving gel	6% w/v stacking gel
dH$_2$0	1.1	2.6
30% w/v protogel acrylamide	2.5	1.0
Tris	1.3 (1.5 M, pH 8.8)	1.6 (1.0 M, pH 6.8)
10% w/v ammonium persulphate	0.05	0.05
10% w/v SDS	0.05	0.05
100% v/v Temed	0.002	0.004

Table 2.6 Recipes 15% resolving and 6% stacking for SDS-page.

2.7.2.2 Native polyacrylamide gel electrophoresis

ApoFt variants were dialysed against buffers of varying pH ranges (50 mM Tris pH 7-9, 50 mM sodium acetate pH 3-6 and 50 mM potassium chloride pH 2-3) in a 6-8 kDa molecular weight cut-off (MWCO) dialysis bag at 4 ^{o}C for 24 hrs. The next day, the samples were removed from the dialysis bag, centrifuged at 13,000 g for 10 mins, and the supernatant containing the

protein was quantified using NanoDrop® A-1000. Samples were mixed with native PAGE sample buffer (Table 2.1) in a 1:1 v/v ratio. After mixing, 10 μL of 0.7-1 mg/ mL protein samples and 5 μL of NativeMark™ unstained protein standard were quickly mixed and loaded onto a mini SDS plate (0.7 mm thick gels cast using Biorad's Mini-PROTEAN spacer plates). Native PAGE was run for 2 hrs at 100 V. After electrophoresis, the gel was stained with activated GelCode™ Blue Safe Protein Stain for 1 hr. Following this, the gel was de-stained using deionised water 3x for 5 mins at RT and then overnight at RT. The de-stained gel was visualised using the Bio-Rad gel imaging system.

2.7.3 Protein overexpression

A starter culture of 10 mL LB, supplemented with ampicillin at a final concentration of 100 μg/ mL, was inoculated with BL21 (DE3) *E.coli* cells containing the respective plasmid vector (Table 2.3). The culture was grown overnight for 14-18 hrs at 37 °C, shaking at 180 rpm. The starter culture was then sub-cultured into 1 L of LB, containing the ampicillin at a final concentration of 100 μg/ mL. The culture was grown for 2-3 hrs at 37 °C, shaking at 180 rpm until the cell density reached OD_{600} 0.6-0.8. Following this, IPTG was added at a final concentration of 1 mM for 3 hrs at 37 °C or 0.5 mM for overnight expression at 20 °C, with shaking at 180 rpm for both conditions. Following expression, the culture was transferred to 400 mL centrifuge tubes and pelleted at 1,200 *g* for 20 mins. Finally, the supernatant was decanted, and the pellet was resuspended in either 50 mL IMAC binding buffer or ion exchange buffer (Table 2.1) and stored at -20 °C for up to one week.

2.7.4 Preparation of *E. coli* cell lysate

Following the procedure outlined in section 2.7.3, cell suspensions were thawed on ice. Next, the cell suspension was transferred to a 50 mL tube and phenylmethylsulfonyl fluoride was added at a final concentration of 1 mM. The cell suspension was then sonicated (SANYO) at 15 mA with 1 min off and 1 min on for 12 mins at 4 °C. Following sonication, the cell suspension was transferred to centrifuge tubes and centrifuged at 35,000 *g* for 30 mins to allow separation of the soluble and insoluble fraction. Following centrifugation, the soluble and insoluble fractions were separated, and the insoluble fraction was resuspended in IMAC binding

buffer containing 8 M urea and 1 mM DTT (Table 2.1) for protein constructs containing cysteine mutations. Finally, fractions were filtered with a 0.22 μm filter before being carried forward for immobilised metal affinity chromatography (section 2.7.5).

2.7.5 Immobilised metal affinity chromatography

An ÄKTA prime system (GE Healthcare life science, Amersham) and a 5 mL column of HiTrap chelating HP affinity resin charged with 50 mM $NiSO_4$ solution were used to purify proteins containing a 6x-Histidine tag. The column was equilibrated with IMAC binding buffer (Table 2.1) before the soluble/insoluble fraction was loaded onto the column. The whole purification was run at 5 mL/ min. Once the protein solution was loaded onto the column, non-specific protein binding was eluted with 10% v/v IMAC elution buffer (Table 2.1) until the UV_{280} signal reached zero. Next, the desired protein was eluted with 60% v/v IMAC elution buffer and collected in 6 mL fractions. The soluble, insoluble, flow-through and protein samples were collected for SDS-page analysis (section 2.7.2.1). Once SDS-page confirmed the presence of the desired protein in the appropriate fraction, the respective fractions were pooled and dialysed against 2 L of 20 mM Tris pH 8.0 buffer overnight at 4 °C. Samples were then concentrated using a VIVAspin column 10 K MWCO, flash-frozen in liquid nitrogen and stored at -80 °C.

2.7.6 Ammonium sulphate precipitation and gel filtration

Ammonium sulphate: Ammonium sulphate precipitation was performed for proteins that require concentration or further purification. A saturated ammonium sulphate solution (4.06 M at 20 °C) was gradually added to protein solutions to give a final ammonium sulphate concentration at 70% v/v. The protein solution was then incubated for 30 mins on a rotor to precipitate proteins. Precipitated proteins were centrifuged at 12,000 *g* for 10 mins, and the supernatant was discarded. For proteins requiring further purification, the pellet was resuspended in 7 mL of ion-exchange buffer (Table 2.1), loaded onto a super loop injection device and carried forward for gel filtration.

Gel filtration: A 16/20 S300 GL column (GE Healthcare) was pre-equilibrated overnight with ion-exchange buffer (Table 2.1) at a flow rate of 1.0 mL/ min. The protein solution was loaded

onto the column via a super loop at a flow rate of 1.0 mL/ min. Following this, the purification was performed at 1.0 mL/ min. After an initial 68 mL elution which was directed to waste, samples were collected at 7 mL per fraction. Samples were run on SDS-page (section 2.7.2.1) to verify the fractions that contained the desired protein. Fractions containing the desired protein were pooled, diluted 50% v/v in 20 mM Tris pH 7.2 and incubated for 30 mins at RT before being further purified with a Hi-Trap Q HP column.

Anion ion exchange: The Hi-Trap Q HP column (Q-Sepharose High-Performance GE Healthcare, Boston, USA) was pre-equilibrated in ion-exchange buffer (section 2.1) at 5 mL/ min. Fractions from gel filtration were loaded on the column at 2.5 mL/ min. Non-specific proteins were eluted using 10% v/v ratio of ion-exchange elution buffer at 5 mL/ min until the UV_{280} reached zero. The desired protein was eluted using 30% v/v ion exchange elution buffer at 5 mL/ min in 7 mL fractions. Fractions were run on an SDS page gel to confirm the presence of the protein in the eluted fractions. The desired fractions were pooled and dialysed against 2 L of the ion exchange buffer overnight at 4 °C. The following day, samples were concentrated using a VIVAspin column 10K MWCO, flash-frozen in liquid nitrogen and stored at -80 °C.

2.7.7 Δ160 expression

A stop codon at A160 was introduced using DNA mutagenesis and was performed according to the protocol described in section 2.5. The final plasmid was transformed into DH5α *E.coli* for storage and BL21 *E.coli* for expression. The construct was then expressed according to section 2.7.3.

2.7.8 E-helix truncated (ΔDE) expression

Using HuHC-'s protein sequence as a template, the following modifications were applied to the new apoFt construct; an E-helix truncation, a ferroxidase (huH 222) mutation and a transferrin receptor site (TfR-) mutation. The N-terminal histidine-tag and avidin-tag were also removed. The final construct was termed ΔDE. The gene was codon-optimised for expression in BL21 (DE3) *E.coli* and cloned into a PJexpress414 vector with an ampicillin resistance gene by ATUM (formerly known as DNA 2.0 (DNA 2.0 CA, USA)). The construct was cloned into

both DH5α *E.coli* for storage and BL21 *E.coli* for expression. The BL21 *E.coli* glycerol stock was inoculated into a 10 mL LB starter culture with the ampicillin at a final working concentration (in medium) of 100 μg/ mL. The culture was then expressed according to section 2.7.3 and purified according to 2.7.6.

2.8 Circular Dichroism

To measure the secondary structure of different apoFt variants, circular dichroism (CD) measurements were performed using an Applied Photophysics Chirascan-Plus fitted with a Quantum Northwest temperature controller (JASCO UK). Purified proteins were filtered using a 0.22 μm filter to remove large aggregates and then diluted to 4 μM. The spectra were recorded using four accumulations in the UV range between 200 nm and 250 nm at 25 °C. The cuvette path length used was 10 mm.

2.9 Mass spectrometry

ESI-mass spectrometry (MS) was employed to analyse conjugation efficiency (section 2.14) and the MW of apoFt subunits. Each protein sample was diluted to a final concentration of 30 μM with 20 mM Tris pH 8.0. A C4 ZipTip® Pipette tip was used to prepare samples for ESI-mass spectrometry. The ZipTip was cleaned using wetting solution (section 2.1) with six repeat aspirations and waste removal steps. Following this, the ZipTip was pre-equilibrated with wash buffer (section 2.1). The ZipTip was then used to aspirate 10 μL of 30 μM protein. Repeat aspiration and dispersal of the protein solution was performed 15x to saturate all the binding sites in the ZipTip. The bound protein was then washed with wash buffer 6x. Finally, the protein was eluted using 4 μL of elution buffer (section 2.1). Samples were analysed using MALDI-ToFMS Bruker ultraFlex III.

2.10 Dynamic light scattering

huHC- and ΔDE were dialysed against buffers of varying pH ranges (50 mM Tris pH 7.5, 50 mM sodium acetate pH 6, 5 and 4, as well as 50 mM potassium chloride pH 2) in a 6-8 kDa MW cut-off dialysis bag at 4 °C for 24 hrs. The protein samples were then filtered with a 0.22

µm filter and loaded into a folded DTS1070 capillary zeta cell. Samples were analysed at 0.4 mg/ mL using a Zetasizer Nano ZS for intensity, volume, and number. The samples were run according to the following parameters, material: protein (RI: 1.45), dispersant: water (25 °C, RI: 1.33). 11 runs per sample with a three-second delay.

2.11 *In vitro* assays

2.11.1 Cell culture

Culture conditions: Human cell lines were cultured in 5 mL of suitable cell culture medium supplemented with 10% v/v heat-inactivated foetal bovine serum (FBS). Cultures were incubated in a 25 cm^2 cell culture flask at 37 °C in a humidified atmosphere of 5% CO_2. Cells were grown at 70% confluency before being sub-cultured twice weekly to maintain logarithmic growth.

Subculture: Once cultures reached 70% confluency, the flasks were washed twice with 5 mL PBS. Following this, 1 mL of 1x trypsin solution (2.5% v/v in solution PBS of 1:250 trypsin) was added and incubated for 5-10 mins at 37 °C in a humidified atmosphere of 5% CO_2. Finally, trypsin was inactivated in complete culture medium, and 1:5 of the cell suspension was transferred into a new flask containing 5 mL culture medium.

Cell lines: All cell lines were sourced from the American Type Tissue Collection (ATCC).

- BT474 (Breast Ductal Carcinoma, ATCC No: HTB-20™)
- SKBR3 (Breast Adenocarcinoma, ATCC No: HTB-30™)
- MDA-MB-231 (Breast Adenocarcinoma, ATCC No: HTB-26™)
- MDA-MB-468 (breast adenocarcinoma, ATCC No: HTB-132™).
- MRC-5 (foetal lung fibroblast, ATCC No: CCL-171™). This non-tumourigenic cell line was also cultured as a control.

Cell culture media:

Cell lines were maintained and grown according to the conditions outlined in Table 2.7.

Cell line	Culture medium	Supplementation
BT474	RPMI-1640 medium	10% v/v FBS and 1% v/v penicillin/streptomycin
SKBR3		
MDA-MB-231		
MDA-MB-468	Minimum essential medium (MEM)	10% v/v FBS, 1% v/v penicillin/streptomycin and 2 mM L-glutamine.
MRC-5		10% v/v FBS, 1% v/v non-essential amino acids, 1% v/v penicillin/streptomycin, 2 mM L-glutamine and 10 mM Hepes buffer.

Table 2.7 Cell culture medium and supplementation for different cell lines.

2.11.2 Confocal imaging

Cells (SKBR3 or MRC-5) were seeded at a density of 1.5×10^4 cells/well in 200 µL of their respective growth media into eight-well µ-slides ibiTreat (Ibidi). Cells were allowed 24 hrs to adhere at 37 ºC in a humidified atmosphere of 5% CO_2. The medium was then aspirated, and the cells were washed twice with 200 µL of PBS. 200 µL of new growth media containing 100 µM 5-FAM-apoFts (huHC-, TfR-, both with and without HIS-tag (-HIS)) was added. Next, the cells were incubated for 1 hr at 37 ºC in a humidified atmosphere of 5% CO_2. After 1 hr incubation, cells were washed once with complete medium and left in 200 µL of respective

growth medium at 37 °C. Cells were imaged on a Zeiss LSM710 phase-contrast and fluorescent microscope. Gains were adjusted to match the fluorescence emission intensity of each sample. Images were acquired with x20 and x40 objectives, and excitation and emission filters were set at 488 nm and 520 nm, respectively. Confocal and phase-contrast channels were acquired in the same window.

2.11.3 Flow Cytometry

2.11.3.1 Human epidermal growth receptor 2 and transferrin receptor 1 quantification

Flow cytometry standard curve: BD Quantibrite[TM] beads (BD biosciences) and flow cytometry were employed to estimate the number of surface receptor-bound phycoerythrin-conjugated antibodies. The BD Quantibrite PE tube was reconstituted in ice-cold blocking buffer (section 2.1) and run on the Beckman Coulter FC50 using the excitation channel at 566 nm and emission channel at 574 nm. The linear regression of Log_{10} PE molecules per bead was calculated according to the manufacturer's instruction. Flowing Software (Turku Biosciences) was used to gate and analyse the data.

Receptor quantification: Cells were seeded at a density of 5 x 10^4 cells/well in 1 mL of complete media per 12-well plate. Cells were then allowed 24 hrs to adhere at 37 °C in a humidified atmosphere of 5% CO_2. After 24 hrs cells were washed with PBS and trypsinised with 1x trypsin buffer for 5 mins at 37 °C in a humidified atmosphere of 5% CO_2. The trypsin was inactivated with ice-cold respective cell growth media, and the cell suspensions were transferred to ice-cold fluorescence-activated cell sorting (FACS) tubes. After that, the cells were centrifuged at 300 *g* for 10 mins at 4 °C. The supernatant was decanted, and the pellet was resuspended in 1 mL of ice-cold PBS. The suspension was briefly vortexed and centrifuged at 300 *g* for 10 mins at 4 °C. The supernatant was decanted, and 1 mL of ice-cold blocking buffer was added (section 2.1). The suspension was briefly vortexed and left to incubate at 4 °C for 30 mins. The suspension was then centrifuged at 300 *g* for 10 mins at 4 °C. The supernatant was decanted, the cells resuspended in antibody buffer (section 2.1) and left to incubate for 45 mins at 4 °C while being protected from light. The suspension was then centrifuged at 300 *g*

for 10 mins at 4 °C. The supernatant was discarded, washed with ice-cold PBS, and centrifuged at 300 g for 10 mins at 4 °C. The supernatant was discarded, the cells resuspended in fixation solution (section 2.1) and incubated at RT for 10 mins. The suspension was then centrifuged at 300 g for 10 mins at 4 °C. The supernatant was discarded, and the cells were resuspended in 5 mL of ice-cold blocking buffer. Cell suspensions were analysed using Beckman Coulter FC500, using the excitation channel at 566 nm and emission channel at 574 nm, with a cell count of 10, 000 cells. Flowing Software (Turku Biosciences) was used to gate and analyse the data. An n=3 was performed for each internal replicate and N=3 for the number of independent experiments.

2.11.3.2 Relative apoferritin uptake

Cells (SKBR3, MDA-MB-231, MDA-MB-468, BT474 and MRC-5) were seeded at a density of 5 x 10^4 cells in 1 mL complete media per 12-well plate in their respective media and allowed 24 hrs to adhere at 37 ℃ in a humidified atmosphere of 5% CO_2. The media was then removed, washed twice with PBS and apoFt variants at a final concentration of 200 μM in complete media were added to the cells. The cells were incubated for 1 hr at 37 °C in a humidified atmosphere of 5% CO_2. The cells were then washed twice with ice-cold PBS and trypsinised for 5 mins at 37 °C in a humidified atmosphere of 5% CO_2. The addition of ice-cold complete medium inactivated trypsin and the cell suspension was collected into FACS tubes. The suspension was then centrifuged at 300 g for 10 mins at 4 °C, the supernatant decanted, and the cells washed with 5 mL of ice-cold PBS. The suspension was again centrifuged at 300 g for 10 mins at 4 °C and the supernatant decanted. The cells were resuspended in fixation solution (section 2.1) and left to incubate at RT for 10 mins. The suspension was then centrifuged at 300 g for 10 mins at 4 °C. The supernatant was discarded, and the cells were resuspended in 5 mL of ice-cold blocking buffer. Cells were analysed using a Beckman Coulter FC500, using the excitation channel at 488 nm and emission channel at 512 nm, with a cell count of 10,000 cells. Flowing Software (Turku Biosciences) was used to gate and analyse the data. An n=3 was performed for each internal replicate and N=3 for the number of independent experiments.

2.11.4 MTT assay

The MTT (3-(4,5-dimethylthiazol-2-yl)-2,5 diphenyltetrazolium bromide) assay was performed to measure inhibition of cell proliferation. The reduction of the tetrazolium

component of MTT into an insoluble formazan product by the mitochondria can reflect proliferation.

Day 1: Cells (SKBR3, MDA-MB-231 and MRC-5) were seeded into 96-well plates at a density of 3 x 10^3/well in a total volume of 180 μL of their respective growth media. The cells were left to adhere overnight at 37 ℃ in a humidified atmosphere of 5% CO_2.

Day 2: Time zero (T0) plates containing all three cell lines were treated with 50 μL of MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) at a final concentration of 400 μg/mL). The plates were then left to incubate for 3 hrs. During this period, the MTT is metabolised by dehydrogenases to form a blue insoluble formazan product. After the 3 hrs, the media was aspirated, and 150 μL of DMSO was added to each well to solubilise the formazan crystals. The T0 plates were shaken for 20 seconds, and the absorbance was monitored at 570 nm using the EnVision Multilabel Plate Readers (Perkin Elmer®). The T0 values represent absorbance at the time of test agent addition. Treatment plates were also prepared, by adding 20 μL of either ΔDE, encapsulated Cyt C or Cyt C only to the cells at concentrations ranging from 50-1000 nM and left to incubate for 72 hrs.

Day 5: Similar to the T0 plate, the experimental plates were treated with 50 μL of MTT solution, followed by 150 μL of DMSO and the data collected. Comparing the absorbance reading at 72 hrs to the T0 value can determine the inhibition of cell proliferation caused by the test agent. The test agent required to inhibit cell growth/proliferation by 50% (estimated GI_{50}) were calculated according to the equation below.

Data analysis: MTT was carried out twice, and the mean GI_{50} values were calculated using the following equations:

Abs. GI_{50}:

$$\left(\frac{\text{ControlAbs} - \text{T0 Abs}}{2}\right) + \text{T0Abs})$$

GI_{50}:

$$\left(\frac{\text{High Abswhere GI50 falls} - \text{Abs.GI}_{50}}{\text{Low Abswhere GI50 falls} - \text{Abs.GI}_{50}}\right) \times (\text{High Concentrationswhere GI}_{50} \text{ falls} -$$

Low Concentrationswhere GI_{50} falls) $+$ Low Concentrationswhere GI_{50} falls.

2.12 Modelling

2.12.1 Electrostatic potential of apoferritin's channels

To evaluate the electrostatic potential of apoFt variants at various pH points, basic modelling was performed using PyMol's inbuilt electrostatic calculator, APBS 2.0. PDBPPQR and CHARMM FORCE FIELDS were used to generate Van der Waals radii for huH subunits at the desired pH and optimise solvent calculations. PPB: 2FHA was used for huHC- and PDB: 5XB1 for ΔDE. PROPKA was used to assign protonation states at the selected pH, and the Debye-Hückel model was used to represent the behaviour of sodium and chloride ions surrounding the protein. The dielectric field for both the solvent and protein were set to $\Sigma s = 80.0$ and $\Sigma p = 4.0$, respectively.[1, 2] The solvent surface radius was set to 0 Å to correspond to a van der Waals surface. The ionic strength was set to 0.15 M to represent the spatial distribution of the electrostatic potential of sodium chloride in physiological conditions, with +1/-1 charges for the counterions, Na+ (0.95 Å) and Cl- (1.81 Å).[209] The final output was visualised at ±5 kBT/e.

2.12.2 Outer channel surface hydrophobicity

Hydrophobicity was achieved using the color_h.py python script in PyMol. The script defines the relative hydrophobic and hydrophilic value of amino acids. The scale was based on Eisenberg's normalised consensus hydrophobicity scale.[210]

PPB: 2FHA was used for huHC- and 5XB1 for ΔDE. The scale was based on Eisenberg's normalised consensus hydrophobicity scale.[210]

The script, color_h.py, had the following code. The code was acquired from http://www.protein.osaka-u.ac.jp.

```
# color_h
# -------
# Amino acid scale values:
#
# Ala:  0.620
# Arg: -2.530
# Asn: -0.780
# Asp: -0.900
# Cys:  0.290
# Gln: -0.850
```

```
# Glu: -0.740
# Gly:  0.480
# His: -0.400
# Ile:  1.380
# Leu:  1.060
# Lys: -1.500
# Met:  0.640
# Phe:  1.190
# Pro:  0.120
# Ser: -0.180
# Thr: -0.050
# Trp:  0.810
# Tyr:  0.260
# Val:  1.080
#
# Usage:
# color_h (selection)
#
from PyMol import cmd

def color_h(selection='all'):
        s = str(selection)
      print s
        cmd.set_color('color_ile',[0.996,0.062,0.062])
        cmd.set_color('color_phe',[0.996,0.109,0.109])
        cmd.set_color('color_val',[0.992,0.156,0.156])
        cmd.set_color('color_leu',[0.992,0.207,0.207])
        cmd.set_color('color_trp',[0.992,0.254,0.254])
        cmd.set_color('color_met',[0.988,0.301,0.301])
        cmd.set_color('color_ala',[0.988,0.348,0.348])
        cmd.set_color('color_gly',[0.984,0.394,0.394])
        cmd.set_color('color_cys',[0.984,0.445,0.445])
        cmd.set_color('color_tyr',[0.984,0.492,0.492])
        cmd.set_color('color_pro',[0.980,0.539,0.539])
        cmd.set_color('color_thr',[0.980,0.586,0.586])
        cmd.set_color('color_ser',[0.980,0.637,0.637])
        cmd.set_color('color_his',[0.977,0.684,0.684])
        cmd.set_color('color_glu',[0.977,0.730,0.730])
        cmd.set_color('color_asn',[0.973,0.777,0.777])
        cmd.set_color('color_gln',[0.973,0.824,0.824])
        cmd.set_color('color_asp',[0.973,0.875,0.875])
        cmd.set_color('color_lys',[0.899,0.922,0.922])
        cmd.set_color('color_arg',[0.899,0.969,0.969])
        cmd.color("color_ile","("+s+" and resn ile)")
        cmd.color("color_phe","("+s+" and resn phe)")
        cmd.color("color_val","("+s+" and resn val)")
        cmd.color("color_leu","("+s+" and resn leu)")
        cmd.color("color_trp","("+s+" and resn trp)")
        cmd.color("color_met","("+s+" and resn met)")
        cmd.color("color_ala","("+s+" and resn ala)")
        cmd.color("color_gly","("+s+" and resn gly)")
        cmd.color("color_cys","("+s+" and resn cys)")
        cmd.color("color_tyr","("+s+" and resn tyr)")
        cmd.color("color_pro","("+s+" and resn pro)")
        cmd.color("color_thr","("+s+" and resn thr)")
        cmd.color("color_ser","("+s+" and resn ser)")
        cmd.color("color_his","("+s+" and resn his)")
```

```
        cmd.color("color_glu","("+s+" and resn glu)")
        cmd.color("color_asn","("+s+" and resn asn)")
        cmd.color("color_gln","("+s+" and resn gln)")
        cmd.color("color_asp","("+s+" and resn asp)")
        cmd.color("color_lys","("+s+" and resn lys)")
        cmd.color("color_arg","("+s+" and resn arg)")
cmd.extend('color_h',color_h)


def color_h2(selection='all'):
        s = str(selection)
        print s
        cmd.set_color("color_ile2",[0.938,1,0.938])
        cmd.set_color("color_phe2",[0.891,1,0.891])
        cmd.set_color("color_val2",[0.844,1,0.844])
        cmd.set_color("color_leu2",[0.793,1,0.793])
        cmd.set_color("color_trp2",[0.746,1,0.746])
        cmd.set_color("color_met2",[0.699,1,0.699])
        cmd.set_color("color_ala2",[0.652,1,0.652])
        cmd.set_color("color_gly2",[0.606,1,0.606])
        cmd.set_color("color_cys2",[0.555,1,0.555])
        cmd.set_color("color_tyr2",[0.508,1,0.508])
        cmd.set_color("color_pro2",[0.461,1,0.461])
        cmd.set_color("color_thr2",[0.414,1,0.414])
        cmd.set_color("color_ser2",[0.363,1,0.363])
        cmd.set_color("color_his2",[0.316,1,0.316])
        cmd.set_color("color_glu2",[0.27,1,0.27])
        cmd.set_color("color_asn2",[0.223,1,0.223])
        cmd.set_color("color_gln2",[0.176,1,0.176])
        cmd.set_color("color_asp2",[0.125,1,0.125])
        cmd.set_color("color_lys2",[0.078,1,0.078])
        cmd.set_color("color_arg2",[0.031,1,0.031])
         cmd.color("color_ile2","("+s+" and resn ile)")
         cmd.color("color_phe2","("+s+" and resn phe)")
         cmd.color("color_val2","("+s+" and resn val)")
         cmd.color("color_leu2","("+s+" and resn leu)")
         cmd.color("color_trp2","("+s+" and resn trp)")
         cmd.color("color_met2","("+s+" and resn met)")
         cmd.color("color_ala2","("+s+" and resn ala)")
         cmd.color("color_gly2","("+s+" and resn gly)")
         cmd.color("color_cys2","("+s+" and resn cys)")
         cmd.color("color_tyr2","("+s+" and resn tyr)")
         cmd.color("color_pro2","("+s+" and resn pro)")
         cmd.color("color_thr2","("+s+" and resn thr)")
         cmd.color("color_ser2","("+s+" and resn ser)")
         cmd.color("color_his2","("+s+" and resn his)")
         cmd.color("color_glu2","("+s+" and resn glu)")
         cmd.color("color_asn2","("+s+" and resn asn)")
         cmd.color("color_gln2","("+s+" and resn gln)")
         cmd.color("color_asp2","("+s+" and resn asp)")
         cmd.color("color_lys2","("+s+" and resn lys)")
         cmd.color("color_arg2","("+s+" and resn arg)")
cmd.extend('color_h2',color_h2)
```

2.12.3 Channel radius

CAVER Analyst 2.0 was used to calculate the width of each apoFt channel.[211] PPB: 2FHA was used for huHC- and PDB: 5XB1 for ΔDE. Tunnel calculations were optimised for the three-fold channel and four-fold channel of both huHC- and ΔDE. The weight coefficient was set to 1.0 and the clustering threshold to 4.0 to reduce tunnel splitting. Tunnels going through the channel were selected for analysis.

huHC- (four-fold): Starting residue 161, probe radius 0.7 Å, shell radius 6 Å, and shell depth 12 Å.

ΔDE (four-fold): Starting residue 160, probe radius 5 Å, shell radius 25 Å and shell depth 5 Å.

Three-fold channel: Starting residue 123, probe radius 0.8, shell and radius 6 and shell depth 6 Å.

2.12.4 Modelling the apoferritin - affibody

A homology model for huHC- Afb was generated using MODELLER, version 9.24 (Sali Lab of California).[212] Structural data was acquired using the protein database for the human heavy chain subunit (PDB ID: 2FHA), apoFt human heavy chain cage (PDB ID: 6GSR) and $Z_{HER2:342}$ Afb (PDB ID: 2KZI). The AA sequence of huHC- (with the linker domain: VDGGGSGGSEF) was aligned and structurally modelled against the trimer subunit structure of 2FHA. Similarly, the structure of Afb (with the HIS-tag) was modelled against 2KZI. The trimer apoFt and Afbs were then modelled together, and the model with the lowest Discrete Optimised Protein Energy (DOPE) was carried forward. The linker domain was optimised using MODELLER's loop model class. The final cage structure was structurally aligned and superimposed against 6GSR.

2.13 Iron core mineralisation

huHC- and HuH 222 were rendered iron-free by adding 1 mM ethylenediaminetetraacetic acid

(EDTA) in 20 mM Tris at pH 8.0. The protein solution was then left to incubate on a roller at 15 rpm (Cole-Parmer™ Stuart™) for 24 hrs at 4 $^{\circ}$C to ensure iron chelation. The resulting solution was dialysed (6-8 kDa dialysis bag) against 20 mM Tris pH 8.0 overnight at 4 $^{\circ}$C, to remove the EDTA-iron complex. The protein solution was then removed from dialysis and filtered using a 0.22 µm filter. To measure the formation of Fe(III) oxo(hydroxo) species, 1 mL of 0.2 µM iron-free apoFt (cage) was added to a quartz cuvette, and Fe(II) solutions (50 mM $FeSO_4$ in 1 mM HCl, pH 2.0) was added at a final concentration of 50 µM. The final solution was rapidly mixed to achieve a molar ratio of 250 iron atoms per apoFt cage (10.4 iron atoms per subunit). The formation of the oxo(hydroxo) species was measured using a UV-vis spectrometer (BioDrop Duo), and time-dependent absorbance kinetic readings were collected at 305 nm. The reference sample was 0.2 µM apoFt solution.

2.14 Ammonium sulphate cut conjugation of 5-FAM to apoferritin

5-Carboxyfluorescein (Sigma-Aldrich) (5-FAM) was dissolved in DMSO to give a final concentration of 100x the molar concentration of the apoFt subunit concentration. DTT and EDTA were added to the apoFt solution at a final concentration of 2 mM DTT and 1 mM EDTA, respectively. The mixture was left to incubate on a roller (Cole-Parmer™ Stuart™) at 15 rpm for 2 hrs. The protein solution was precipitated using a saturated solution of ammonium sulphate (4.06 M at 20 $^{\circ}$C). The Ammonium sulphate solution was gradually added to the protein solution until a final concentration of 70% v/v ammonium sulphate was reached. The final solution was left to incubate on a rotor for 20 mins at RT and then centrifuged at 13,000 g for 10 mins. The supernatant was removed, and the pellet was quickly resuspended in 400 µL of conjugation buffer (section 2.1). The mixture was quickly vortexed and left to react overnight at 4 ˚C in the dark on a shaker. The protein solution was dialysed against 8 M urea, 20 mM Tris pH 7.2 for 8 hrs in a 6-8 kDa dialysis bag at RT and then against 20 mM Tris pH 8.0 at 4 $^{\circ}$C. The final apoFt solution (250 µM, 24-mer) was stored in 20 mM Tris pH 8.0 at 4 $^{\circ}$C for use within one week.

2.15 Free and encapsulated cytochrome c concentration determination

Bovine Cyt C (Sigma-Aldrich) stock solution (200 µM) was diluted to 100-0 µM using 20 mM Tris pH 8.0. Cyt C solution (100 µL) was added onto a 96 Greiner clear bottom well plate and measured using a Tecan Spark 10 M. Wavelength measurements were recorded between 300-600 nm. Readings at 408 nm corresponding to the soret band of Cyt C (Extinction coefficient 11460 M^{-1} cm^{-1}, assuming all Cys residues are reduced) were corrected for buffer background absorbance and a standard curve was plotted.[213] Free and apoFt encapsulated Cyt C concentrations were calculated using the following linear regression equation:

$$\text{Cyt C Molar concentration (M)} = 48.4x + 0.2249$$

2.16 Cytochrome C encapsulation experiments.

2.16.1 Urea-based method for cytochrome C encapsulation in apoferritin

huHC- at 3 µM (cage complex concentration) and Cyt C at 36 µM were dialysed against 8 M urea 50 mM Tris pH 8.9 for 18 hrs at 4 °C. The two protein solutions were then mixed at a huHC- cage 1:12 Cyt C molar ratio. The mixture was incubated on a shaker for 1 hr and then dialysed (6-8 kDa dialysis bag) against 20 mM Tris pH 8.0 for 18 hrs at 4 °C to promote apoFt cage formation. Following encapsulation, free Cyt C and huHC- subunits were removed using a spin concentrator with a 100,000 MWCO. The encapsulated complex's total protein concentration (mg/ mL) was determined using NanoDrop® A-1000 at UV_{280}. Following this, the concentration of encapsulated Cyt C was measured using a 96 Greiner clear bottom well plate and a Tecan Spark 10 M at UV_{408}. The molar concentration was then calculated using the standard curve created in section 2.15. The final huHC- cage concentration (mg/ mL) was calculated by deducting the Cyt C concentration from the total protein concentration measured using NanoDrop® A-1000 at UV_{280}.[213]

2.16.2 pH-based method for cytochrome C encapsulation in apoferritin

pH Disassembly: huHC- or ΔDE were pre-equilibrated in 50 mM KCl pH 3.0 or 50 mM sodium acetate pH 4.0, respectively, for 18 hrs at 4 °C via dialysis (6-8 kDa dialysis bag). Cyt C was also preequilibrated in 50 mM Tris pH 9.0 overnight via dialysis (6-8 kDa dialysis bag). ApoFt was then added to Cyt C solution at a apoFt cage 1:12 Cyt C molar ratio. The mixture was incubated with shaking for 2 hrs at RT and then dialysed for 18 hrs at 4 °C against 20 mM Tris pH 8.9 to allow complete apoFt cage formation. Free Cyt C was then removed using a spin concentrator with a 100,000 Dalton MWCO. The encapsulated complex's total protein concentration (mg/ mL) was determined using NanoDrop® A-1000 at UV_{280}. Following this, the concentration of encapsulated Cyt C was measured on a 96 Greiner clear bottom well plate using a Tecan Spark 10 M at UV_{408}. The molar concentration was calculated using the standard curve in section 2.15. The final huHC- cage concentration (mg/ mL) was calculated by deducting the Cyt C concentration from the total protein concentration measured NanoDrop® A-1000 at UV_{280}.[213]

3.0 Evaluating uptake of affibody-apoferritin via transferring receptor 1 and human epidermal growth receptor 2

3.1 Results and discussion

As detailed in the introduction, huH has an innate ferroxidase activity as part of its natural function to store iron.[46] In addition, huH is also able to bind to TfR1, triggering clathrin-coated endocytosis in normal cells, as well as TfR1 over expressing cancer cells such as breast cancer cells.[210, 211] Quantitative studies using immunohistochemistry (IHC) and single-cell inductively coupled plasma mass spectrometry (SC-ICP-MS) have reported TfR1 overexpression in breast cancer cells (~5x) than normal cells.[216, 217] Therefore, TfR1 has been the focus of some cancer-directed therapies, including apoFt-encapsulated agents.

The current chapter begins by describing key mutations that have been introduced into a cysteine-free human heavy chain apoferritin (huHC-) to diminish ferroxidase activity (huH 222). This was followed by kinetic studies to assess the effect of the ferroxidase mutation on iron core formation in the huH 222 mutant. The study then identified and removed potential residues on huHC- (TfR-) that could be involved in TfR1 binding. The effects of the TfR- on TfR1-dependent uptake were then studied *in vitro* using immunofluorescence microscopy and flow cytometry. Finally, the effects of an Afb engineered huHC- on HER2- and TfR1-dependent uptake were evaluated using variants huHC-, TfR- huHC-Afb, TfR-Afb and Afb variant. Uptake studies were performed against human breast cancer cell lines with varying levels of HER2 and TfR1.

3.1.1 Production of ferritin lacking ferroxidase activity

During ferroxidation in huH, the intermediate species, hydroxide ion ($\cdot$OH), is produced. Even momentarily, the presence of hydroxide ions can negatively impact encapsulated therapeutics. This includes reducing peptide/protein disulphide bonds or early activation of hydroxyl activated pro-drugs.[218, 219] However, a study by Levi *et al.*, 1992 identified key residues in the ferroxidase centre that can significantly reduce Fe(III) core formation.[78] Levi *et al.*, 1992 concluded that residues E61, E62, E64, H65 and E67 were essential for ferroxidase activity. This innate ferroxidase activity is not seen in huL, and as such residues, A61, K62, A64, G65

and A67 on huL were substituted into huHC-'s B-helix. The new mutant was termed 'huH 222'. Recombinant protein expression of huH 222 at 37 ˚C for 3 hrs showed overexpression in both the soluble and insoluble fraction, as seen in Figure 3.1. huH 222's expression profile was similar to its parent protein huHC-, indicating that these modifications did not significantly affect huH 222's solubility during protein expression (SFigure 3.1).
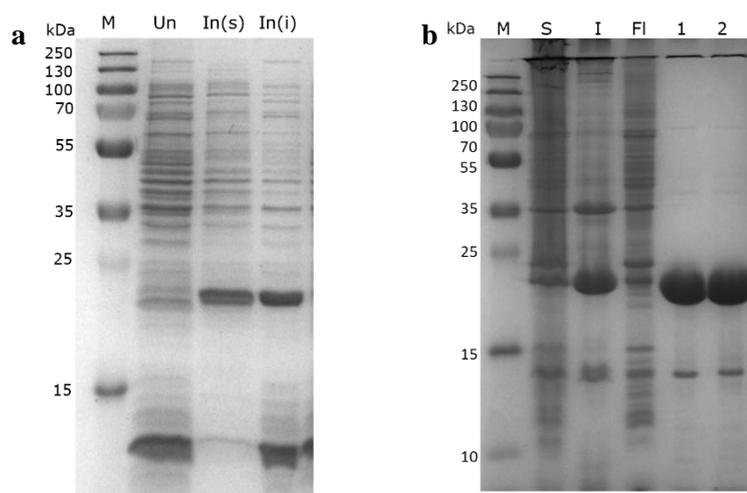

Figure 3. 1 Expression and purification profile of huH 222. **a,** huH 222 expression test at 37 ºC for 3 hrs and **b**, huH 222 over expression. M: Marker, Un: Uninduced, In(s): Induced soluble, In(i): Induced insoluble, S: Soluble fraction, I: Insoluble fraction, Fl: flow through and 1/2: Fraction 1, and 2.

Following successful purification, CD was performed to determine if changes along the B-helix of huH 222 would affect apoFt's native α-helical secondary structure. CD spectra of both huHC- and huH 222 showed similar profiles, with negative bands at ~222 nm and ~208 nm and a positive band at ~193 nm (Figure 3.2). These results indicated an α-helical profile and confirmed that mutations along the B-helix do not significantly affect huH 222's secondary structure. Once it was confirmed that huH 222's secondary structure was not affected, kinetic studies were performed to monitor the formation of the diferric peroxo (DFP) catalytic intermediate at UV $_{305\ nm}$. Kinetic studies indicated a reduced Fe (III) core formation rate for huH 222 compared to huHC-, suggesting diminished ferroxidase activity (Figure 3.3). However, huH 222's ferroxidase activity was only 40 secs slower than huHC-, which indicated that the iron core is still being formed. Given Levi *et al.*, 1992 showed that these mutated residues within huH 222 completely diminish ferroxidase activity, it can be suggested that

experimental conditions could be affecting iron core formation.[78] One condition that would likely affect iron core formation is the buffer in which the kinetic study was performed. In the current kinetic study, Tris buffer at pH 8.0 was chosen as the buffer system, as huHC- and huH 222 precipitate at < pH 7.5. However, kinetics studies performed at high pH, e.g., >pH 8.0 and in Tris buffer, have been shown to increase iron autooxidation.[220] Therefore, it can be hypothesised that the ferroxidase activity seen for huH 222, could be caused by Tris pH 8.0. To verify the effects of Tris or high pH on iron core formation, kinetic studies could be performed in Tris or PBS at pH 7.5, elucidating the effects. However, given that huH 222 showed at least a partial reduction in iron core formation time, the next step focused on removing huHC-'s second innate characteristic, its ability to bind to TfR1.
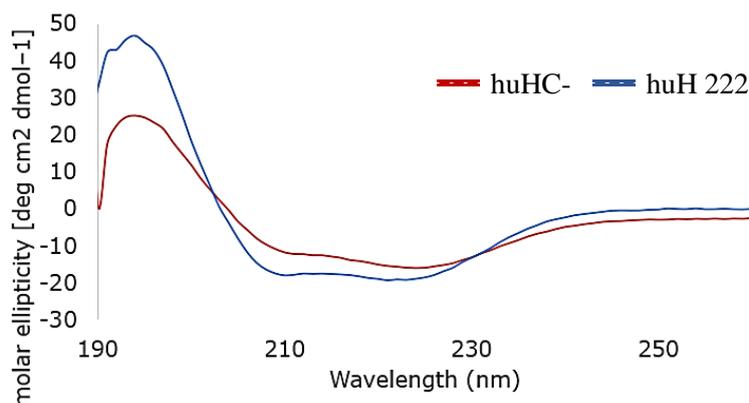

Figure 3. 2 Circular dichroism study on the secondary structure of huHC- and huH 222. Both variants were recorded at 20 mM Tris pH 8.0 at 25 °C. ApoFt variant concentrations 4 μM.
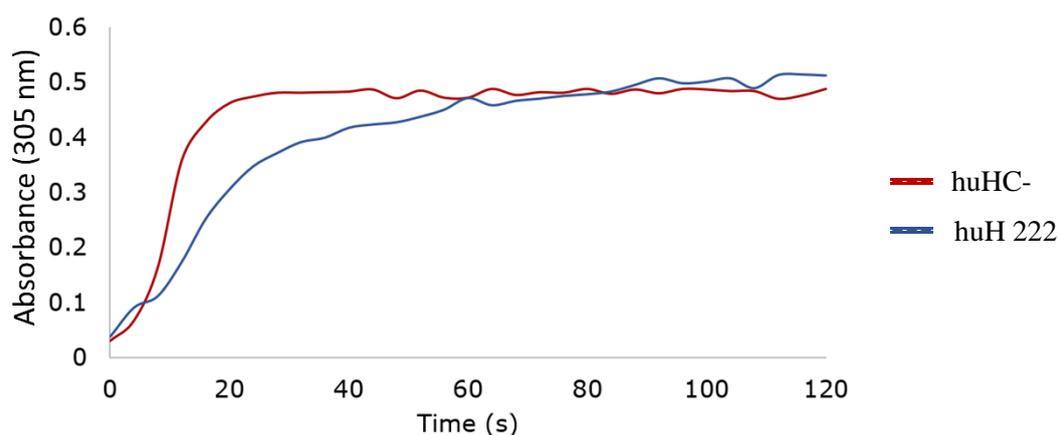
Figure 3. 3 Kinetic curves measuring Fe(III) core formation in huHC- and huH 222. Absorbance at 305 nm was used to measure the formation of Fe(III) oxo(hydroxo) species formation. Conditions: 0.2 μM apoFt, 50 μM FeSO$_4$, 20 mM Tris, pH 8.0, 20 °C.

3.1.2 Identifying and mutating potential transferrin receptor 1 binding sites on apoferritin

At the time of this study was initiated, no known TfR1-huH binding residues had been identified. However, experiments investigating TfR1 binding against the Machupo virus (MACV) and Tf reported extensive contact along the flexible loop region of each ligand.[45, 46] Therefore, it was speculated that huHC-'s flexible BC loop could also be important in huHC-TfR1 interaction. This hypothesis was further supported by another study which investigated the addition of huH's BC loop onto an archaeal apoFt. Unlike archeal apoFt, which is unable to undergo TfR1-mediated endocytosis in human cells, the new humanised archaea (humAf) was successfully endocytosed by TfR1.[93] To narrow down further the residues that could be involved in TfR1 binding, the BC loop sequence from huHC-, huL, and archaeal Ft (38% sequence identity and 58% sequence similarity to huHC-) were aligned. The alignment of all three variants indicated that three key residues on huHC- could be involved in TfR1 binding. These residues include I80, F81 and L82, as seen in Figure 3.4a. In addition, superimposed models of huL's and huHC-'s BC loop at positions 80, 81 and 82 showed a phenylalanine shift from position 81 to 82, which could affect huHC-'s interaction with TfR1, given its large bulky side group, as seen in Figure 3.4b. Therefore, residues on the flexible loop of huH 222 were mutated to the equivalent amino acid present on huL: I80A, F81L and L82F. The new apoFt

variant was termed huHC- transferrin receptor mutant (TfR1-) and was expressed and purified, as shown in Figure 3.5. The expression profile of TfR1- was similar to its parents, huHC- and huH 222, indicating residues I80A, F81L and L82F had little effect on protein solubility during expression.

a

huH (80-82)	I F L	Q D I K K P
Archaeal	V K L	Y A V E E P
huL	A L F	Q D I K K P

b
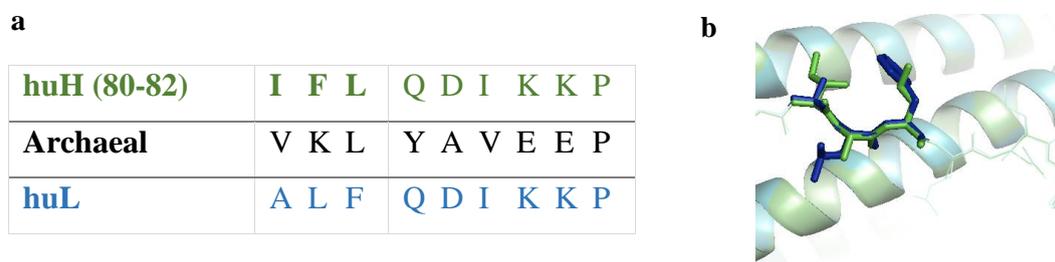
Figure 3. 4 Superimposed BC loop of apoFt variants. **a**, Sequence alignment of BC loop from three different apoFt variants (huH, archaeal and huL) and **b**, Superimposed stick model of residues 80-82 from A, L, F in huL and I, F, L in huH. huH: Human heavy chain apoFt, Archaeal: *Archaeoglobus fulgidus* apoFt and huL: Human light chain apoFt.
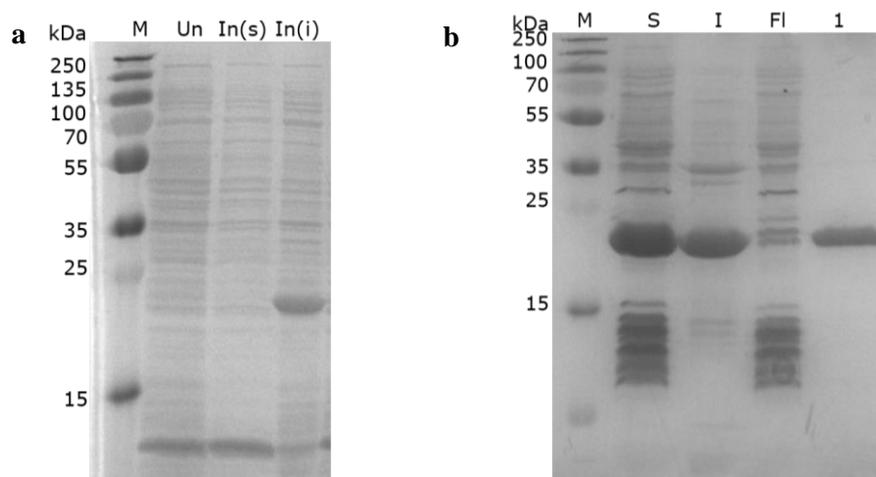
Figure 3. 5 Expression and purification of TfR-. **a**, TfR- expression test at 37 ℃ for 3 hrs and **b**, TfR- His-tag purification. M: Marker, Un: Uninduced, In(s): Induced soluble, In(i): Induced insoluble, S: Soluble fraction, I: Insoluble fraction, Fl: Flow through and 1: Fraction 1.

For future *in vitro* uptake studies, a single cysteine residue was also introduced into huHC- and TfR1-. The presence of a cysteine would permit conjugation of maleimide-linked fluorescent probes, allowing for cellular imaging and quantitative studies. The cysteine was introduced within apoFt's cavity, rather than the exterior, to avoid any possible adverse effects the fluorescent probe might have on TfR1 interaction. The new variants were termed huHC- K68C

and TfR- K68C. Following the introduction of the K68C mutant, each apoFt variant was overexpressed at 37 °C for 3 hrs. The addition of the cysteine resulted in each apoFt variant being overexpressed in the insoluble fraction, as seen in Figures 3.6a,b. One possible reason for overexpression in the insoluble fraction could be due to non-native disulphide bridge formation during protein expression. Disulphide formation would likely change the overall fold of the protein during protein synthesis and, as such, shift the expression to the insoluble fraction. To allow apoFt to re-fold properly, the insoluble fraction was resuspended in buffers containing urea and DTT.
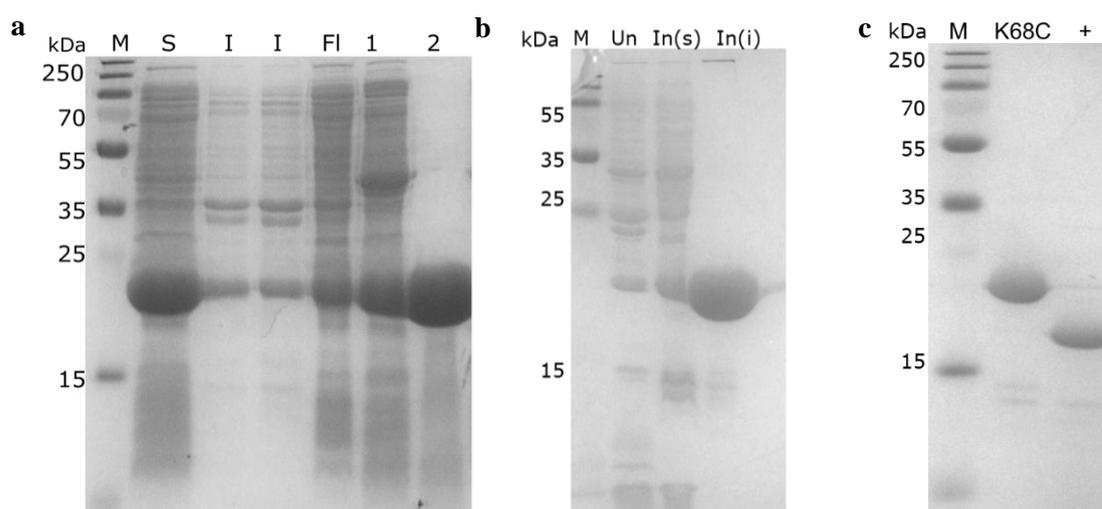
Figure 3.6 K68C purification **a**, Purification of huHC- K68C **b**, Purification of TfR- K68C, M: Marker, Un: Uninduced, In(s): Induced soluble, In(i): Induced insoluble, S: Soluble fraction, I: Insoluble fraction, Fl: Flow through and 1/2: Fraction 1 and 2., and **c**, TfR- HIS-tag cleavage using thrombin (+): 0.5 μg/mL Thrombin, 2 mM CaCl$_2$, 20 mM Tris, pH 8.0, 24 hrs, 20 °C.

In addition to studying the effects of TfR1- on cellular uptake, the N-terminal region (HIS-tag) was also evaluated. The N-terminal region, consisting of a histidine-tag and an avidin-tag, was added to allow purification using IMAC. The HIS-tag region is thought not to affect apoFt's ability to encapsulate therapeutic agents. However, the presence of a positively charged hexahistidine region within the HIS-tag region could interact with the negatively charged cell membrane, impacting apoFt internalisation. Therefore, the effect of the HIS-tag was evaluated. A thrombin cleavage reaction at 20 °C overnight was performed to ensure complete removal of the HIS-tag, as seen in Figure 3.6c. The four constructs containing the K68C mutations were termed huHC-, TfR1-, huHC- (HIS-) and TfR- (HIS-) for simplicity (Figure 3.7). Each

construct showed a slight variation in molecular weight, extinction coefficient and pI (SFigure 3.2). As all four constructs contained a K68C mutation, the next step was to conjugate a fluorophore, optimise conjugation efficiency and evaluate if conjugation is possible with an internal cysteine.
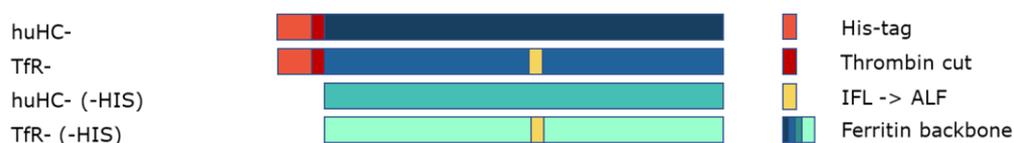
Figure 3.7 Schematic representation outlining key regions in each apoFt variants.

3.1.3 Successful conjugation of a maleimide dye to K68C

In the following experiment, urea was used to relax/partially disassemble the apoFt cage and allow fluorophore 5-carboxyfluorescein (5-FAM) conjugation to K68C (Figure 3.8). The use of urea to relax/partially disassemble huHC- for 5-FAM entry is supported by literature investigating the effects of urea on the structure of apoFt. [223, 224] Using TEM images, the study showed that in the presence of 8 M urea pH 7.4, apoFt's channels 'relax' compared to apoFt's usually rigid structure at pH 7.4 as observed by TEM.[223, 224] The study also highlighted the effects of urea on DOX encapsulation. Under normal physiological conditions, apoFt's channels are 3-4 Å wide and prevent diffusion of large molecules like DOX, which is 13.6 x 9.9 Å (Figure 3.9). However, using 8 M urea, it was shown that large molecules such as DOX could diffuse through the channel.[223] Given that under 8 M urea conditions apoFt is able to encapsulate DOX, it was hypothesised that 8 M urea could also allow entry of 5-FAM into apoFt's core for K68C conjugation, given that both 5-FAM and DOX have similar dimensions (Figure 3.8 and Figure 3.9).

Figure 3.8 Stick-frame chemical structure of 5-FAM. Dotted lines indicate distance in angstrom (Å) between hydrogen atoms. Grey: carbon, white: hydrogen, red: oxygen, blue: nitrogen and red circle: maleimide group. The maleimide group reacts with sulfhydryl groups on cysteines to form a stable thioester bond.

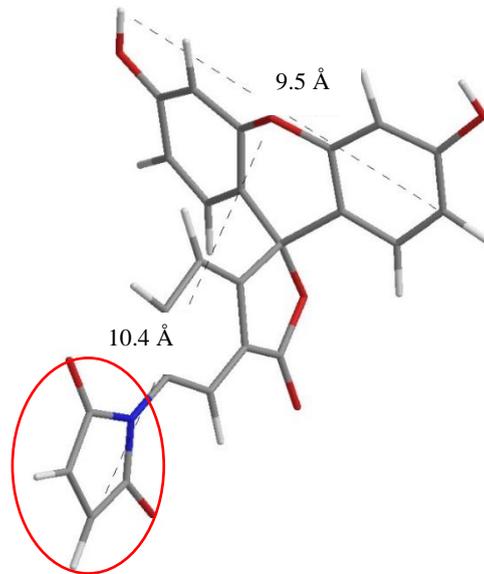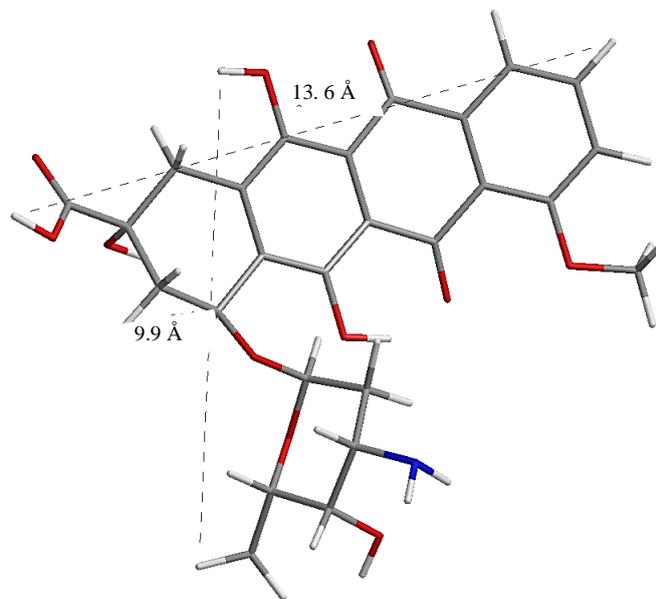**Figure 3.9** Stick-frame chemical structure of doxorubicin. Dotted lines indicate distance in angstrom (Å) between hydrogen atoms. Grey: carbon, white: hydrogen, red: oxygen and blue: nitrogen.

Initial conjugation tests involved incubation of huHC- K68C in 20 mM Tris, 2 mM DTT and 8 M urea, followed by overnight dialysis against 20 mM Tris and 8 M urea buffer. Dialysis against

20 mM Tris and 8 M urea buffer was performed to ensure excess DTT removal and prevent DTT from reacting with the maleimide fluorophore. However, prolonged incubation of K68C variants in 20 mM Tris and 8 M urea buffer urea resulted in low conjugation efficiency (SFigure 3.3). Therefore, an ammonia sulphate step was introduced to reduce the time a K68C mutant was in solution without a reducing agent. The ammonium sulphate step involved the addition of ammonium sulphate (70% w/v) to an apoFt solution (~400 μM), thereby causing the protein to precipitate out of the solution, a process known as 'salting-out'. The ammonium sulphate protocol termed the 'ammonium sulphate precipitation method' in this study took an average of ~40 mins, compared to the dialysis protocol, which took >4 hrs. The short time that the DTT was absent in the ammonium sulphate precipitation method compared to the dialysis protocol was predicted to keep the K68C variants in a reduced state until conjugation. This was confirmed by mass spectrometry (MS) which showed that the ammonium sulphate precipitation method produced a high ratio of conjugated to unconjugated K68C species (SFigure 3.3). Therefore, the ammonium sulphate precipitation method was used to perform the maleimide conjugation reaction for all apoFt variants (Figure 3. 10a). The next step involved evaluating the effects of the conjugation process and the presence of 5-FAM on the native structure of each apoFt variant. As such native PAGE was performed to confirm if each variant was forming a typical 440kDa apoFt cage-like structure post conjugation (Figure 3.10b, c). Native PAGE analysis showed similar migration bands between the 'parent' unconjugated species and the conjugated species at ~440 kDa (Figure 3.10b, c), indicating that each variant had successfully formed a cage-like structure.
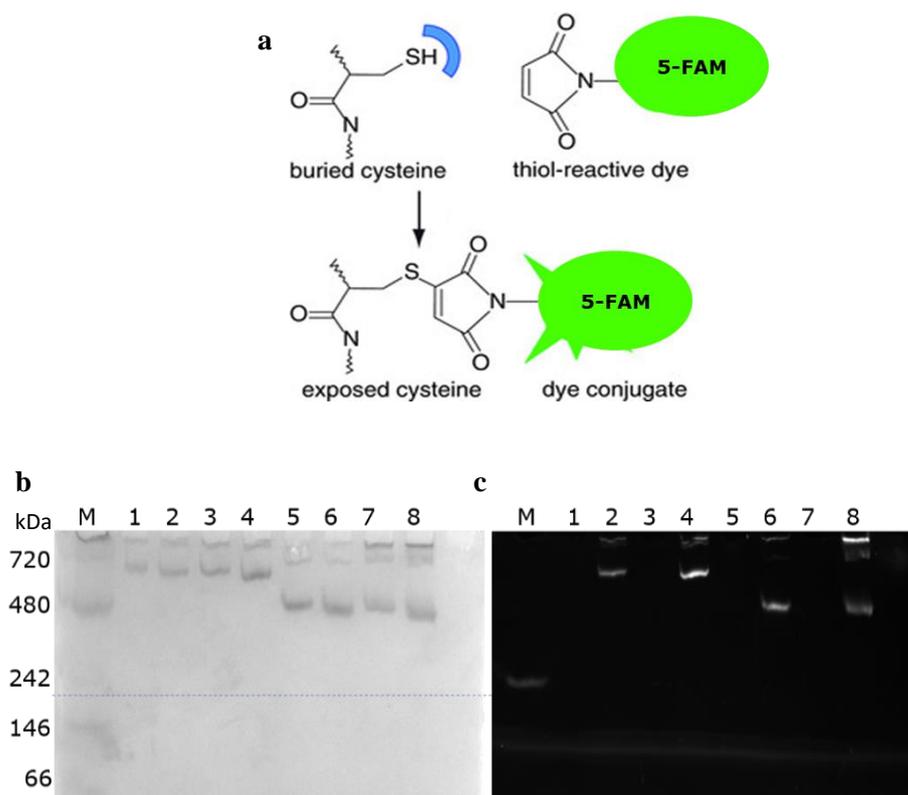
Figure 3.10 ApoFt variants after conjugation. **a**, Maleimide reaction for chemical conjugation to sulfhydryl on K68C. **b**, Native PAGE of intact ApoFt variants after conjugation to 5-FAM under normal light, and **c**, Native PAGE of intact ApoFt variants after conjugation to 5-FAM under UV light. M: Marker, 1: HC-, 2: huHC- (5-FAM), 3: TfR-, 4: TfR- (5-FAM), 5: huHC-(-HIS), 6: huHC-(-HIS) (5-FAM), 7: TfR-(-HIS) and 8: TfR- (-HIS) (5-FAM).

The successful attachment of 5-FAM on K68C variants was also confirmed by MS analysis, which showed varying levels of conjugation efficiency between variants, as seen in Figure 3.11a, b: HC (30%), TfR- (36%), HC-(-HIS) (11%) and TfR-(-HIS) (36%). A 100% conjugation efficiency was not achieved under the current conjugation conditions. However, this could be due to the proximity of the K68C residues to each other and the channel, sterically hindering further conjugation of 5-FAM. In addition, a notable 135 Da increase in MW was also seen for huHC- (-HIS-tag)'s MS profile post-conjugation, compared to the 'parent ' peak pre-conjugation, as seen in Figure 3.11a. Several possible reasons could account for this, including metal cation contamination in the MS machine or contamination in the elution buffer during the sample preparation phase. The presence of salts such as sodium chloride in either the sample solution or mass spectrometry can lead to multiple cations to bind to AAs, including sodium ions, which could lead to peak broadening or adduct formation and, as such, a higher Da.[225]

As both native PAGE and MS indicated that 5-FAM can successfully enter apoFt's cavity, it can be suggested that the presence of 8 M urea can 'relax'/partially disassemble apoFt's channels, allowing entry of large molecules (>3-4 Å). Therefore, 8 M urea could be used for future experiments to facilitate the entry and encapsulate large molecules as much as 9.4 Å (width of 5-FAM). It would also be interesting to see if larger molecules or lower urea concentrations would also achieve encapsulation.
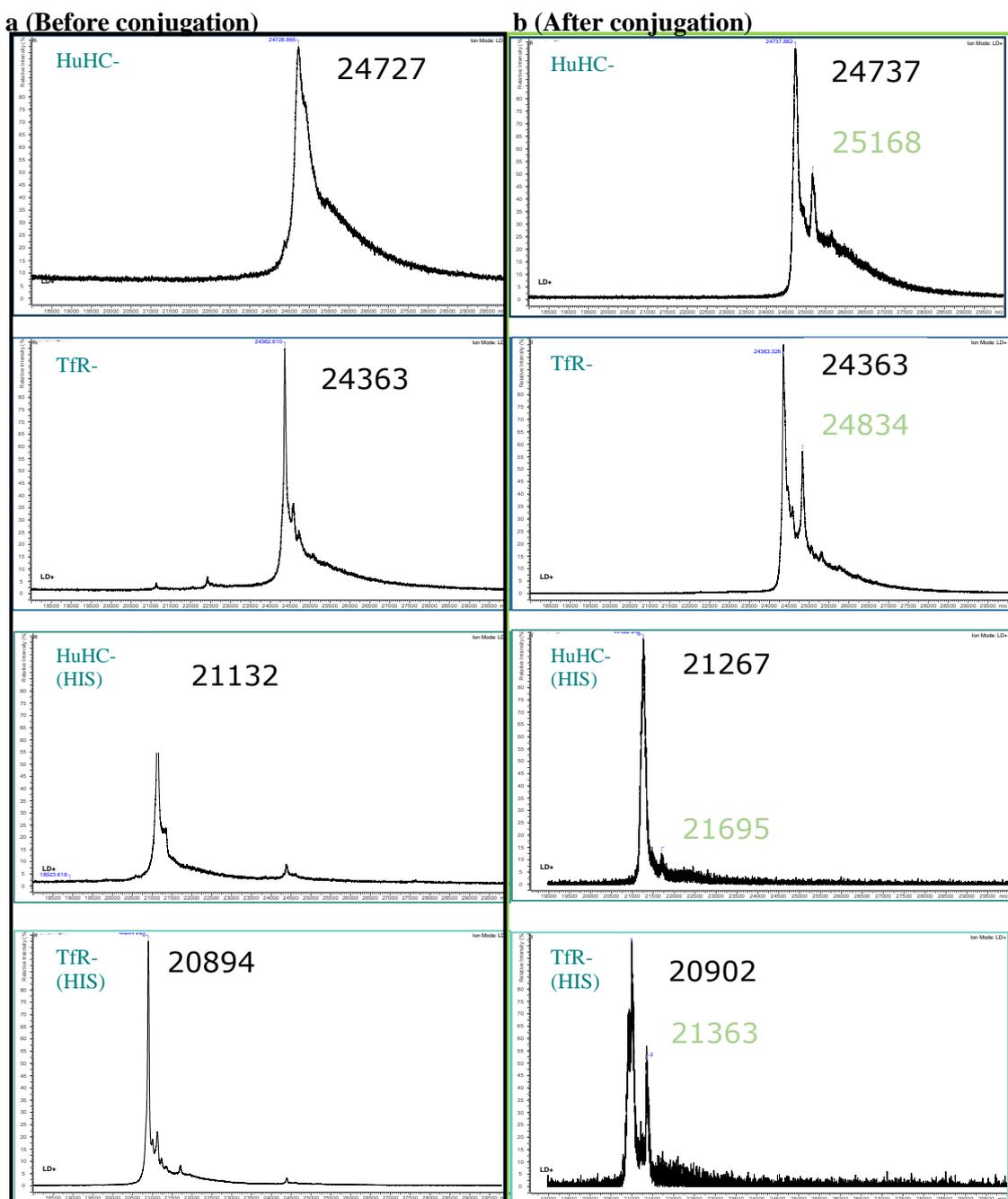

Figure 3.11 Conjugation efficiency of apoFt variants. **a,** Mass spectra of apoFt variants before (black box) and **b,** after conjugation to 5-FAM (green box).

Following successful conjugation, the apoFt variants were carried forward for immunofluorescence imaging to verify the effects of the TfR- mutant. However, the difference in conjugation efficiency between each apoFt variant and the photo-instability of the 5-FAM molecule can lead to varying fluorescence signals. Therefore, it was essential to adjust the fluorescence intensity between conjugates (Figure 3.11a, b). To account for such factors, conjugated apoFts were scanned, and the linear relative fluorescence signal was calculated between different apoFts (Table 3.1). Following this, the equation $y = \log(x)/0.9$ was used to calculate the logarithmic gain difference between different samples and adjust the digital gain of each image captured during confocal microscopy. Once the gains were calculated, the conjugates were put forward for *in vitro* testing against breast cancer cells with varying TfR1 and HER2 expression.

	Fluorescence abs	Relative to TfR-	Log10
huHC-	77063	0.62	-0.21
TfR-	47844	1.00	0.00
huHC- (-HIS)	11393	4.20	0.62
TfR- (-HIS)	11888	4.02	0.60

Table 3. 1 Example of the relative fluorescence from a conjugated apoFt variants. ApoFt variants diluted to 45 μg/ mL and scanned at Ex 485 nm and Em 535 nm.

3.1.4 Quantifying human epidermal receptor 2 and transferrin receptor 1 expression on breast cancer cells

Many NCs such as cyclodextrin and polyion complex micelles have been modified to target TfR1 overexpressing cancer cells.[226, 227] Such systems have used transferrin for TfR1 targeted therapy. However, different cancer cell lines have varying expression levels of TfR1, which could affect the level of NC uptake. To account for this difference, methods such as WB and IHC have been developed to semi-quantify TfR1 expression levels.[228, 229, 230.] Utilising a more

quantitative approach, with SC-ICP-MS and neodymium (^{143}Nd)-labelled anti-TfR1 mAb, the absolute expression levels of TfR1 on breast cancer cells from different malignancies can also be calculated.[216] Using this method, an absolute value of ~23 $\pm$ 0.4 x 10^3 TfR1 molecules/cell was calculated to be present on the surface of MDA-MB-231 cells. Although this method is very accurate, it requires conjugation and purification of ^{143}Nd-labelled antibodies, thereby increasing the experimental time and cost.

In addition to TfR1, HER2 overexpression has also been associated with some breast malignancies (as well as other tumour phenotypes, e.g., gastric carcinoma). Therefore there has been growing interest in developing reliable and accurate HER2 receptor quantification methods.[231] As with TfR1 quantification, the standard WB and IHC scoring methods can be used for preliminary protein verification. However, these techniques are not accurate and measure the whole-cell HER2 receptor number rather than just the cell surface number.[180] In comparison, the use of antibody-conjugated fluorescent dyes including Alexa Fluors, R-phycoerythrin or fluorescein isothiocyanate in conjunction with fluorescence quantitation kits and flow cytometry has enabled accurate HER2 receptor quantification.[232, 233, 234, 235]

Using similar quantification techniques to those stated in the literature, a fast in-house method was developed to measure HER2 and TfR1 expression. The current method used commercially available phycoerythrin-conjugated anti-HER2 (anti-HER2 mAb) and anti-TfR1 antibodies (anti-TfR1 mAb) as well a flow cytometry. Flow cytometry was used over WB for TfR1/HER2 quantification because, unlike WB, flow cytometry can distinguish intracellular and cell surface receptors. This was essential as the current study only needed to know the surface receptor expression levels of TfR1/HER2 to elucidate the effects of the TfR- mutant and future HER2 targeting ligands on huHC- uptake. Using a standard curve, a 1:1 ratio of phycoerythrin: mAb was confirmed for each conjugated antibody (SFigure 3.4 a,b). In addition, BD Quantibrite™ fluorescence beads were used to quantify receptor cell membrane expression, as seen in Figure 3.12a. The beads have four levels of conjugated phycoerythrin, allowing a fluorescence standard curve to be computed from flow cytometry.

Breast cancer cell lines BT474 (luminal, HER2 IHC score: 3+), SKBR3 (luminal, HER2 IHC score: 3+), MDA-MB-231 (basal B1+, HER2 IHC score: 0/1+), MDA-MB-468 (basal B, HER2 IHC score: 0/ 1+) and normal lung fibroblast MRC-5 have been stated in the literature to show varying expression levels of TfR1 and HER2 and therefore were used to evaluate this quantification method (Figure 3.12b). [236, 237, 238, 239] Although all four breast cancer cell lines have been shown to have similar levels of TfR1 expression, the HER2 expression levels vary.

Both BT474 and SKBR3 cell lines have been shown to overexpress a high number of the HER2 receptor, making them helpful in identifying the upper limits for the current quantification method.[240] The BT474 cell line was also employed as previous experiments from our group have utilised BT474 for *in vivo* apoFt experiments. Therefore, it would be good to evaluate the expression of HER2 and the effects of the TfR- mutant on BT474 *in vitro* before carrying it forward for *in vivo* work. To understand the lower limits of this newly developed receptor quantification method, MDA-MB-231 and MDA-MB-468 were used, given that both cell lines have been shown to have low HER2 expression levels, with MDA-MB-468 exhibiting little/no HER2 expression in WB.[241]

3.1.4.1 Transferrin receptor 1 quantification

Using flow cytometry 11 ± 8 x 10^3, copies of TfR1/cell were detected for MRC-5 cells, as shown in Figure 3.12c. This is much higher than those seen in WB experiments, where levels of TfR1 were below the detectable range.[242] However, given the high sensitivity of flow cytometry compared to WB, the TfR1 expression level detected for MRC-5 cells is acceptable, as most normal cells will have some level of TfR1 expression. Unlike 'normal' cells, breast cancer cells have a high TfR1receptor copy number, with the current method calculating a value of 113 ± 16 x 10^3 TfR1/cell for SKBR3 cells and 73 ± 23 x 10^3 TfR1/cell for MDA-MB-231 cells. The value obtained for MDA-MB-231 is ~3x than those reported with SC-ICP-MS (23 x 10^3).[216] As previous WB studies have shown that the levels of TfR1 in MDA-MB-231 cells are only marginally less than in SKBR3 cells, it can be suggested that the TfR1 value calculated for MDA-MB-231 is accurate.[207] A similar TfR1 surface expression value was also calculated for MDA-MB-468 and BT474 cells, at 77 ± 7 x 10^3 and 54 ± 4 x 10^3 TfR1/cell, respectively. As expression levels obtained by all four breast cancer cell lines were similar to those reported by WB, SC-ICP-MS and flow cytometry studies, it is indicative that this method is sensitive enough for detecting and quantifying TfR1 expression.[93, 94]

3.1.4.2 Human epidermal growth receptor 2 quantification

Both BT474 and SKBR3 cells showed high receptor copy numbers compared to the other cell lines, as seen in Figure 3.12c. However, BT474 showed a relatively low/receptor number of 912 ± 185 x 10^3 HER2/cell, compared to most values reported in the literature, stating a value of 1800-3600 x 10^3 HER2 /cell.[86, 95] A low receptor copy number of ~700 x 10^3 has previously

been reported using R-PE–conjugated antibodies and DAKO QUIFIKIT as calibration beads.[246] One possible reason for the difference in HER2 expression levels on BT474 could be the culture conditions in which the cells were grown (2D monolayer or 3D spheroids). BT474 cells cultured in 3D spheroids showed higher HER2 copy numbers, a likely cause of cell clustering, leading to HER2 upregulation.[247] Therefore, the considerable variation in HER2 expression detected for BT474 cells could be a consequence of slight variations in culture conditions. SKBR3 cells were found to express 864 ± 51 x 10^3 HER2/cell. The values detected for SKBR3 cells are slightly higher than those reported in the literature, ranging from 112-140 x 10^3 HER2/cell, but still relatively close.[232, 235] A much lower HER2 copy number of 9 ± 1 x 10^3 HER2/cell was detected for MDA-MB-231 cells. The HERE2 values calculated MDA-MB-231 cells are similar to those reported in Rodallec *et al.,* 2018 who detected 5 x 10^3 HER2/cell, solidifying flow cytometry's capability to quantify HER2 accurately.[235] A value of 5 ± 1 x 10^3 HER2/cell was detected for MDA-MB-468 cells, comparable to literature data stating 1 x 10^3 HER2/cell.[232] In comparison to WB, which shows no/minimal HER2 expression on MDA-MB-468 cells, the current flow cytometry method was able to detect and quantity low HER2 numbers. Therefore confirms the sensitivity and accuracy of flow cytometry for detecting low receptor copy numbers.

Comparing the receptor expression values reported in the literature to the values acquired using the current method, it can be suggested that the current method can accurately detect and quantify TfR1/HER2 expression for SKBR3, MDA-MB-231, MDA-MB-468, BT474 and MRC-5 cells. Therefore, the next step was to use SKBR3, which expresses high TfR1 and MRC-5 cells, which expresses very low TfR1, to verify the effects of uptake of huHC- and TfR-.
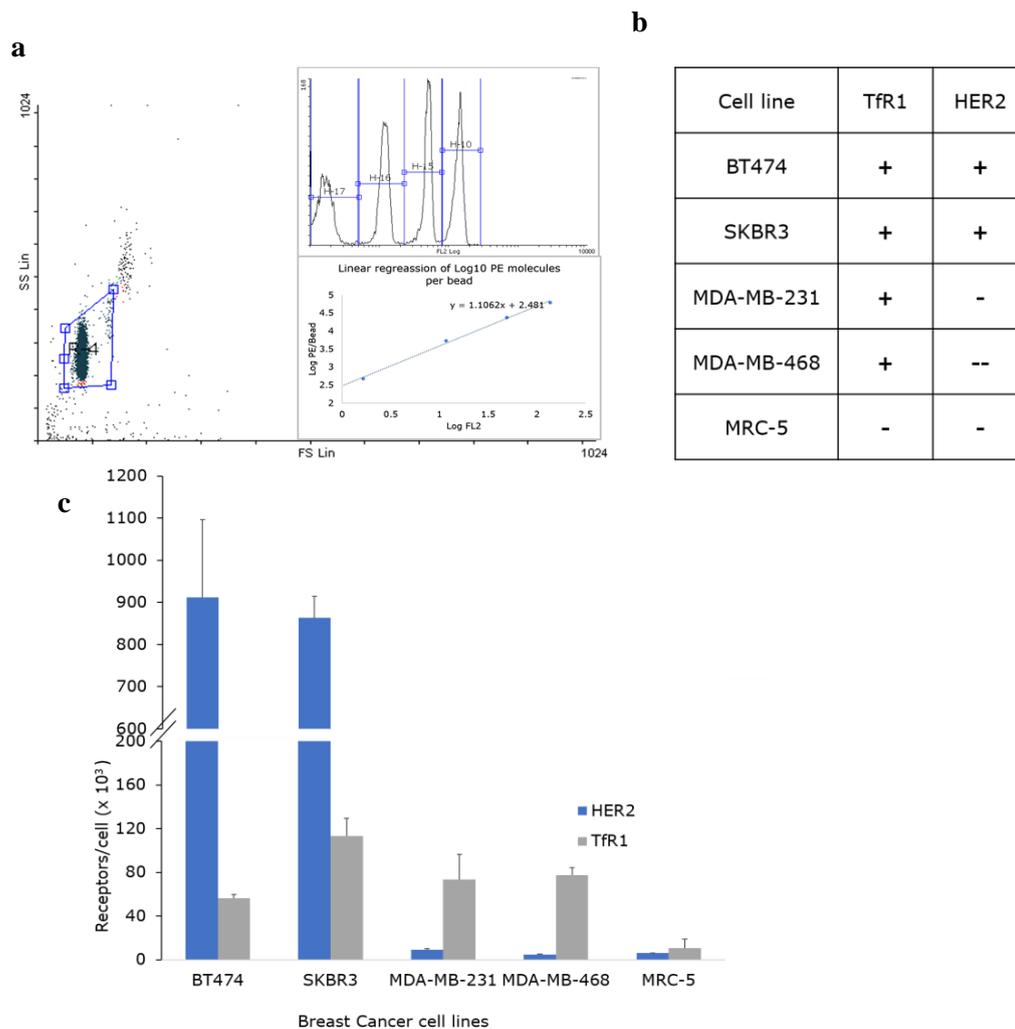
Figure 3.12 Quantitative analysis of TfR1 expression in BT474, SKBR3, MDA-MB-231, MDA-MB-468 and normal lung fibroblast MRC-5 cell lines. **a,** PE standard curve calculated from BD Quantibrite ™ fluorescence beads, **b,** HER2/TfR1 status of breast cancer cell lines and **c,** HER2 and TfR1 receptor levels in breast cancer cell lines BT474, SKBR3, MDA-MB-231 and MDA-MB-468 and MRC-5. (+) Receptor overexpression and (-) normal receptor expression.

3.1.5 Effect of TfR- and HIS-tag on apoferritin uptake

To corroborate and compare the uptake of huHC- and TfR- by TfR1 expressing cells, the study employed SKBR3 cells, as this cell line was shown to have the highest TfR1 expression (section 3.14). MRC-5 cells were also used as a control to verify if the TfR- mutations could increase uptake in non-transformed cells.

Live-cell confocal microscopy was performed to verify the TfR- variant's effects on TfR1-

mediated uptake. The use of live-cell confocal imaging would allow visual confirmation of whether the constructs were bound only to the exterior of the cells or were sequestered by cells. The initial cell seeding density was optimised to allow visualisation of single SKBR3 cells but not too low to affect the growth rate. The seeding density for the current experiment was established at 1.5 x10^4 cells per well or ~50% confluency. After 24 hrs incubation, cells were treated with 120 μg/ mL of huHC- (+/- HIS tag) or TfR- (+/- HIS tag). Previous literature data monitoring huH uptake using confocal experiments states that a minimum of 1 hr is required for huH to be internalised entirely by cells, and incubation times > 24 hrs lead to a reduction in huH signal, possibly due to huH degradation in the lysosome.[248, 249] Taking this into account, each apoFt variant was incubated for 1 hr before the first imaging time point.

A preliminary control test with each non-conjugated apoFt variant (huHC and TfR-, both +/- HIS-tag) indicated no adverse morphological changes to SKBR3 and MRC- 5 cells post 1hr. Following this, SKBR3 cells were incubated with each conjugated apoFt variant, and images acquired showed a high number of SKBR3 cells with huHC- and huHC- (-HIS) within their cytoplasm compared to TfR- mutant and TfR- (-HIS) (Figure 3.13). The significant reduction in cellular endocytosis by the TfR- variant supports the notion that mutations, I80A, F81L and L82F severely hinder huH-TfR1 interaction. Experiments with MRC-5 cells (which express TfR1 below detection capabilities) showed no/very little internalisation by all four variants, supporting the hypothesis that huHC- uptake is TfR1-specific (SFigure 3.6). In addition to comparing the uptake between TfR- and huHC-, the effects of the HIS-tag on apoFt uptake were also monitored using confocal microscopy. Neither the presence nor absence of the HIS-tag promoted nor reduced uptake of huHC- and TfR- in SKBR3 cells (Figure 3.13). This suggests that the HIS-tag cannot bind strongly to the cell membrane or is too flexible to form multiple attachments to the cell membrane to promote cellular uptake.
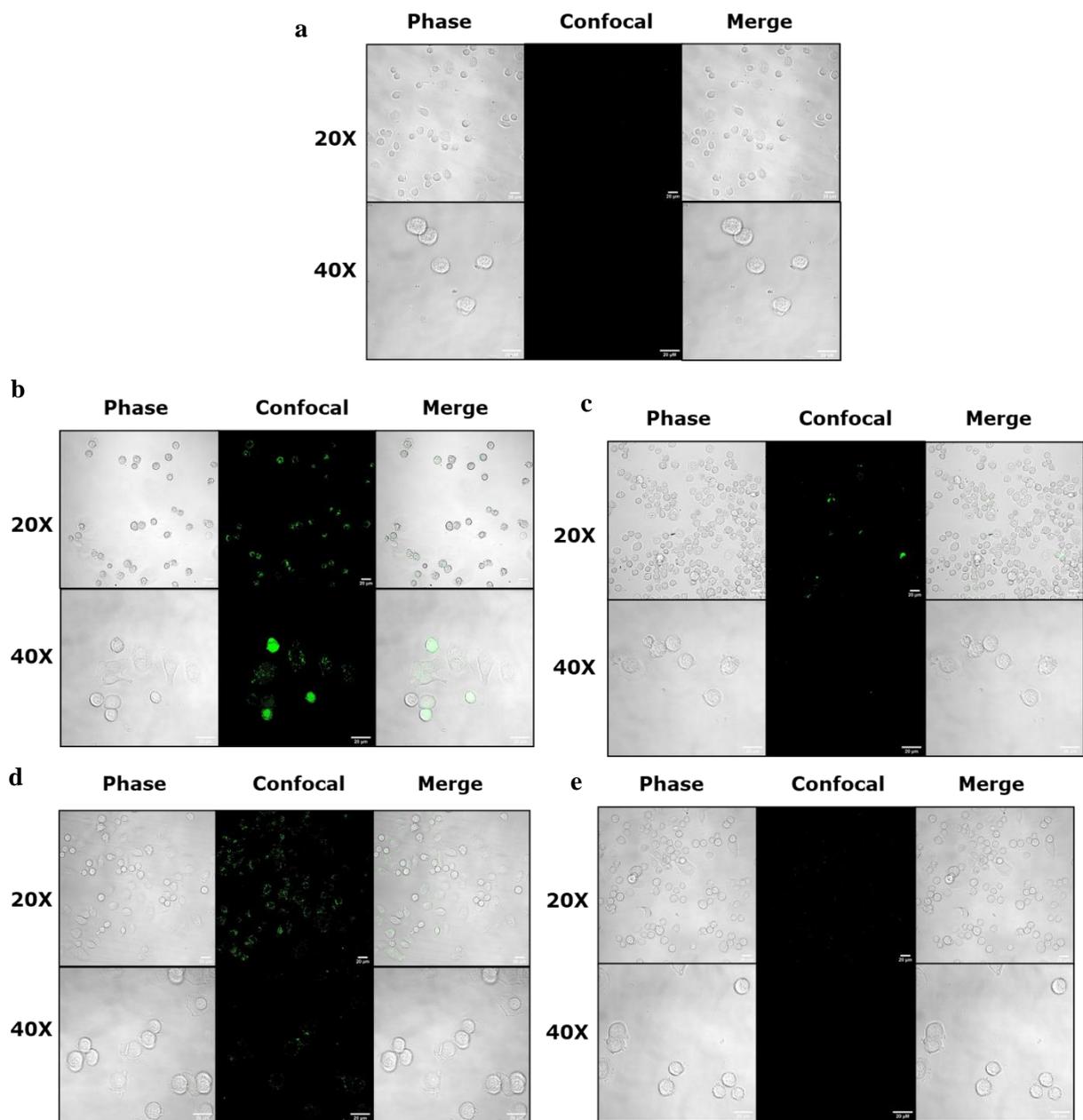
Figure 3.13 Internalisation of apoFt variants by SKBR3 cells. SKBR3 cells were incubated with 60 μg/ mL of apoFt for 1 hr. Images were taken at X20 and X40 under confocal (488 nm channel) and phase-contrast microscopy. Scale bar 20 μm. **a,** Control (no treatment), **b,** huHC-, **c,** TfR-, **d,** huHC- (-HIS) and **e,** TfR- (-HIS).

As confocal images confirmed the reduction of apoFt uptake by mutant TfR-, flow cytometry was employed to allow fast quantification to corroborate this observation. Flow cytometry analysis of SKBR3 cells showed a ~70% uptake for huHC-. A ~70% uptake is expected, given that not all the cells will be able to uptake apoFt (e.g., different stages of the cell cycle). A similar level of uptake has been previously reported with huH and TfR1; therefore, ~70% is used as the maximum threshold for apoFt uptake.[93, 94] In comparison, a significant reduction in uptake was seen in cells treated with TfR-, reducing from ~70% to <10% (Figure 3.14). A complete reduction in uptake was not seen for the TfR- mutant, probably because of partial unspecified uptake, e.g., pinocytosis. In addition, comparing the uptake of huHC- (~70%), which contains a HIS-tag and huHC- (-His-tag) (~70%), it was seen that the presence of HIS-tag did not increase or decrease the level of uptake. This supports confocal microscopy data, which also showed no effect of HIS-tag on TfR1 mediated uptake.
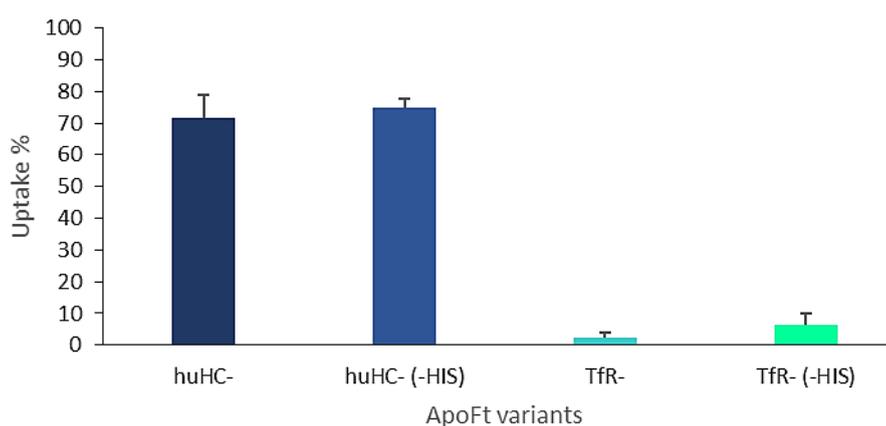
Figure 3.14 Uptake of apoFt variants by SKBR3 cells. Quantification of cells that have internalised different ferritin variants in SKBR3 cells. 10, 000 cells were sampled. 3 independent repeats (N=3) and 3 internal repeats (n=3) were tested SKBR3 cells.

While this work was being performed, a cryo-EM structure of the huH–TfR1 complex was published.[94] In response to the cryo-EM data, Montemiglio *et al.*, 2009 created Mutant A. Mutant A was created based on the cryo-EM data, which showed that the BC loop interacts with TfR1 using specific residues: R79, F81, Q83, K86, K87. Mutant A, which harboured mutation along the external BC-loop (F81A/Q83A), showed reduced cell uptake. The TfR-mutant created in this work showed a similar reduction in uptake to those reported by Montemiglio *et al.*, 2009. Comparison between the huHC- : TfR1 model to TfR- : TfR1 model

shows that the presence of F82 in TfR- prevents Q83 from forming hydrogen bonds with N348 on TfR1 (Figure 3.15). A similar effect was also caused by F81A, in mutant A. For this reason, it can be suggested that the hydrogen bond formed by Q83- N348 is very important in promoting TfR1 mediated endocytosis, with the remaining residues playing a part in stabilising the huH- TfR1 interaction. Given these promising results, the TfR- mutant was next employed to test the effects of a secondary protein or targeting agents, e.g., a HER2 Afb on binding and uptake via TfR1. This is an essential consideration as it will determine whether the HER2 Afb is the sole targeting agent or if dual uptake of huHC- Afb occurs by both TfR1/HER2 receptors. In addition, it may also establish whether the TfR- mutant can promote HER2-mediated uptake only.
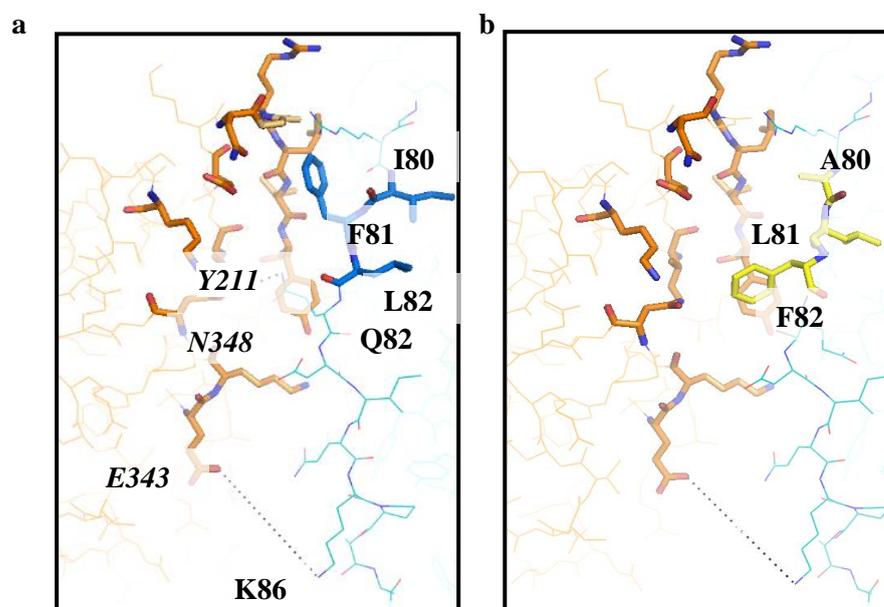
Figure 3.15 View of TfR1 (orange)-ferritin (blue) interaction region. **a**, Stick view of huHC- interaction with TfR1 and **b**, Stick view of TfR- interaction with TfR1. Predicted interacting region on HC- (IFL) (Blue), mutated region on TfR- (ALF) (Yellow) and Y211 on TfR1 βII-2 and Q83 with N348 αII-2 on TfR1 (Orange). Important interactions have been marked between αII-2 helix of TfR1 and the BC loop of huHC-: N348 to Q83 and E343 to K86 respectively. Hydrogen bonds: (---).

3.1.6 Modelling apoferritin-affibody variants

As shown in section 3.1.5, the TfR- mutant reduces uptake in TfR1-expressing cells. However, it was unknown if the presence of an N-terminally engineered ligand such as an EGFR-targeting Afb peptide could affect uptake in TfR1-expressing cells. Given the large structure protruding out of an EGFR-targeting Afb there was a probability that the presence of an Afb could hinder TfR1 binding. One such ligand is the ZHER2:342 HER2 Afb, which has its pico molar binding affinity (Kd: 22 pM) for the HER2 receptor and high thermal stability, as outlined in section 1.4.2. Given its therapeutic benefits, the ZHER2:342, which will henceforth be referred to as Afb, was genetically engineered onto huHC- by previous and current group members (Zhang L, 2015). Before the effects of the huHC- N-terminal fused ZHER2:342 Afb (Afb-huHC-) on TfR1 interaction could be studied *in vitro*, a national lockdown was put into effect due to the COVID-19 pandemic. Preliminary *in vitro* studies would have included confocal microscopy to visualise uptake of Afb-huHC-and surface plasmon resonance experiments to study the molecular effects of Afb on huHC-/TfR1 interaction. However, given the lack of experimental work during lockdown, some basic computational work was carried out to understand if the presence of the Afb would affect huHC- binding to TfR1. Using MODELLER and PyMol, a 3D homology model was created to visualise the distribution of the Afbs on the surface of the huHC-, as shown in Figure 3.16a-c. The diameter of the apical domain of a single TfR1 monomer was also used to evaluate the spatial environment for TfR1 entry and access to the three key residues (ALF) on the BC loop of huHC- Afb (Figure 3.16d). Only the apical domain of TfR1 was used, given that previous literature data from TfR1 binding studies have shown that the apical domain of the TfR1 receptor is a crucial area for protein binding, including huH and arenaviruses.[250, 94]

Using PyMol, it was seen that the flexibility of the linker region (VDGGGGSGGS) promoted distance variation between each Afb subunit, ranging between 55-60 Å (Figure 3.16 a- c). This flexibility could aid TfR1 entry to the BC- loop, as the Afbs would be able to move upon TfR1 entry. In addition, the 3D model created also showed that the Afbs were only distributed around the three-fold axis and the 2-fold axis (Figure 3.16a,b). Afbs localisation around the 3-fold axis is expected, given that the N-terminus of each huHC- subunit protrudes from the 3-fold channel, and the Afbs are fused to the N-termini of each huHC- subunit. In comparison, the four-fold axis area is mainly unoccupied, with a diameter of 65-70 Å between each Afb (Figure 3.16c). Therefore, it can be hypothesised that TfR1's apical domain, with a diameter of 57-62 Å (Figure 3.16d), would experience steric hindrance around the 2- and 3- fold axis, affecting access to the

BC-loop. However, TfR1 would be able to access the BC-loop via the 4-fold axis, as schematically shown in Figure 3.17a-c. Considering this and the fact that there are only six four-fold axes, plus 24 possible TfR1 binding sites on huHC-, it can be presumed that an Afb would decrease huHC-TfR1 binding from 100% to 25%. Therefore, for any future *in vitro* uptake experiments involving 50:50 uptake via both TfR1 and the HER2 receptor, it would be preferable to have 12 affibodies per cage. This adjustment would free up to four 3-fold channels, as more of the BC-loop would be accessible to TfR1, thereby increasing the probability of TfR1 binding to huHC- from 25% to 50%.
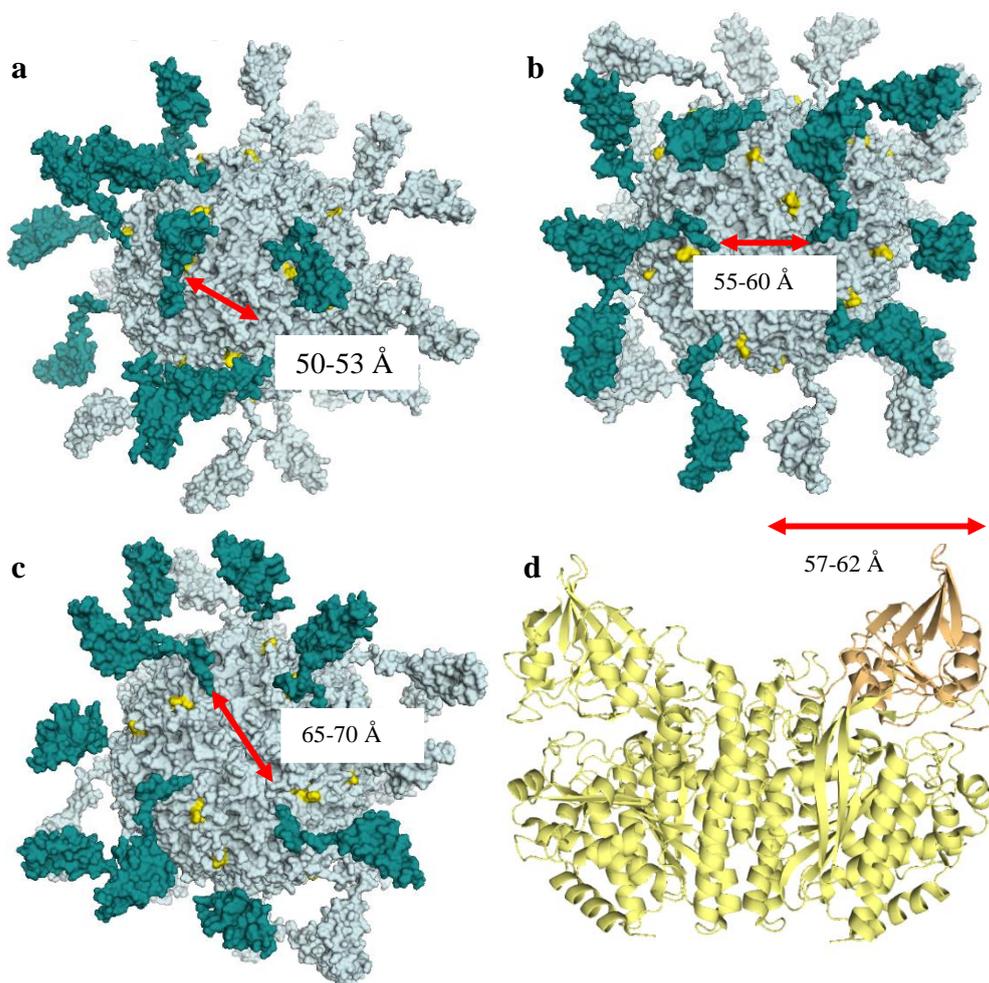
Figure 3.16 Computational modelling of huHC-Afb. **a**, 2-fold view, **b**, 3-fold view, **c**, 4-fold view of huHC- Afb and **d**, TfR1. Red arrows indicate the maximum distance available in each space, as measured by PyMol. Extracellular domain of TfR1, apical domain of TfR1, huHC-, Afb and TfR- mutation (isoleucine, phenylalanine, and leucine) areas are coloured yellow, orange, light blue, green and dark yellow, respectively. Modelled after PDB: 2ZKI (Affibody) and 6GSR (huHC-).

The effect of charge distribution on both TfR1 and Afb subunits was also considered because a large positive or negative charged area on the Afb could affect the entry of TfR1. The charge distribution on both TfR1 and huHC- Afb was calculated at pH 7.4, as this reflects physiological pH outside the cell membrane. At this pH, the Afb is shown to have both negative and positive charged regions, but neither one dominates the surface (Figure 3.17). In comparison, TfR1 is predominantly positively charged in the apical domain, with negatively charged areas surrounding the apical domain. Although the regions surrounding TfR1's apical domain do not have any key apoFt binding residues, the negative regions could encounter electrostatic repulsion when in close proximity to the Afb's negative regions, preventing TfR1 from binding. However, given the flexibility of the Afb, as a result of the flexible linker, the electrostatic repulsion might only be momentary,

Considering both the homology modelling and electrostatic data, it can be suggested that TfR1 might not completely be hindered by the presence of Afb on huHC-. Therefore, *in vitro* studies would need to be performed to verify and quantify the effects on huHC- Afb uptake against breast cancer cell lines with varying degrees of TfR1/HER2 expression, including HER2+ BT474 HER2+ SKBR3, HER2- MDA-MB-231 and HER2- MDA-MB-231.

,

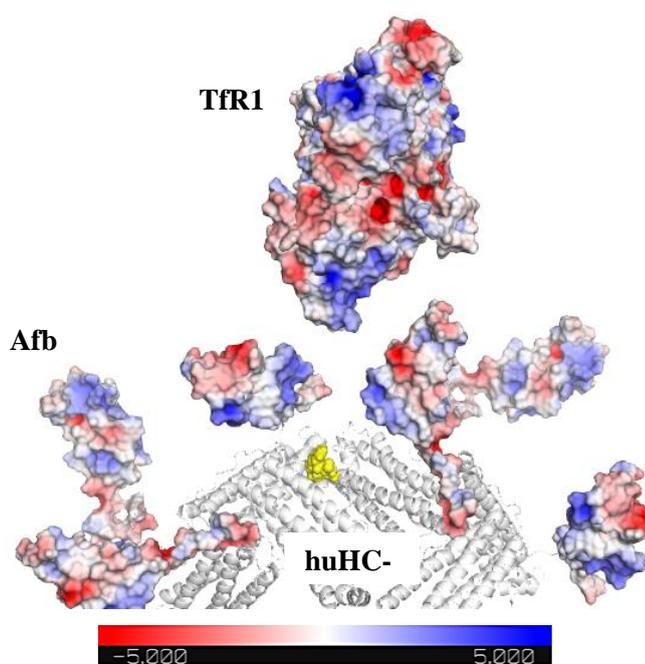**Figure 3.17** Electrostatic potential of TfR1 monomer and an huHC-Afb at pH 7.4. TfR-mutation (IFL) is coloured yellow and the positively charged and negatively charged regions on the TfR1 monomer and Afb are coloured in, blue, and red, respectively.

3.1.7 Effect of an affibody on apoferritin uptake

To evaluate the effects of an Afb on huHC- uptake via TfR1, three Afb constructs were purified: huHC- Afb, Afb and TfR- Afb. huHC- Afb and TfR- Afb (SFigure 3.7) All three constructs contained the A102C mutation to allow 5-FAM conjugation. In addition, a cysteine was introduced on the exterior of both Afb-apoFt variants. An external cysteine was selected over an internal cysteine as the flexible linker region around the 3-fold channel could have hindered the uptake of 5-FAM. The Afb-only construct contained the S40C mutation, located on the flexible loop region between the second and third helix. Following protein purification, SDS-PAGE analysis showed a MW of ~30 kDa for both huHC-Afb and TfR-Afb, consistent with their theoretical calculated MW of 31.7 kDa and 31.6 kDa, respectively (Figure 3.18). In comparison, the Afb-only construct migrated just below the 15 kDa mark, which is not consistent with its theoretical MW of 9.8 kDa. However, this slow SDS migration has been previously reported with other Afbs and is caused by their thermal resilience and ability to refold even after 10 mins at 100 ºC, leading to a non-linear migratory protein.[91, 92] Once the constructs were purified, the ammonium sulphate precipitation method was used to conjugate 5-FAM to all three constructs. Native PAGE analysis showed that huHC-Afb and TfR-Afb maintained their cage integrity even after 5-FAM conjugation (Figure 3.19a). SDS-PAGE also confirmed conjugation of 5-FAM to Afb (Figure 3.19b). Unfortunately, due to the coronavirus lockdown, mass spectrometry for the Afb containing constructs could not be obtained, and therefore the conjugation efficiency could not be calculated.
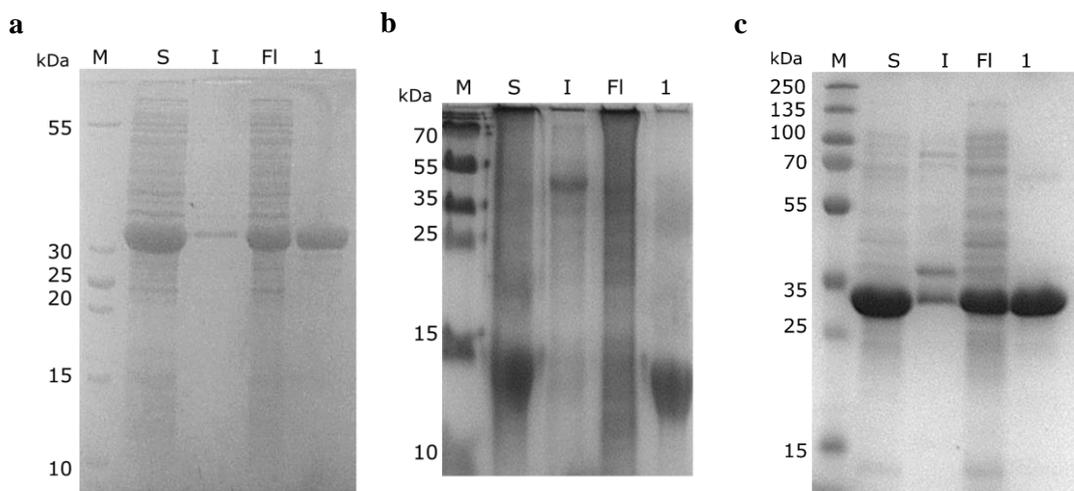
Figure 3.18 Expressions of Afb-containing variants. **a**, huHC- Afb A102C, **b,** Afb S40C and **c**, TfR- Afb A102C. M: Marker, S: Soluble, I: Insoluble, Fl: Flow through and 1: Fraction 1.
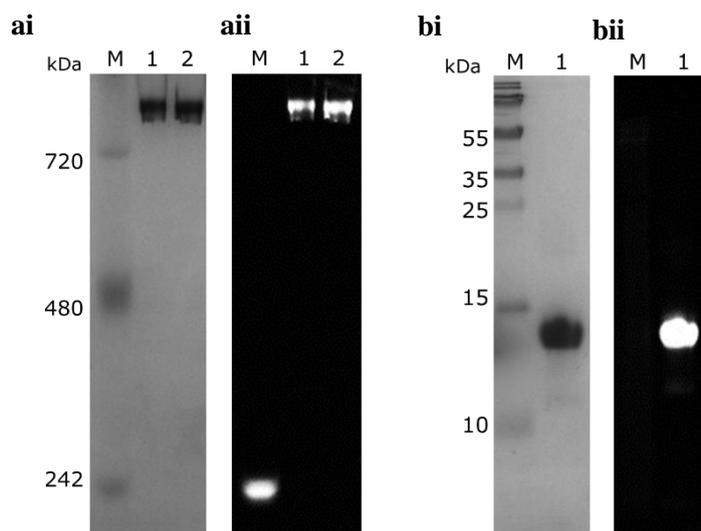
Figure 3.19 Native PAGE of intact affibody containing variants after conjugation to 5-FAM. **a** M: Marker, 1: huHC- Afb A102C and 2: TfR- Afb A102C and **b** M: Marker and 1: Afb S40C. **ai** and **bi**: Native PAGE of affibody constructs under normal light. **aii** and **bii**: Native PAGE of affibody constructs under UV light.

Once purified, all three Afb-containing constructs (huHC- Afb, Afb and TfR- Afb) and those created in section 3.1.3 (HuHC- and TfR-) were carried forward to assess the effects of Afb-HuHC- on TfR1 uptake. Uptake was quantified by flow cytometry in five cell lines, each possessing varying degrees of TfR1/HER2 expression (Figure 3.20a).

SKBR3, which has been shown to have high HER2 and TfR1 receptor expression levels,

showed uptake of ~60% for huHC-. A 60% uptake indicates that after incubating SKBR3 cells with huHC- for 1 hr, only ~2/3 of the population either have up taken huHC- or have not yet degraded huHC-. Therefore, using ~60% as the maximum for cellular uptake and the information acquired from the modelling data in section 3.16, it was calculated that the presence of the Afb would reduce TfR1-mediated uptake of huHC- Afb to ~15%. However, this was not the outcome with SKBR3 cells treated with Afb- huHC-, whereby the level of uptake was still at ~60% (Figure 3.20b). This is expected given that the Afb segment of the Afb-huHC-construct could permit endocytosis by the HER2 receptor. This is further supported by data showing that the TfR- Afb only had an uptake value of 30%, even with the TfR- mutation. This suggests that the TfR-Afb has undergone receptor-mediated endocytosis via HER2 on SKBR3 cells. When Afb-containing constructs including Afb, huHC-Afb and TfR-Afb were tested on MDA-MB-231 cells, a considerable reduction in uptake (> 40%) was seen compared to their counterparts in SKBR3 cells. This reduction is caused by the low HER2 receptor number in MDA-MB-231 cells. In addition, it was seen that huHC- Afb had a ~10% higher uptake than TfR- Afb and Afb. This is possibly caused by huHC- Afb being endocytosed by TfR1 and HER2 on MDA-MB-231. As quantification studies confirmed a lower HER2 copy number for MDA-MB-468 compared to MDA-MB-231, it was chosen to confirm the effects of huHC- Afb on uptake via both the TfR1 and HER2. huHC- showed a similar uptake level in MDA-MB-468 compared to SKBR3 and MDA-MB-231 (~60%), owing to the similar levels of comparable TfR1 expression in both cell lines. In comparison, a reduction of <~10% was seen both Afb and TfR- Afb for MDA-MB-468 cells, indicating the primary uptake in MDA-MB-468 cells was only by TfR1. HuHC- Afb was also incubated with MDA-MB-468 cells, resulting in a 20-30% uptake. These results support the theory outlined in section 3.1.6, whereby the presence of the Afb on the huHC- hinders TfR1 binding but does not completely abolish it.

Three constructs (huHC-, Afb and TfR- Afb) were also tested on BT474 cells (Figure 3.20c). This cell line demonstrates similar HER2 and TfR1 expression to SKBR3 but has been shown to develop spheroids for use in 3D cell culture and is able to generate tumour xenografts for *in vivo* studies. All three constructs showed similar uptake levels to those seen in SKBR3 cells. In addition, the percentage uptake for Afb and TfR- Afb were similar in both MDA-MB-468 MRC-5 cells, with an uptake of <10%. This indicates that Afb-containing variants can only be internalised by cells possessing high HER2 expression and reinforces the Afbs specificity for HER2.
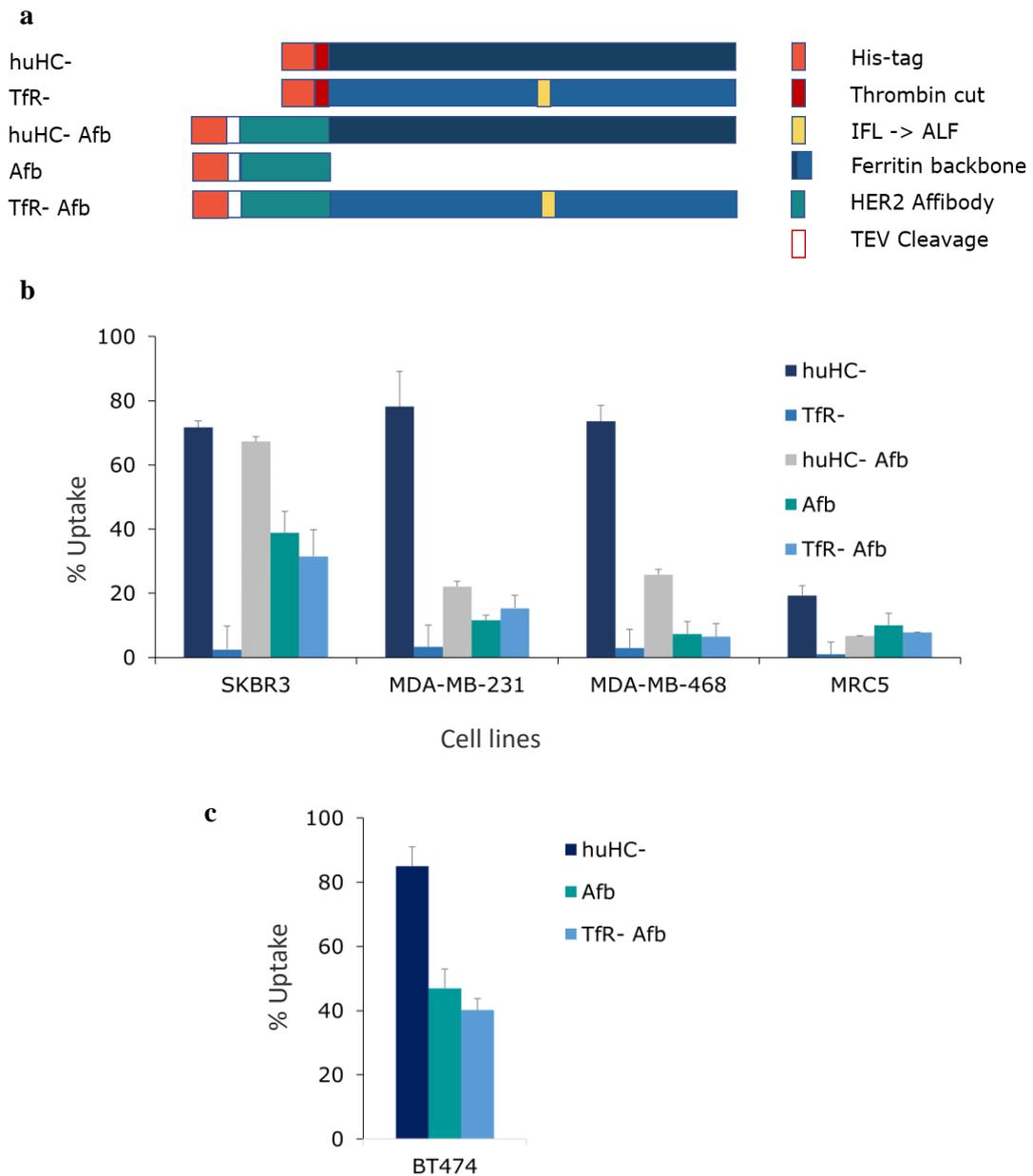
Figure 3.20 Quantifying different huH and Afb variants internalisation using flow cytometry. **a**, Schematic representation of huH and Afb variants, **b**, Mean + Standard deviation (SD) uptake of huH and Afb variants by breast cancer cells and **c**, Mean + SD uptake of huH and Afb variants by BT474 cells. Quantification was performed on 10, 000 cells per run. 3 independent repeats (N=3) and 3 internal repeats (n=3) were tested for each construct for each cell line .

3.2 Conclusions and future work

The conclusions that can be drawn from the data presented in this chapter are outlined below. Furthermore, some future considerations are also presented to improve and build upon the data acquired in this work.

1. Introduction of mutations E61A, E62K, E64A, H65G (huH 222) did not change the protein solubility during protein expression in *E.coli* and the secondary structure profile compared to that of huHC- as determined by CD. Although the mutations reduced the time taken for iron core formation, the formation time was still relatively rapid, reaching maximum saturation only 40 seconds after huHC-. As the Tris buffer could have affected the kinetic rate of iron core formation and is not a typical buffer for biological assays, kinetic studies could be performed in various buffers, including PBS, RPMI/MEM + 10% FBS and in different pHs, e.g., 6.5 and 7.4. Both PBS and RPMI/MEM + 10% FBS are standard buffer/media in which cells usually proliferate. Therefore kinetic studies performed in these conditions would elucidate if the rate of iron core formation would increase, especially as 10% (v/v) FBS has 2–5 μM iron.[253] Monitoring iron core formation at both pH 6.5 and pH 7.4 would also be necessary, as it would indicate if iron core formation would be altered in an acidic pH such as those surrounding cancers cells (pH ~6.5) compared to those surrounding normal cells (~pH 7.4). This will also show if iron core formation in the huH 222 mutant is affected under different conditions.

2. Introduction of additional mutations I70A, F71L and L72F (to give the TfR- mutant) did not change the variant's protein expression profile, but the presence of a K68C mutation resulted in the variant being expressed mainly in the insoluble fraction.

3. The TfR- mutant showed reduced uptake in SKBR3 cells compared to huHC- in confocal microscopy and flow cytometry studies. The presence of a HIS-tag on either TfR- or huHC- also had no effect on uptake in SKBR3 cells. To further confirm this observation, future analysis could quantify the ratio of non-fluorescent cells vs fluorescent cells captured by confocal microscopy. This type of analysis would account for any huHC- variants binding to the exterior of the cells vs those internalized. This distinction is not possible by flow cytometry as the technique cannot quickly differentiate between fluorescently labelled proteins that are internalised or surface-bound.

4. Initial modelling carried out during the COVID-19 lockdown showed that any future addition

of Afbs' on the surface of huHC- could affect TfF1 uptake. However, due to the limited computational training and expertise that could be acquired during the COVID-19 lockdown, the data acquired are limited and therefore needs to be further developed, and the parameters optimised. Parameters include understanding the molecular movement of the flexible loop region between huHC- and Afb using molecular dynamic stimulations with different force fields (CHARMM/AMBER). This is important because it will elucidate if the linker loop region is flexible or partially rigid, affecting the Afb's protrusion. In addition to optimising the huHC-- Afb model, protein-protein docking software, including ClusPro could also be used to analyse the attraction or repulsion of neighbouring Afb's and TfR1 as well as consulting experts in the field of computational biology.

5. *In vitro* testing using flow cytometry and breast cancer cell lines with various levels of TfR1/HER2 expression confirmed that the presence of Afbs on the surface of huHC- reduced uptake from ~70% to ~25%. Future work could include testing huHC- Afb and TfR- Afb in a breast cancer xenograft model to evaluate the effects of a TfR- mutation on the distribution and half-life of Afb variants. Introduction of the TfR- mutation could enhance tumour localisation of apoFt constructs, thereby reducing excretion and accumulation in organs, such as the liver and increasing the half-life of Afb variants. In addition, future work could also involve adding a small targeting or therapeutic peptide on the C-terminus or BC-loop of TfR- Afb. This addition would increase TfR- Afb's tumour targeting capabilities and increase its anti-tumour activity.

In summary, the data presented here highlight that removing the ferroxidase and TfR1 binding ability of huHC- could improve its capabilities as a NC. These initial findings provide the preliminary steps to develop huHC- further for not only breast cancer targeted delivery but also cancers such as prostate cancer, by addition of multiple targeting ligands.

4.0 Engineering ferritin for encapsulation of therapeutic proteins

4.1 Results and discussion

As outlined in chapter 1, apoFt's 24-mer cage consists of an 8 nm hollow core and 3-4 Å wide channels. Using apoFt's channels, therapeutics such as cisplatin can diffuse (nanoreact) and be encapsulated within apoFt's core. In addition to using the nanoreactor route, compound encapsulation can also be achieved by disassembling and re-assembling apoFt's cage using either the pH-based or the denaturant-based method. However, the nanoreactor and disassembly routes are restricted to small compounds (3-4 Å, in at least 2 of their dimensions) and compounds that can withstand the relatively extreme denaturation conditions, respectively. The extreme conditions of the disassembly/re-assembly route can also damage loads such as therapeutic proteins/peptides. There are many benefits to using therapeutic proteins/peptides for cancer treatment over compounds such as DOX, including their ability to bind to a larger surface area of the target protein than small molecules. Smaller molecules typically target hydrophobic pockets, which are typically buried within a protein. Therefore, this limits small compounds to proteins that have solvent-accessible pockets. In contrast, proteins/peptides are more flexible and can bind to a greater area of the target protein, increasing target specificity. Therefore, this decreases a therapeutic protein's off-target effects and increases the range of targetable proteins.[254]

The current chapter focused on apoFt's ability to encapsulate and deliver a therapeutic protein to breast cancer cells. Cyt C's cytotoxic characteristic and capacity to resist extreme pH enabled it to be selected as a candidate for encapsulation. [170, 172] Initial studies focused on using both the pH- and the urea-based disassembly method to encapsulate Cyt C within huHC-. Following this, huHC- was engineered to disassemble at a higher pH (>4.0) through the expression of the ΔDE variant. The ΔDE variant was created to reduce any possible denaturing effects Cyt C would encounter during the pH-disassembly route, requiring apoFt disassembly at pH 3.0-2.0. The pH disassembly points of huHC- and ΔDE were then verified using DLS and native PAGE. Unfortunately, subsequent *in vitro* work was hindered by the COVID-19 pandemic, and therefore basic computational modelling was performed on huHC-'s and ΔDE's 3- and 4-channels for ~6 months. Following this, Cyt C was encapsulated within ΔDE using the pH-based method, and the cytotoxic effects of Cyt C were evaluated against SKBR3 and MDA-MB-231 human breast cancer cell lines, and non-transformed MRC-5 human foetal fibroblast.

4.1.1 Cytochrome C encapsulation

Prior to this work, no therapeutic proteins/peptides had been encapsulated within huH. Among the selection of proteins available, Cyt C was identified as a potential candidate for multiple reasons, including its unique absorption spectrum. Cyt C's distinct spectrum is owed to its native structure, consisting of 4 α-helices with a central porphyrin ring. The presence of the porphyrin ring results in a distinct Soret absorption band at 410 nm when monitored using UV-vis (Figure 4.1a). Using Cyt C's unique 410 nm band and extinction coefficient of 106 mM^{-1}·cm^{-1}, it would be possible to quantify the molar concentration of Cyt C independent of the standard protein A280 method (Figure 4.1b).[255] In addition to the quantifiable nature of Cyt C, another benefit of using Cyt C is its pH and denaturant stability. Cyt C, which has a pI of 9.5, is stable across a range of conditions, retaining a 50% folded state at pH 2.5 and in 8 M urea solutions at pH 9.0.[169] This characteristic would enable Cyt C to retain its structure following either the urea- (8 M urea) or pH-based encapsulation.

As the outcome of this chapter was to test huHC- capabilities to deliver therapeutic proteins, it was necessary to encapsulate a protein that could produce a cytotoxic effect in breast cancer cells. Cyt C met this criterion, with previous literature data showing that breast cancer cells, such as ER+ (MCF-7), triple-negative (MDA-MB-231) and ER+/HER2+ (BT474), are hypersensitive to injected Cyt C, leading to enhanced caspase activity and apoptosis.[173, 171] As such the ability of Cyt C to promote apoptosis would provide a functional readout on huHC-'s capability to deliver cytotoxic proteins.

Given that Cyt C is both quantifiable and cytotoxic towards breast cancer cells, it was carried forward for encapsulation studies. An initial standard curve was produced using series Cyt C concentrations (5 µM - 80 µM), and their respective 410 nm peak intensities were plotted, as shown in Figures 4.1a,b. The standard curve outlined the linear relationship between Cyt C concentrations and 410 nm absorption and therefore was used as a reference to calculate the number of encapsulated Cyt C molecules.
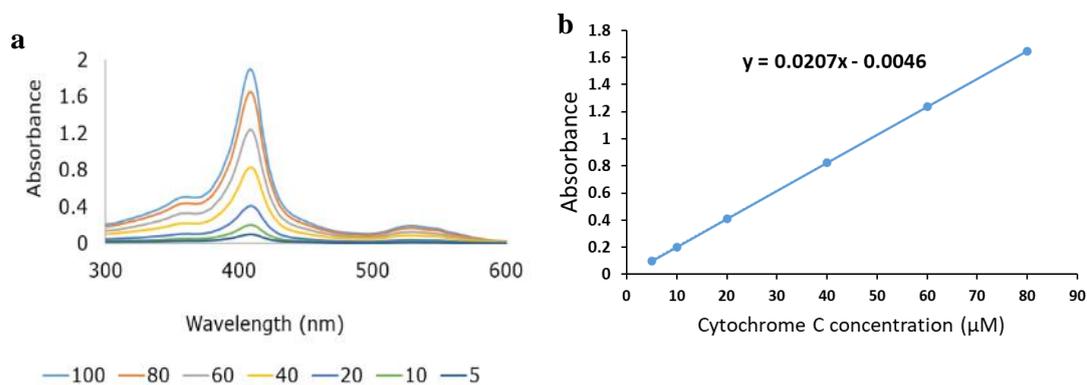
Figure 4.1 a. Determination of Cyt C concentration. **a**, UV-vis spectra (300 nm – 600 nm) of Cyt C at different concentrations (5-100 μM) and **b**, Linear regression plot for Cyt C concentration vs absorbance at 410 nm ($R^2 = 1$), using an extinction coefficient of 106 mM^{-1}.cm^{-1}.[255]

Following the production of a Cyt C standard curve, two methods were used to encapsulate Cyt C in huHC-: the 8 M urea and the pH disassembly methods. Both the urea and pH disassembly methods yielded ~3 molecules of Cyt C per apoFt cage (Figure 4.2). This value is low compared to the maximum theoretical value of ~5 molecules per cage, which was calculated based on an inner 8 nm diameter of huHC- cavity and a 3.9 nm diameter of Cyt C, with a 75% packing capacity. The reduction in encapsulation efficacy could be owed to Cyt C's positively charged surface at pH 8.9, leading to electrostatic repulsion between neighbouring encapsulated Cyt C molecules (Figure 4.3). Consequently, this would increase the space between neighbouring Cyt C molecules and decrease the total volume available within huHC-.
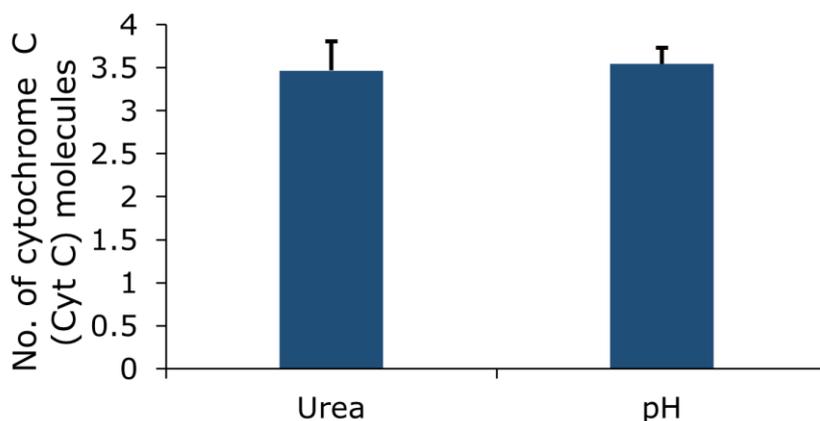
Figure 4.2 Number of encapsulated Cyt C molecules using either the 8 M urea-based method or the pH-based method. N=3 independent trials were performed. Mean ±SD n= 3.

113
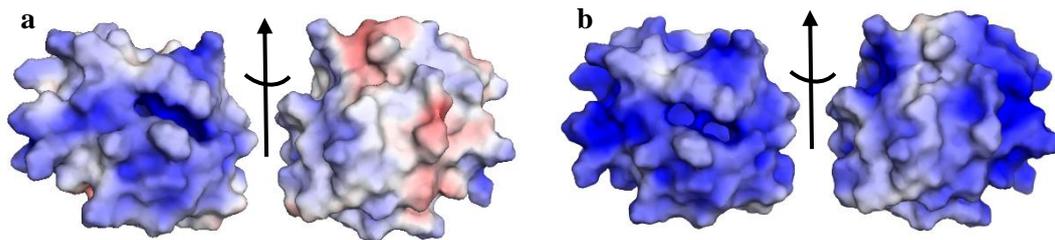
Figure 4. 3 Surface charge distribution of Cyt C at different pHs values. **a**, pH 8.9 and **b**, pH 2.5. Positively charged residues (blue) and negatively charged residues (red). PyMol: 2B4Z

Once Cyt C encapsulation was confirmed, native PAGE analysis was performed to determine if huHC-'s tertiary structure had been altered by post encapsulation. huHC- and huHC- with Cyt C encapsulated samples retrieved from the urea-based method showed similar migration bands, indicating that huHC-'s architecture post encapsulation had not been altered, as shown in Figure 4.4a. In comparison, the pH-based encapsulation method seemed to have affected huHC- structure, with huHC- with Cyt C encapsulated, showing a less defined band at 480 kDa and another faint lower band, as shown in Figure 4.4b. The appearance of a lower molecular weight band could be attributed to an incomplete huHC- cage. Literature data that used synchrotron small-angle X-ray scattering to analyse the effect of pH-based disassembly/re-assembly of Ft reported that upon re-assembly, some Ft cages contained a hole defect the size of a single subunit.[256] In addition, given that the global structure of huHC- would change with a missing subunit, it can be suggested that the pI would also be altered, thus affecting the migration of the defective huHC- on native PAGE compared to native huHC-. Therefore, it can be theorised that the lower band seen in Figure 4.4b is an incomplete huHC- cage. However, this theory needs to be further investigated using techniques such as native mass spectrometry. Nevertheless, the incomplete cage with a missing subunit (4.8 nm length x 1.9 nm width) would still have the capacity to encapsulate and retain Cyt C (3.9 nm, globular). Hence, the pH-based method would give a similar encapsulation efficiency as the urea-based method (Figure 4.2).

Although the urea- and pH-based methods enabled Cyt C encapsulation, the extreme disassembly conditions, e.g. 8 M urea and pH 2-3, could still denature some Cyt C molecules. In addition, most therapeutic peptides/proteins are not as resilient to such extreme pH conditions and would likely irreversibly denature under these conditions. On that account, the next step involved engineering huHC- to disassemble at a higher pH.
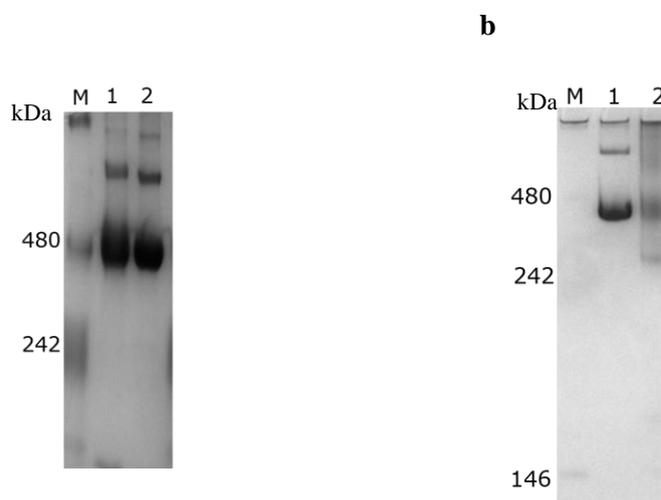
Figure 4.4 Native PAGE analysis of Cyt C encapsulation. **a**, Urea based encapsulation (1: Cyt C encapsulated and 2: huHC- only) and **b**, pH-based encapsulation (1:huHC- only and 2: Cyt C encapsulated). M: Marker.

4.1.2 Cleavage of apoferritin's DE loop and the E-helix

Given that a very acidic pH of pH 2-3 is necessary for huHC- disassembly and Cyt C has been shown to partially unfold at pH 2-3; there is a risk that some of the encapsulated Cyt C would be functionally inactive.[169] To account for this, the current work sought to engineer huHC- to disassemble at a higher pH. Previous literature data indicated that removal of the DE loop and E-helix increased apoFt's disassembly pH.[6, 7] Therefore, a stop codon was introduced at the beginning of the DE loop at position 160 of huHC-, as shown in Figure 4.5a. The new variant was termed Δ160. Small scale tests showed an expression of Δ160 in the insoluble fraction (Figure 4.5b). However, during large scale overexpression, the insoluble fraction could not be solubilised in either 8 M urea or 6 M guanidinium hydrochloride solution. As a consequence, the Δ160 variant could not be purified. The presence of Δ160 in the insoluble fraction contradicted previous literature, which showed the expression of a similar construct in the soluble fraction.[6, 8] The only notable difference between the construct stated in the literature and Δ160, which could affect solubility, was the presence of an N-terminus segment, consisting

of polyhistidine tag (HIS-tag), an avidin-tag (Avi-tag) and a thrombin cleavage site, as shown in Figure 4.5a.

Figure 4.5 Synthesis of Δ160. **a**, Schematic representation of Δ160 primary structure. Orange label: polyhistidine tag and avidin tag, Red label: thrombin tag and green boxes : the location of an α-helix and **b**, Expression of Δ160. M: Marker, S: Soluble, In: Insoluble, Fl: Flow though, 2.5: 2.5% of buffer B and 40: 40% of buffer B.

To elucidate the effects of the N-terminus segment (+HIS) on huHC- stability, native PAGE and pH precipitation studies were performed, as shown in Figure 6. Both huHC-(-HIS) and huHC-variants showed similar disassembly profiles, producing bands at 480 kDa, between pH 8.0 and 4.0. Migration bands at 480 kD indicate that huHC- and huHC-(-HIS) had retained their globular structure and not disassembled into their constituent subunits. However, between pH 3.0 and pH 2.0, multiple bands are visible below 480 kDa for huHC- and huHC-(-HIS), indicating that both variants had undergone stepwise protein disassembly.[258] This suggests that the presence +HIS region does not affect pH disassembly. However, during pH precipitation studies, only huHC- was shown to precipitate between pH 7.5 and pH 5.0, as shown in Figure 4.6c, d. The precipitant is likely caused by huHC- aggregating rather than denaturing, as there were no bands visible below 480kDa on native PAGE and a denatured huHC- would constitute of subunits rather a cage. However, an equivalently darker 720 MW band was present for huHC- (-HIS), indicating that huHC- had formed dimers or higher molecular weight aggregates. Considering these observations, it can be suggested that the histidine-tag (pKa of 6.0), which consists of polar residues, is unlikely to be the source of huHC-'s pH instability but rather the avidin-tag which is comprised of several hydrophobic residues. Therefore, it can be hypothesised that when the pH draws close to pH 6-7, the histidine-tag would have a neutral

charge and add to the hydrophobicity of the avidin-tag region. This combined effect of the histidine-tag and avidin-tag at pH 6-7 could be more pronounced in Δ160, which has a theoretical pI of 6.1, leading to the expression of Δ160 in the insoluble fraction during expression. Therefore, a new Δ160 construct without the N-terminal segment was designed. However, it should be noted that new purification steps would also need to be considered, given the absence of the histidine-tag, a region required for nickel-based purification.
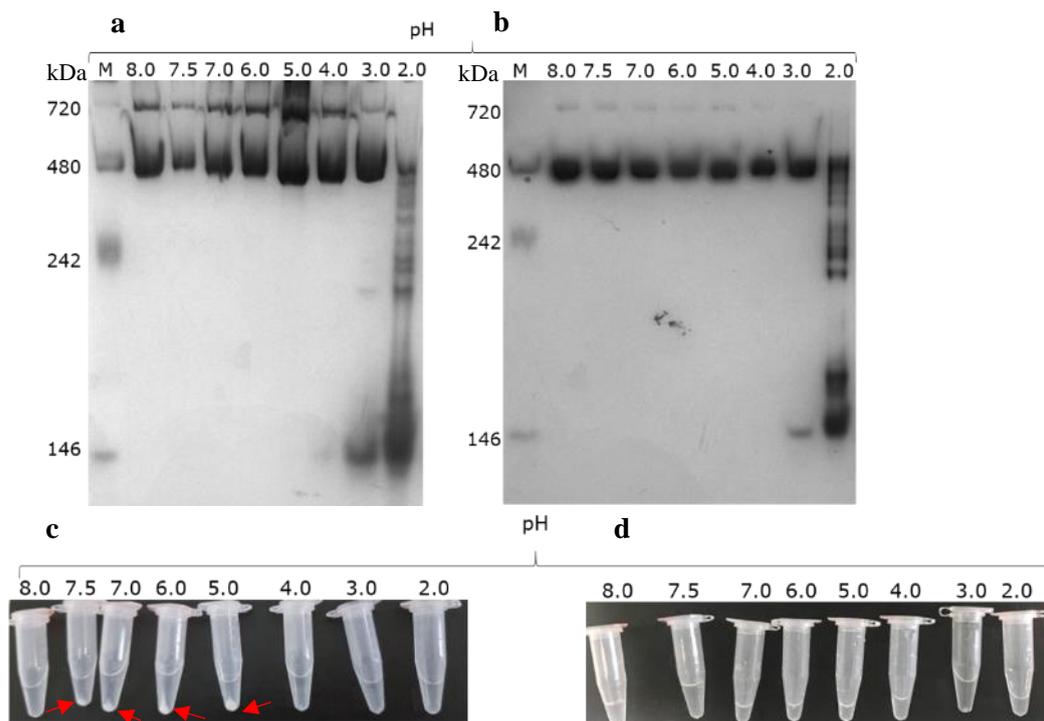
Figure **4.6** huHC- stability in different pH conditions. Native PAGE of huHC- disassembly. **a**, huHC- without thrombin cleavage and **b**, huHC- (-HIS) with thrombin cleavage. pH 2.0 (50 mM potassium chloride), pH 4.0/5.0/6.0 (50 mM sodium acetate) and pH 7.0/7.5/8.0 (20 mM Tris). Visual view of huHC- without thrombin cleavage and d. Visual view of huHC- with thrombin cleavage. Red arrows indicate insoluble huHC-.

4.1.3 Production of E-helix truncated apoferritin

As the presence of the histidine- and avidin-tag affected the solubility of Δ160 between pH 6-7.5, a new truncated version of Δ160 was synthesised. The variant was termed ΔDE, and its sequence is schematically shown in Figure 4.7a. The sequence also contained the ferroxidase

and TfR- mutations created in chapter 3. ΔDE was overexpressed in both the soluble and insoluble fractions when expressed for 16 hrs at 20 °C (Figure 4.7b). Given the absence of the histidine-tag, which is necessary for nickel-based purification, multiple purification methods were trialed, including thermal denaturation, gel filtration and ion-exchange chromatography, as shown in Figure 4.8a. A heat-based purification method was initially employed because apoFt is highly stable at 75 °C for 10 mins.[104] Therefore, it was hypothesised that a thermal denaturation method would aggregate and remove a significant number of impurities from the soluble fraction, leaving the intact apoFt cage in solution. Heat denaturation at various temperatures, including 10 mins at 55 °C and 60 °C, were investigated to purify ΔDE. However, the heat-based method led to low apoFt yields and high impurities and was consequently omitted from the final protocol. The final product was purified using only gel filtration and ion-exchange chromatography. SDS-PGE analysis of purified ΔDE showed a MW of ~15 kDa, ~3 kDa lower than the predicted theoretical MW of 18 kDa, as shown in Figure 4.8a. One possible explanation for this could be the SDS-PAGE gel percentage. Previous literature indicates that resolving a lower molecular weight protein (<~20kDa) with high percentage gels, e.g. 20%, results in a faster protein migration rate than low percentage gel, e.g. 15%.[259] However, to further confirm that the purified product was ΔDE, native PAGE and DLS analyses were performed. Both techniques would corroborate the presence of a cage-like structure, which is unique to apoFt variants. Native PAGE analysis showed a faster migration rate for ΔDE, which resulted in a band below 480 kDa, the MW for huHC-. This shift in migration is likely caused by ΔDE's high pI (pI 7.11) compared to huHC-'s (pI 5.74) (Figure 4.8b), and given that native PAGE resolves protein structures based on both size and overall charge, it is likely that ΔDE's high pI affected its migration. However, there was also a possibility that ΔDE had formed a smaller-size cage structure. Hence DLS was employed to confirm and analyse the size of ΔDE.
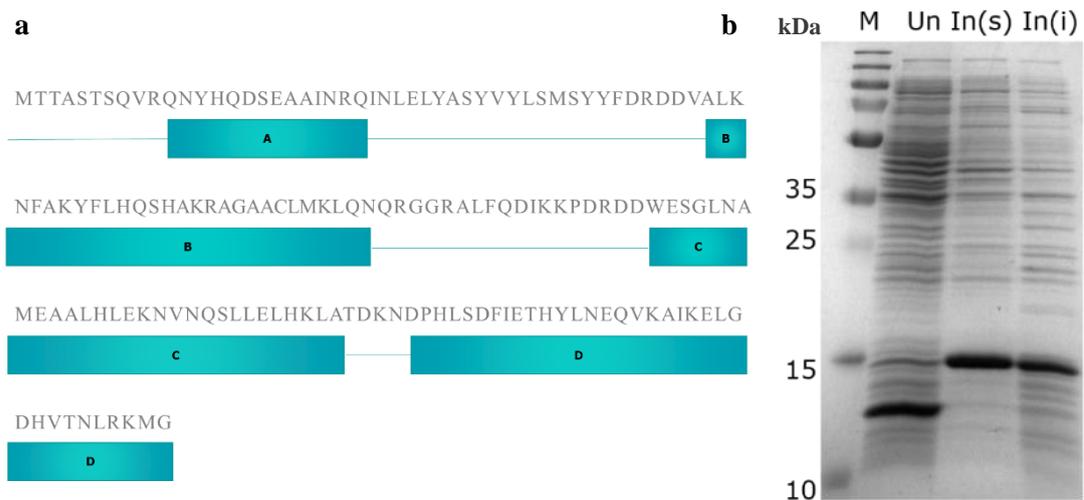
Figure 4.7 Synthesis of ΔDE. **a**, Schematic representation of ΔDE primary structure and **b**, Expression of ΔDE. M: Marker, Un: Un-induced, In(s): Soluble and In(i): Induced soluble. Expressed for 16 hrs at 20 °C with 0.5 mM IPTG.
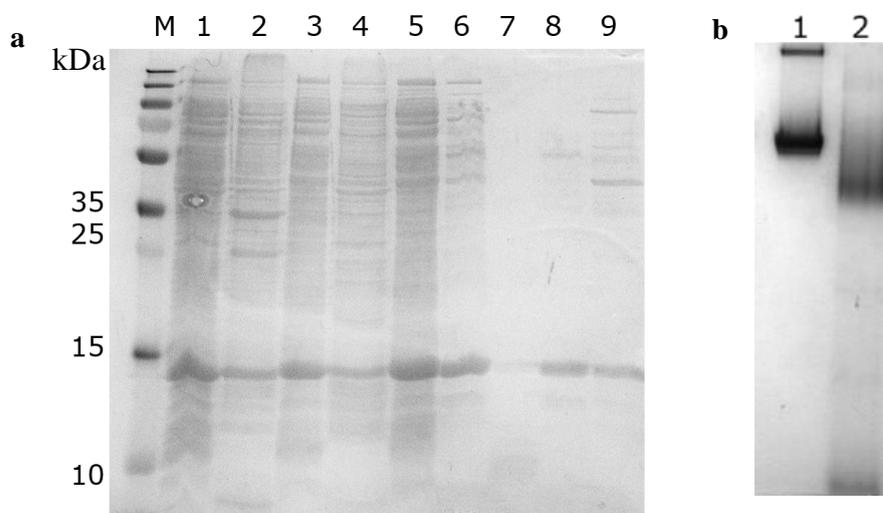
Figure 4.8 Purification and tertiarystructure of ΔDE. **a**, Purification of ΔDE (M:Marker, 1: Soluble, 2: Insoluble, 3: Heat treatment soluble fraction, 4: Heat treatment insoluble fraction, 5: 70% ammonium sulphate, 6: Gel filtration, 7: Ion exchange flow though, 8: 15% ion exchange buffer B and 9: 30% ion exchange buffer B) and **b**, Native PAGE of huH (1: huHC- and 2: ΔDE).

4.1.4 Confirming ΔDE pH disassembly point

As native PAGE could not verify if ΔDE had retained a globular structure similar to huHC-, DLS was utilised. Unlike native PAGE, which employs both size and charge to resolve the MW of a protein, DLS uses the protein's hydrodynamic shell to calculate the diameter of a protein in solution. In addition to using DLS to verify the native globular structure of both huHC- and ΔDE at pH 7.2-7.5, DLS was also utilised to verify the pH disassembly point of ΔDE. Similar structures to ΔDE have been reported to disassemble at a higher pH than pH 2.0. [257, 107]. However, the pH disassembly points differed, with Chen *et al.,* 2016 reporting a pH disassembly point of pH 4.0 and Ahn *et al.,* 2018 reporting pH 6.0.[257, 107] Therefore, to confirm ΔDE's pH disassembly point, ΔDE was dialysed against a range of pHs for 24 hrs and analysed by DLS. Three data sets were acquired for DLS, volume, number and intensity, with each set transforming the data raw data differently. The intensity is the initial raw data acquired and distributes the particles according to scattering intensity, with larger particles dominating the profile. The raw intensity can then be converted to either volume, which weights the profile on a particle's spherical volume or number, which weights the profile on the relative ratio of each particle in the solution. Inspection of the intensity profiles shows that huHC-s' primary peak is 12 nm, consistent with apoFt's 12 nm cage, with a minor peak between 100 -1000 nm, as shown in Figure 4.9ai. However, the primary peak for ΔDE resolves between 100 -1000 nm as shown in Figure 4.9bi. One possible reason for ΔDE's left profile shift, could be attributed to the removal of the DE loop structure near the 4-fold interface, increasing ΔDE dimerisation. However, ΔDE's secondary peak is absent in both the volume and number profile, as shown in Figure 4.9bii and cii. This indicates that the high molecular weight particle seen for both apoFt variants on the intensity profile accounts for a very small percentage of the total particles present. As such it can be suggested that at pH > 7.0, both apoFt variants form stable 12 nm cages.

Following observations that both variants form 12 nm cages, the disassembly point of each variant was investigated. It should be noted that huHC- samples dialysed against buffers at pH 6.0 formed large aggregates. The extensive aggregation prevented sample recovery and therefore was omitted from DLS analysis. In comparison, ΔDE did not show aggregates at pH 6.0 but showed a ~12 nm and 100-1000 nm peak, similar to the intensity profile of huHC- at pH > 7.0, indicating ΔDE's stability at pH 6.0 (Figure 4.9ai and bi). However, as the pH was further reduced to pH 5.0, ΔDE formed larger aggregates, indicated by a single high MW peak at ~200 nm on the intensity profile (Figure 4.9ai). Further analysis of ΔDE at pH 5.0, using both

number and volume profiles, showed two MW particle peaks at ~30 nm and ~100 nm (Figure 4.9aii and aiii). Given that neither MW peak reflects apoFt's 12 nm cage, it can be implied that ΔDE is unstable at pH 5.0, leading to high MW aggregates. At pH 4.0, ΔDE's number and volume profiles shifts to ~6-8 nm, similar to the profile seen for huHC- at pH 2.0 (Figure 4.9 aii-biii). Considering that DLS estimates the globular nm size of a particle from its longest point and that an apoFt subunit is ~4-5 nm in length, it could be suggested that ΔDE had disassembled into monomeric or dimeric subunits at pH 4.0. A disassembly point of pH 4.0 was also verified by native PAGE, which showed smaller MW bands <pH 6.0, indicative of a disassembled ΔDE cage (Figure 4.10).

ai
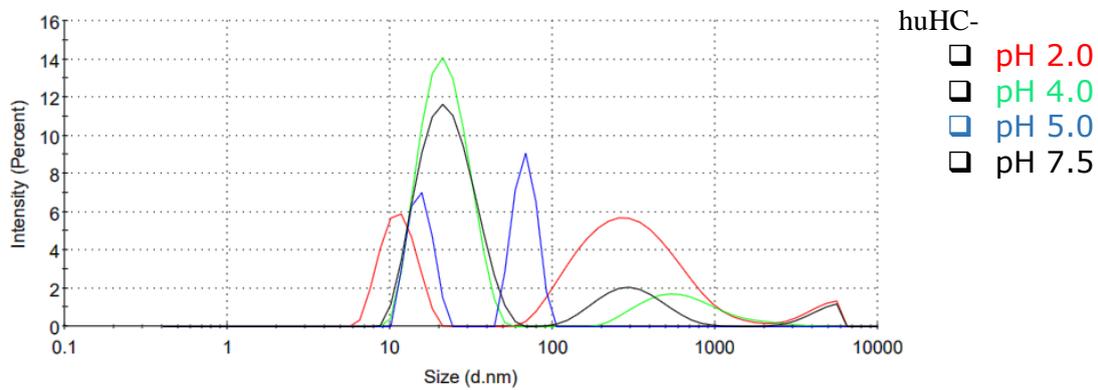
bi
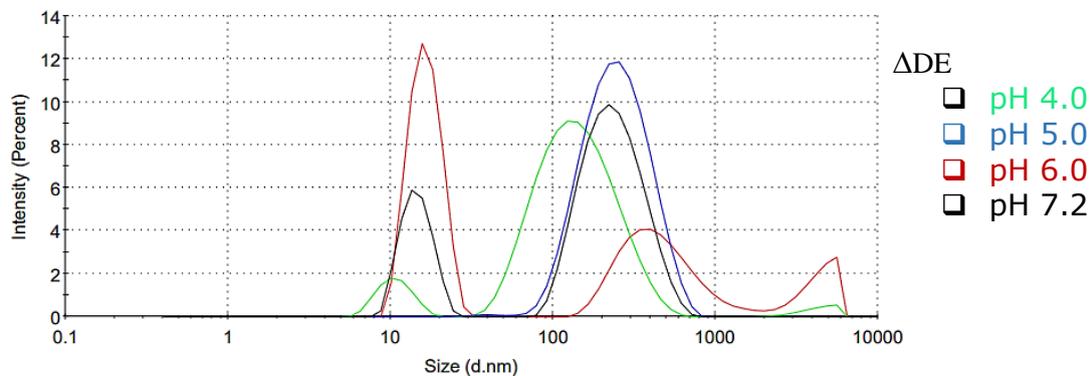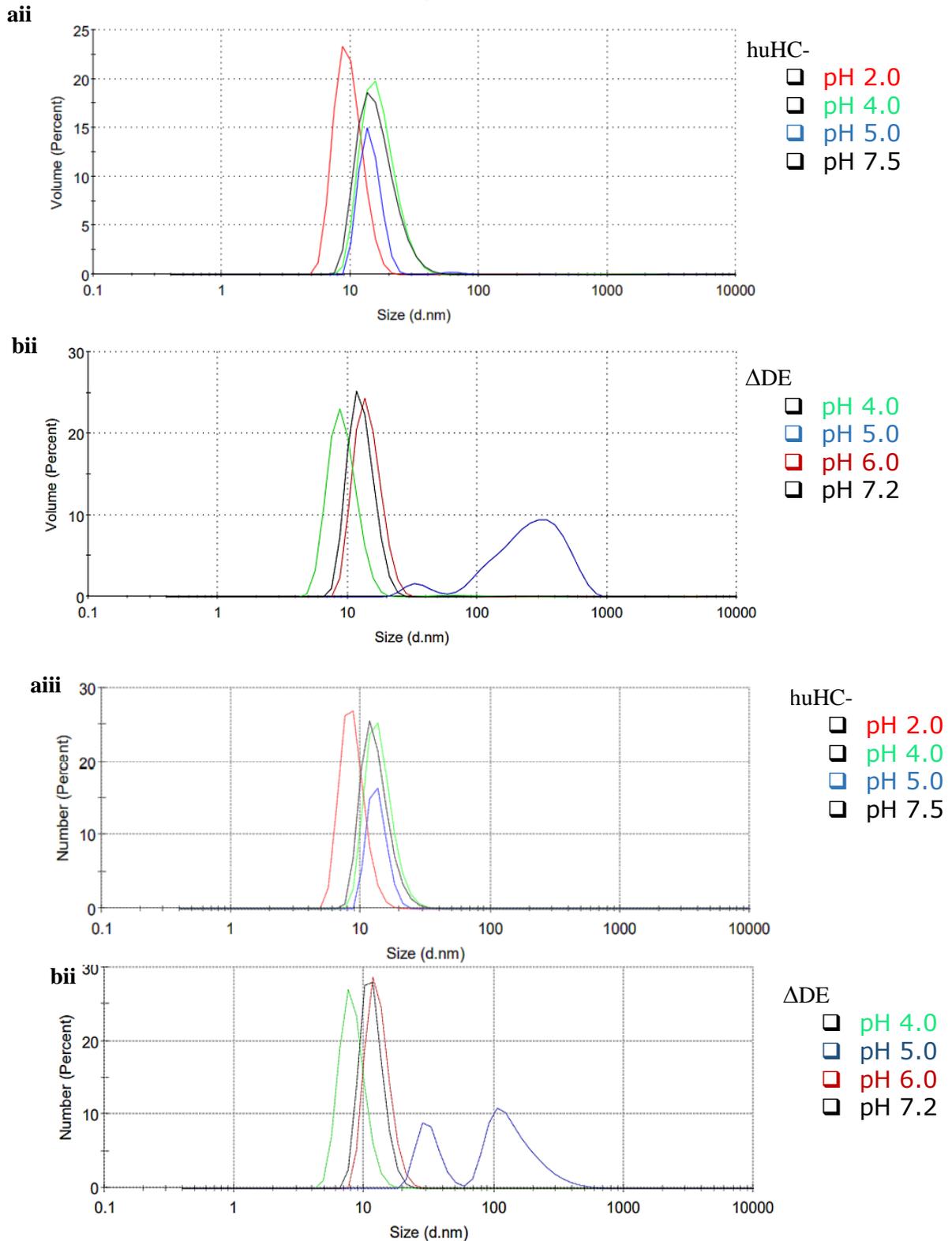
Figure 4. 9 Dynamic light scanning (DLS) spectra show the disassembly spectra of apoFt at pH 2.0 (50 mM potassium chloride), pH 4.0/5.0/6.0 (50 mM sodium acetate) and pH 7.2/7.5 (20 mM Tris). **ai and bi:** Intensity spectra of huHC- and ΔDE, **aii and bii:** volume spectra of huHC- and ΔDE and **aiii and biii:** number spectra of huHC- and ΔDE. ApoFt concentration: 0.4 mg/ mL
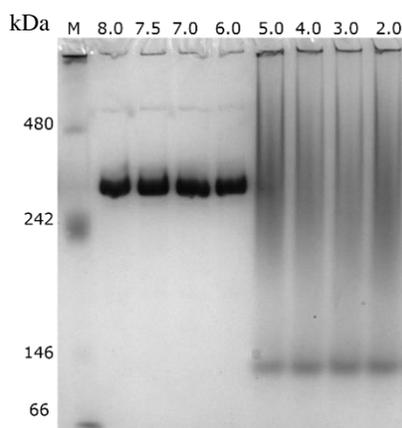
Figure 4.10 6% native PAGE of ΔDE disassembly at pH 2.0 (50 mM potassium chloride), pH 4.0/5.0/6.0 (50 mM sodium acetate) and pH 7.0/7.5/8.0 (20 mM Tris). M, Marker

Following DLS and native PAGE analysis of huHC-'s and ΔDE's disassembly points, a national COVID-19 lockdown was enforced, and no experimental work could be performed over six months. To overcome the lack of *in vitro* data, fundamental *in silico* work was performed. *In silico* experiments were self-directed under a short time. Therefore the data presented here should only be used as guidance to understand huHC-'s and ΔDE's four-fold channel characteristics as the variants undergo pH-dependent disassembly. An initial electrostatic potential analysis was performed using PyMol to elucidate the charge distribution of huHC-'s and ΔDE's four-fold channel at various pH. At pH 7.5, the surface charge surrounding ΔDE's four-fold channel is seen to be less negative than huHC-, as shown in Figure 4.11. This indicates that removing 23 AAs at the C-terminal, which is mainly composed of negatively charged residues, decreases the electrostatic potential surrounding the fourfold channel. Therefore, it can be hypothesised that as ΔDE undergoes pH-based disassembly, the four-fold channels are likely to reach a neutral charge at a higher pH, promoting a higher pH disassembly point than huHC-. As well as revealing the electrostatic charge distribution, PyMol analysis also confirmed the presence of a wide four-fold channel on ΔDE, as shown in Figure 4.11e. To further understand the structure of ΔDE's wide four-fold channel, a new analysis software was employed, given that PyMol's built-in tools were unable to map out channels.
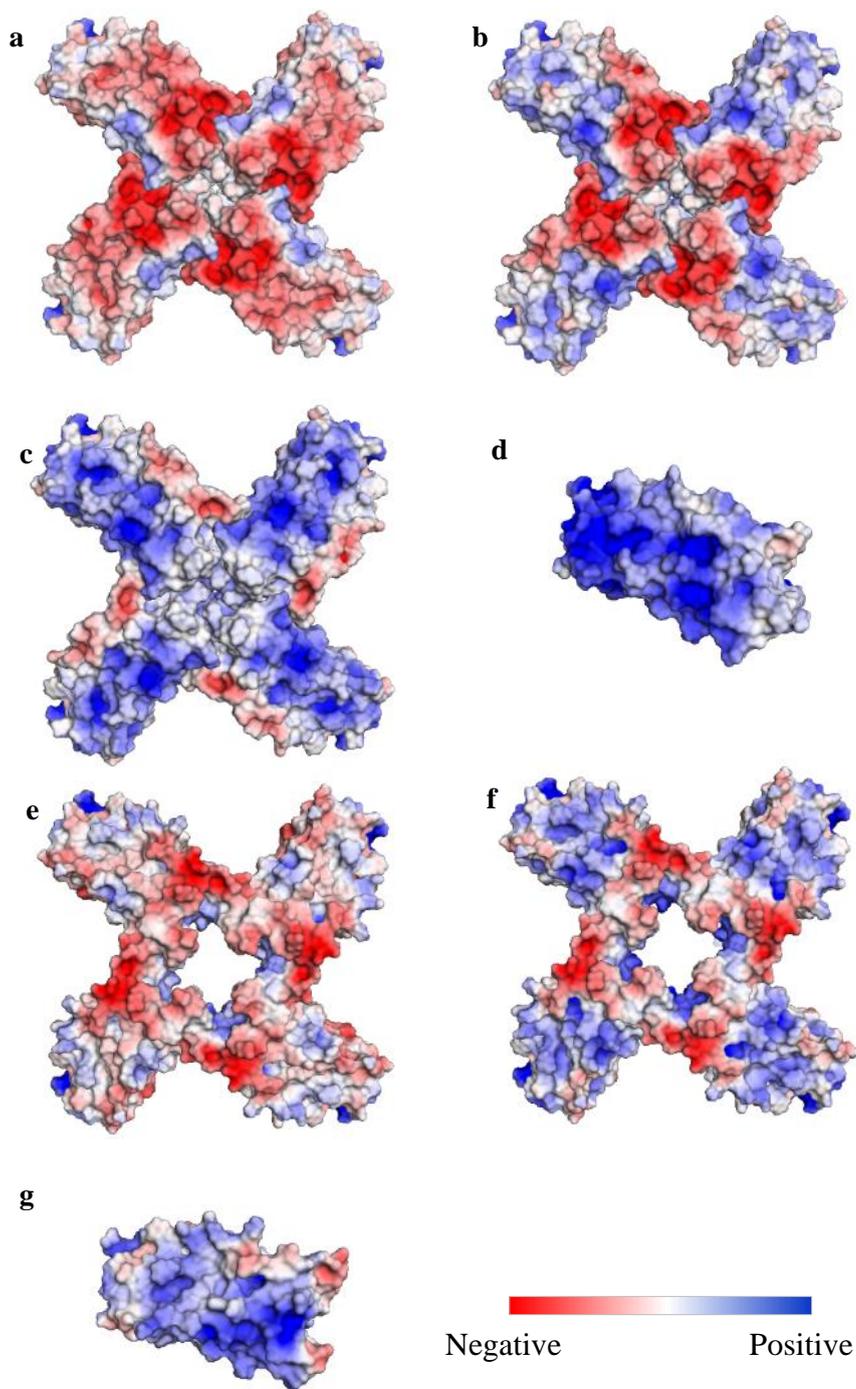
Figure 4.11 Electrostatic charge of each apoFt variant at different pHs. huHC- at **a**, pH 7.5, **b**, pH 5.0, **c**, pH 4.0 and **d**, pH 2.0. ΔDE at **e**, pH 7.5, **f**, pH 5.0 and **g**, pH 4.0.

The computational modelling software, Cavern, was employed to analyse both huHC-'s and ΔDE's four-fold channel. Cavern was selected over channel analysis software such as Molaxis for many reasons, including its compatibility with multiple operating systems, data visualisation using PyMol, and ease of data transfer into a user-friendly read Excel file.[260] All these factors were necessary given that during the COVID-19 lockdown, there was minimal support and

technology available for *in silico* analysis. Cavern was able to map out the architecture of huHC-'s four-fold channel, which showed a distinct funnel shape, lined with charged residues at either end of the channel and two hydrophobic leucine residues at its centre, as shown in Figure 4.12a. With regards to encapsulation via diffusion, the passage of ions/compounds through the four-fold channel would likely be hindered by the methyl groups on residues L165 and L169 and the imidazole groups on H173. The presence of H173 would also cause steric hindrance, given that the four-fold tunnel is narrowest at this point, ~3 Å wide, therefore preventing entry of large compounds.

In comparison to HuHC-, ΔDE's four-fold channel formed a cone-shaped architecture, the shape and the AAs lining the channel, as shown in Figure 4.12. Methionine, asparagine and lysine residues could also be visualised protruding from ΔDE's four-fold channel, as shown in Figure 4.12b. Similar to huHC-'s four-fold channel, there is a risk that the presence of the AAs that contain large side groups such as N51 would hinder ion/compound diffusion. However, given that the channel radius expands above ~10 Å, residues N51 and K50 are unlikely to hinder ion/compound entry, as shown in Figure 4.12c. Changes to the structure and the AAs lining ΔDE's four-fold channel are expected, given the absence of the DE loop, which usually lines huHC- four-fold channel. Further analysis of ΔDE's four-fold channel indicated four areas that 'bulge in' from the main channel, as shown by the arrows in Figure 4.13. These areas are lined with hydrophobic residues and positively charged residues such as K50, K54 and K158. Although ΔDE's four-fold channel has expanded considerably to avoid steric hindrance from the AAs lining the channel, the channel width is insufficient to permit the diffusion of Cyt C, which has a maximum diameter of ~39 Å and a minimum diameter of ~25 Å. Therefore the pH disassembly/assembly route was selected to encapsulate Cyt C.
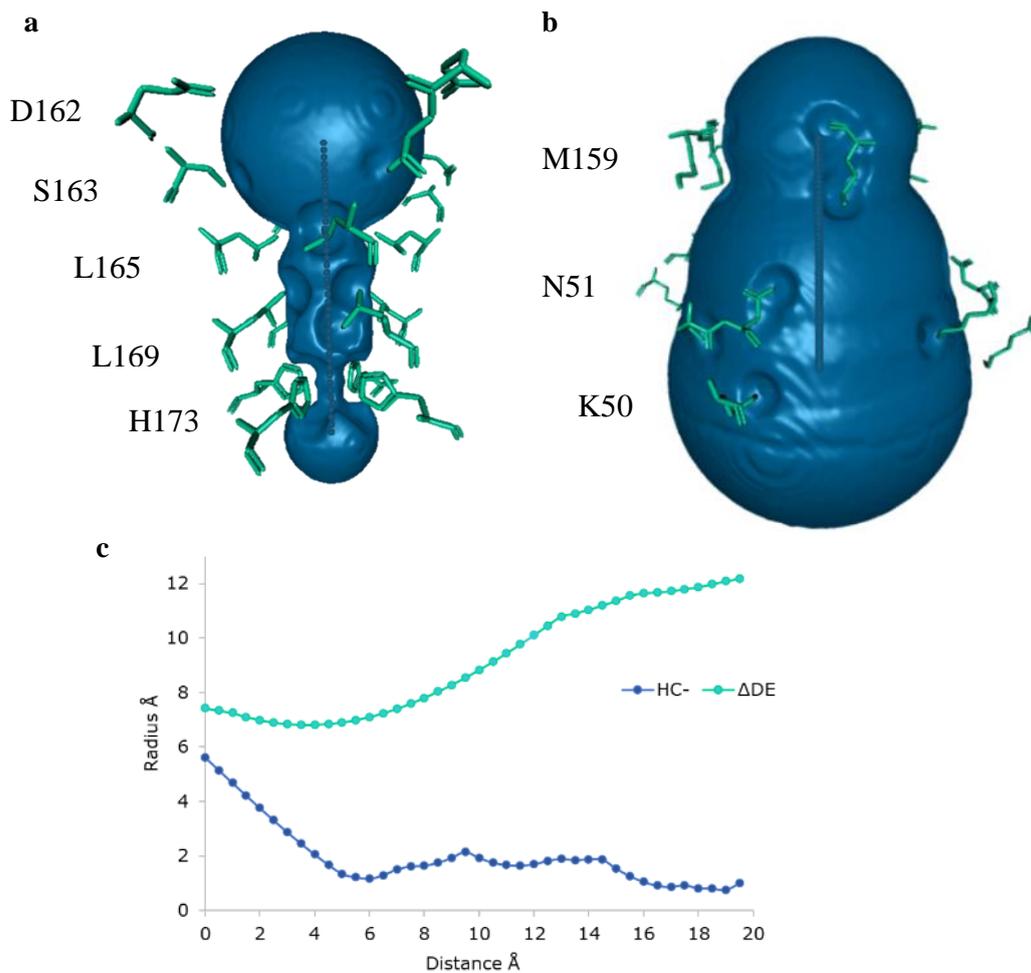
Figure 4.12 Four-fold channel structure of apoFt variants. AA lining of **a**, huHC- and **b**, ΔDE, **c**. channel diameter of huHC- and ΔDE.

Hydrophobic   Hydrophilic
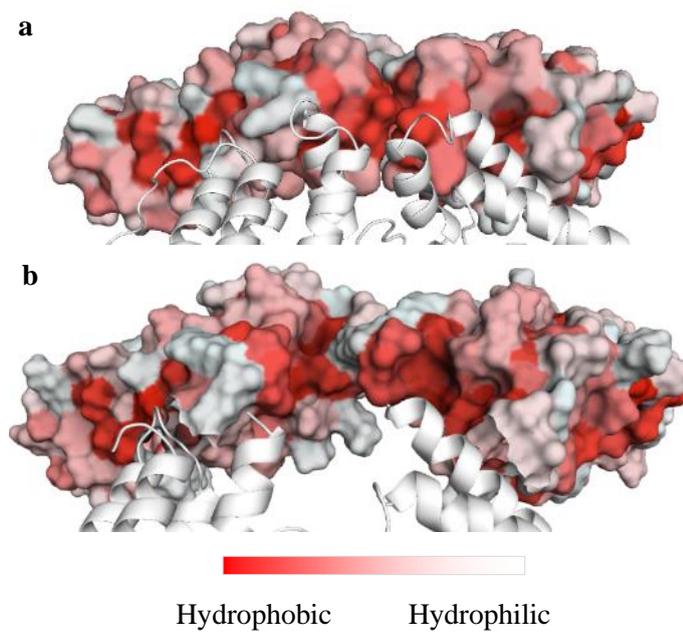
Figure 4. 13 Hydrophobicity profile of **a**, huHC- and **b**, ΔDE

4.1.5 *In vitro* effects of encapsulated cytochrome C

The pH-disassembly method was selected for Cyt C encapsulation, given that *in silico* results illustrated ΔDE's limited four-fold channel size for Cyt C diffusion. The pH-disassembly method was modified to account for ΔDE's new pH-disassembly (pH 4.0), which DLS verified in section 4.13. To analyse and compare ΔDE's tertiary structure, pre- and post- Cyt C encapsulation, native PAGE was utilised (Figure 4.14a). Cyt C encapsulated ΔDE, referred to as the 'encapsulate' from here on, produced a higher migration band than ΔDE, resolving to the same position as huHC-'s dimer form. Therefore, it can be presumed that the encapsulate had dimerised post-encapsulation. However, to test this hypothesis, DLS studies or transmission electron microscopy would need to be carried out to confirm the dimensions of the reassembled ΔDE.

Further analysis of the encapsulate also highlighted the existence of a small molecular species, indicating the presence of ΔDE monomers or an incomplete ΔDE cage (Figure 4.14a). This notion is supported by previous literature, which reports the adverse effects of engineering near the fold-fold channel, thereby hindering the reassembly process and promoting monomerisation.[261] However, as 300,000 Da spin concentrators were used post-encapsulation to separate formed caged from subunits, the small molecular species are unlikely to be monomers. Therefore, it can be presumed that the small molecular species were incomplete ΔDE cages, e.g., the absence of monomers. A similar phenomenon has been previously reported by Kim *et al.,* 2011, whereby pH-based disassembly/assembly resulted in a hole defect in the apoFt cage.[125] As the small molecular species is absent from the huHC- solution post encapsulation (Figure 4.4b), it can be suggested that the absence of the DE segment along the four-fold channel, in conjunction with the pH-based disassembly/assembly method, could have contributed to an incomplete ΔDE cage. However, the presence of ΔDE's possible hole defect would need further investigation, e.g transmission electron microscopy.

The encapsulation efficiency was then calculated using UV-vis, and an average of 2.5 Cyt C molecules were calculated to be encapsulated (Figure 4.14a). This is lower than the value acquired for huHC- (~3.5 molecules of Cyt C) and could be attributed to the presence of some incomplete ΔDE cages, which could not be removed using the 300,000 Da spin concentrators. The absence of a single monomer (~50 Å x 28 Å) from ΔDE's cage could enable Cyt C to escape and affect the ΔDE: Cyt C ratio, thereby underestimating the calculated value of encapsulated Cyt C. Nevertheless, the presence of ~2.5 Cyt C molecules is promising as it confirms that ΔDE has the capacity to encapsulate proteins. Following successful

encapsulation, Cyt C's cancer-specific cytotoxicity was evaluated using breast cancer cell lines, SKBR3 and MDA-MB-231and normal lung fibroblast cell line, MRC-5.
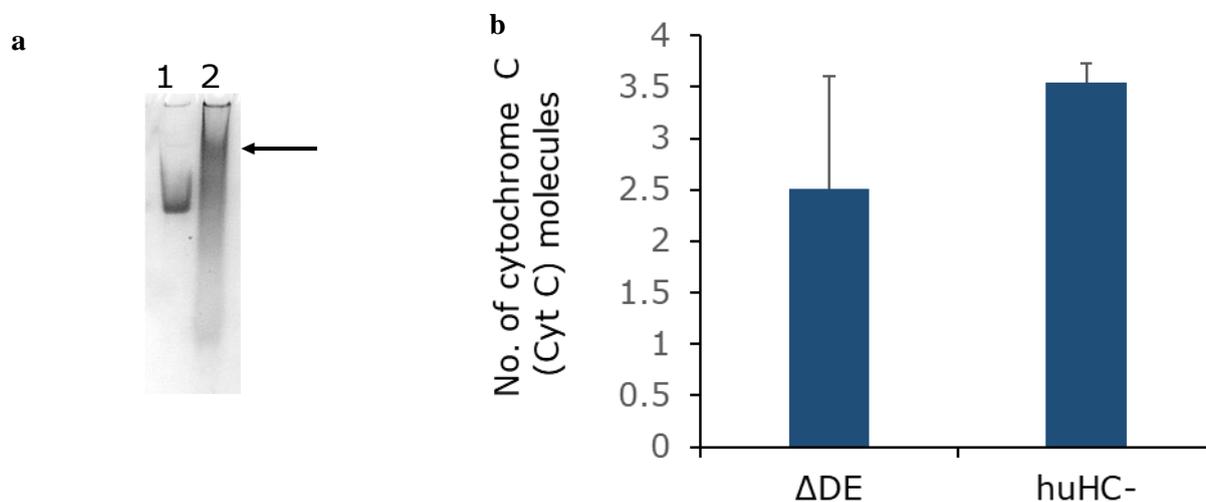
Figure 4.14 Encapsulation of Cyt C molecules in ΔDE and huHC-. **a**, Number of encapsulated Cyt C molecules. **b**, Native PAGE analysis of ΔDE. 1: ΔDE and encapsulated Cyt C. N=3 independent trials were performed. Mean ±SD

After UV-vis confirmed successful encapsulation of Cyt C, the effects of ΔDE, Cyt C and the 'encapsulate' were assessed using two breast cancer cell lines (SKBR3 and MDA-MB-231) and a normal lung fibroblast line (MRC-5). An MTT assay was selected to assess the activity of Cyt C against cells. The MTT assay is a measure of metabolic activity of the cells and, as such, can reflect proliferative inhibition upon treatment with a given test agent such as Cyt C. Following 72 hrs treatment, SKBR3 cells were treated with the encapsulate inhibiting growth with a GI_{50} of 406 ± 42 nM (Figure 4.15a). However, this effect was not visible in MDA-MB-231 cells, which retained a high number following treatment with the encapsulate, similar to the control treatments, ΔDE and Cyt C (Figure 4.15b). The lack of response from MDA-MB-231 cells is unexpected, considering that SKBR3 and MDA-MB-231 cell lines have similar protein expression levels of caspase 9, caspase 3, caspase 7 and XIAP, all of which are involved in Cyt C-induced apoptosis.[262] However, one possible reason for the lack of effect of Cyt C on MDA-MB-231 could be owed to the difference in the expression levels of PHAPI, a tumour repressor protein, in SKBR3 and MDA-MB-231 cells. Previous reports support this hypothesis,

indicating that overexpression of PHAPI in HER2+ cells (MDA-MB-361 and BT474) cells promotes higher sensitivity to Cyt C treatment than HER2- cell lines (T47D and MCF7), which showed minimal PHAPI expression. [171] As MDA-MB-231 is a HER2- cell line, western blot studies would need to be conducted to confirm PHAPI expression levels.

Compared to SKBR3 cells, the encapsulate had very little/no effect on MRC-5 cells (Figure 4.15c). The lack of response could be a consequence of two factors, reduced sensitivity to Cyt C and low expression levels of TfR1 on MRC-5 cells. Schafer *et al*., 2006 showed that 'normal' cell lines such as primary human mammary epithelial cells have a reduced sensitivity to Cyt C treatment due to decreased activation of caspase 3/9.[171] Therefore, it can be suggested that the core apoptosome components in MRC-5 cells are less sensitive to activation by Cyt C. In addition to the change in caspase activity, 'normal' cell lines such as MRC-5, express low levels of TfR1 when compared to SKBR3 cells (Figure 3.12). The low levels of TfR1 would mean decreased uptake of ΔDE and the encapsulate by MRC-5 cells, thereby reducing the effects of Cyt C. Overall, the results from the MTT assays confirmed ΔDE's ability to deliver encapsulated Cyt C and breast cancer selectively of Cyt C.
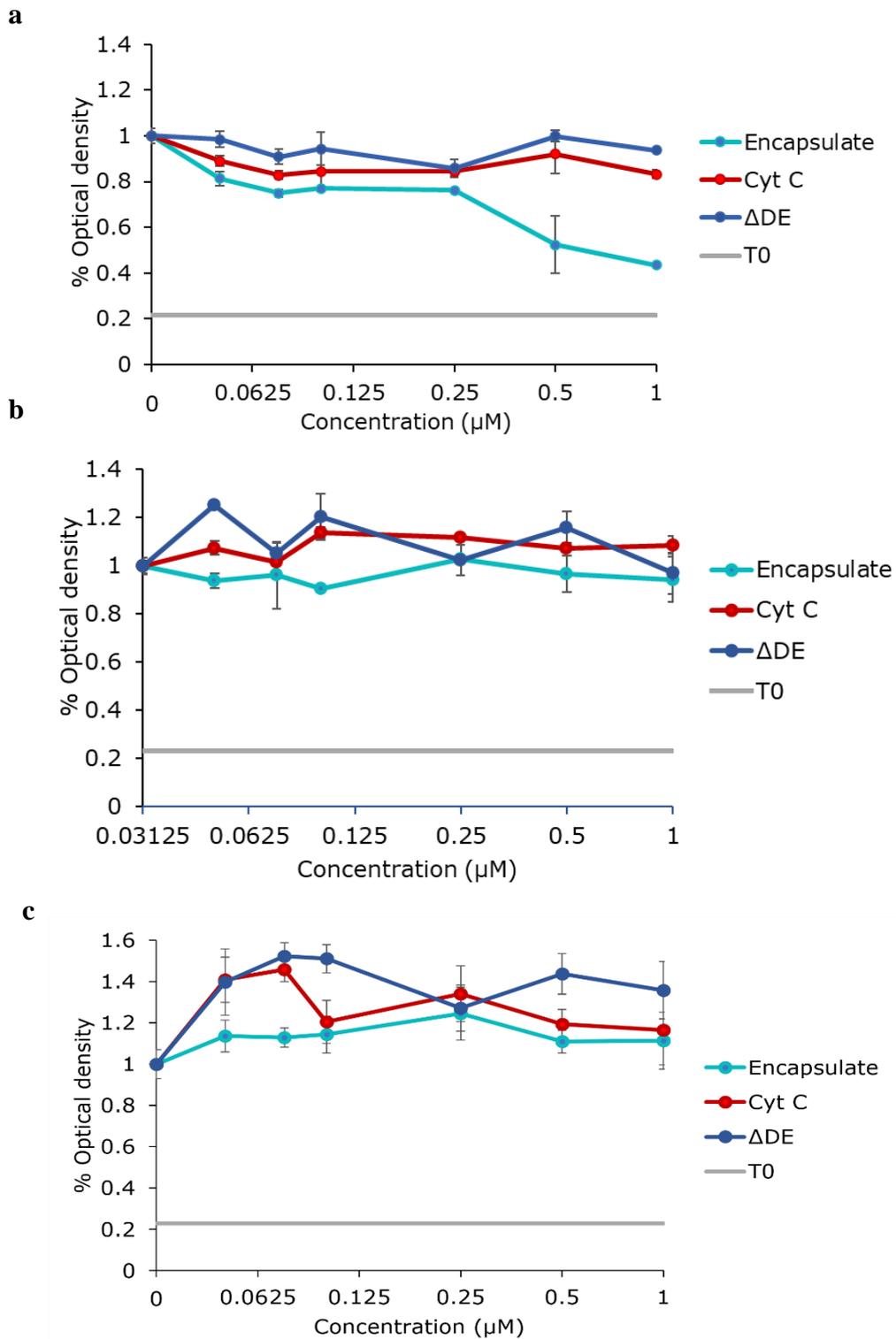
Figure 4.15 The effect of proliferation on ΔDE, Cyt C and ΔDE encapsulated Cyt C (encapsulate) on. **a**, SKBR3 **b**, MDA-MB-231 and **c**, MRC-5 cells. Cells were treated with 50-1000 nM of ΔDE, Cyt C and ΔDE encapsulated Cyt C (encapsulate) for 72 hrs. The results are expressed as the mean ± SD of 2 independent repeats (N=2) and 3 internal repeats (n=3) for each construct for each cell line.

4.2 Conclusions and future work

The following represents a summary of the data presented in chapter 4. Future research is also suggested, to gain a better understanding of the current data but also to build upon the current work.

1. As this work was being conducted, no therapeutic proteins had been encapsulated within huHC. Cyt C was selected due to its physical and pro-apoptotic characteristics. Cyt C was encapsulated within huHC- using either the pH-disassembly or urea-based methods. Using UV-vis, it was calculated that both methods yielded ~3-4 Cyt C molecules per huHC- cage. Future cleavage experiments using protease could also be performed to verify further Cyt encapsulation and hence protection from proteases. Given that all the protease cleavage sites in Cyt C are also present in huHC-, a recombinant Cyt C, containing a unique protease sequence such as the tobacco etch virus cleavage site, would need to be expressed. Following encapsulation of the recombinant Cyt C and protease experiments, SDS-page gel analysis or mass spectrometry could be conducted to determine the presence or absence of any cleaved Cyt C fragments.

2. The last 23 AAs were removed to produce a huHC- variant that could disassemble at a higher pH. However, the new variant, Δ160, was urea-resistant and insoluble. pH studies and native PAGE analysis indicated that the presence of the HIS-tag led to Δ160 instability between pH 4-7. To account for this, the HIS-tag and the last 23 AAs were removed to produce ΔDE. Due to the absence of the histidine purification-tag, multiple purification methods were tested, including heat-inactivation, gel chromatography and ion-exchange chromatography.

3. Following purification, the disassembly points of huHC- and ΔDE were analysed using DLS. At pH 5.0, ΔDE showed signs of aggregation, indicated by the presence of high MW particles in both the volume and number spectra. However, once equilibrated to pH 4.0, ΔDE's volume and number profile shifted below 10 nm. This shift is similar to huHC-'s profile at pH 2.0, indicating that ΔDE had disassembled into monomers. ΔDE's higher pH disassembly point was also confirmed with native PAGE analysis.

4. *In silico* analysis of ΔDE's indicated that removal of the last 23 AAs, reduced the positive charge distribution around the four-fold channel compared to huHC-. In addition, Cavern, an *in-silico* tool for analysis channels, confirmed that the absence of the DE turn and E-helix leads to a broader four-fold channel, reaching as wide as 20 Å. Changes in the four-fold structure also resulted in minor grooves appearing along the channel. The grooves

could further increase the capacity of ΔDE to encapsulate larger compounds than possible with huHC-.

5. pH-based encapsulation of Cyt C in ΔDE was confirmed using UV-vis, with an average of 2-3 Cyt C encapsulated molecules per cage. The structural integrity of ΔDE post-encapsulation was also confirmed using native PAGE. However, the newly re-assembled protein resolved closer to its parent, huHC-, dimer state. Therefore, it can be suggested that ΔDE had either dimerised or had formed a new cage structure. To confirm this hypothesis, future work could involve quantifying the number and size of the particulates present post encapsulation using DLS or mass spectrometry. Other possible work could include transmission electron microscopy, as this would visually confirm the integrity of the ΔDE's cage and the type of particulates present after encapsulation.

6. Once encapsulated, the selectivity of Cyt C to inhibit cell growth in breast cancer cells over normal cells was evaluated using an MTT assay. Encapsulated Cyt C was shown to selectively reduce cell number in SKBR3 cells (breast cancer cell line) and not in MRC-5 cells (normal foetal lung fibroblast cell line). To corroborate the MTT assay, future work could include dual annexin V/PI flow cytometry apoptosis assays and a caspase cleavage/activity assay. In addition, it would be interesting to identify the GI_{50} of the ΔDE encapsulated Cyt C for other breast cancer cells, including MDA-MB-231, and confirm the activation of pro-apoptotic proteins such as caspase 9, 3 and 7.

7. Successful delivery of Cyt C confirmed ΔDE's capabilities to deliver proteins. However, additional work is required to demonstrate ΔDE's ability to successfully encapsulate smaller therapeutic peptides, given ΔDE's large channels. Experiments would involve testing ΔDE's potential to encapsulate peptides with a broad range of characteristics (e.g., hydrophobicity, MW and secondary structure). A recent paper showed that a truncated version of Cyt C, Cyt C 77–101, induced apoptosis in a similar manner to Cyt C.[263] Future work involving the encapsulation of Cyt C 77–101 within ΔDE, would not only elucidate ΔDE's ability to retain smaller MW peptides but also its capability of delivering small therapeutic peptides to cancer cells.

5.0 Concluding remarks

huH being a protein-based nanocarrier, can be easily engineered to exhibit therapeutic characteristics. Combining both the traits of the TfR- Afb variant and ΔDE encapsulated Cyt C construct will increase the tumour specific targeting capabilities of huH. As discussed in Chapter 1, many protein-based carriers exhibit beneficial traits. In the future, it will be important to incorporate these traits within huH and explore more cancer-specific protein/proteins or compounds for encapsulation, as it will increase huH's characteristic profile and make it a more promising nanocarrier for cancer therapy.

6.0 References

1. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2016. *CA. Cancer J. Clin.* **66**, 7–30 (2016).

2. Colditz, G. A. & Bohlke, K. Priorities for the primary prevention of breast cancer. *CA. Cancer J. Clin.* **64**, 186–194 (2014).

3. Collignon, J., Lousberg, L., Schroeder, H. & Jerusalem, G. Triple-negative breast cancer: treatment challenges and solutions. *Breast cancer.* **8**, 93–107 (2016).

4. Yersal, O. & Barutca, S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World J. Clin. Oncol.* **5**, 412–24 (2014).

5. Schnitt, S. J. Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy. *Mod. Pathol.* **23**, S60–S64 (2010).

6. Perez, E. A. *et al.* Trastuzumab plus adjuvant chemotherapy for human epidermal growth factor receptor 2 - Positive breast cancer: Planned joint analysis of overall survival from NSABP B-31 and NCCTG N9831. *J. Clin. Oncol.* **32**, 3744–3752 (2014).

7. Perez, E. A. *et al. HER2* and Chromosome 17 Effect on Patient Outcome in the N9831 Adjuvant Trastuzumab Trial. *J. Clin. Oncol.* **28**, 4307–4315 (2010).

8. Paik, S. *et al.* Benefit from adjuvant trastuzumab may not be confined to patients with IHC 3+ and/or FISH-positive tumors: Central testing results from NSABP B-31. *J. Clin. Oncol.* **25**, 511–511 (2007).

9. Wang, X. *et al.* The prognostic value of androgen receptor (AR) in HER2-enriched metastatic breast cancer. *Endocr. Relat. Cancer* **27**, 199–208 (2020).

10. He, L. *et al.* Targeting Androgen Receptor in Treating HER2 Positive Breast Cancer. *Sci. Rep.* **7**, 14584 (2017).

11. A Study of a New Combination and Schedule of Chemotherapy Drugs for the Treatment of Head and Neck Cancer - Full Text View - ClinicalTrials.gov. Available at: https://clinicaltrials.gov/ct2/show/NCT00148122. (Accessed: 17th June 2020)

12. Palbociclib in Molecularly Characterized ER-positive/HER2-negative Metastatic Breast Cancer - Full Text View - ClinicalTrials.gov. Available at: https://clinicaltrials.gov/ct2/show/NCT02536742. (Accessed: 17th June 2020)

13. TIGER - The Institute of Cancer Research, London. Available at: https://www.icr.ac.uk/our-research/centres-and-collaborations/centres-at-the-icr/clinical-trials-and-statistics-unit/clinical-trials/tiger. (Accessed: 17th June 2020)

14. Auffan, M. *et al.* Towards a definition of inorganic nanoparticles from an environmental, health and safety perspective. *Nature Nanotechnology* **4**, 634–641 (2009).

15. Maeda, H. The link between infection and cancer: Tumor vasculature, free radicals, and drug delivery to tumors via the EPR effect. *Cancer Science* **104**, 779–789 (2013).

16. Landi-Librandil, A. P. *et al.* Study of quercetin-loaded liposomes as potential drug carriers: In vitro evaluation of human complement activation. *J. Liposome Res.* **22**, 89–99 (2012).

17. Khoshnejad, M. *et al.* Ferritin Nanocages with Biologically Orthogonal Conjugation for Vascular Targeting and Imaging. *Bioconjug. Chem.* **29**, 1209–1218 (2018).

18. Jain, R. K. & Stylianopoulos, T. Delivering nanomedicine to solid tumors. *Nature Reviews Clinical Oncology* **7**, 653–664 (2010).

19. Li, Y., Ayala-Orozco, C., Rauta, P. R. & Krishnan, S. The application of nanotechnology in enhancing immunotherapy for cancer treatment: Current effects and perspective. *Nanoscale* **11**, 17157–17178 (2019).

20. Lu, C. *et al.* Phase I clinical trial of systemically administered TUSC2(FUS1)-nanoparticles mediating functional gene transfer in humans. *PLoS One* **7**, (2012).

21. Vergote, I., Brize, A., Lisyanskaya, A. S. & Lichinitser, M. Randomized phase III study comparing paclical-carboplatin with paclitaxel-carboplatin in patients with recurrent platinum-sensitive epithelial ovarian cancer. *J. Clin. Oncol.* **33**, 5517–5517 (2015).

22. Joshi, M. D., Patravale, V. & Prabhu, R. Polymeric nanoparticles for targeted treatment in oncology: current insights. *Int. J. Nanomedicine* **10**, 1001 (2015).

23. Zielinska, A. *et al.* Polymeric Nanoparticles: Production, Characterization, Toxicology and Ecotoxicology. *Molecules* **25**, 3731 (2020).

24. Saeed, R. M., Dmour, I. & Taha, M. O. Stable Chitosan-Based Nanoparticles Using Polyphosphoric Acid or Hexametaphosphate for Tandem Ionotropic/Covalent Crosslinking and Subsequent Investigation as Novel Vehicles for Drug Delivery. *Front. Bioeng. Biotechnol.* **8**, 4 (2020).

25. Hines, D. J. & Kaplan, D. L. Poly(lactic-co-glycolic) acid-controlled-release systems: Experimental and modeling insights. *Crit. Rev. Ther. Drug Carrier Syst.* **30**, 257–276 (2013).

26. Bakó, J. *et al.* Poly- γ -Glutamic Acid Nanoparticles Based Visible Light-Curable Hydrogel for Biomedical Application. *J. Nanomater.* **2016**, 1–10 (2016).

27. Zhang, W., Mehta, A., Tong, Z., Esser, L. & Voelcker, N. H. Development of Polymeric Nanoparticles for Blood–Brain Barrier Transfer—Strategies and Challenges. *Adv. Sci.* **8**, 2003937 (2021).

28. Kim, I. *et al.* Doxorubicin-loaded highly porous large PLGA microparticles as a

sustained- release inhalation system for the treatment of metastatic lung cancer. *Biomaterials* **33**, 5574–5583 (2012).

29. Zhang, T. *et al.* Enhanced peroxidase activity and tumour tissue visualization by cobalt-doped magnetoferritin nanoparticles. *Nanotechnology* **28**, 045704 (2017).

30. Hong, S. *et al.* Protein-based nanoparticles as drug delivery systems. *Pharmaceutics* **12**, 1–28 (2020).

31. Charlton, J. R. *et al.* Biocompatibility of ferritin-based nanoparticles as targeted MRI contrast agents. *Nanomedicine Nanotechnology, Biol. Med.* **12**, 1735–1745 (2016).

32. Kim, R., Kim, K. K., Yokota, H. & Kim, S. H. Small heat shock protein of Methanococcus jannaschii, a hyperthermophile. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9129–33 (1998).

33. Virudachalam, R., Harrington, M. & Markley, J. L. Thermal stability of cowpea mosaic virus components: Differential scanning calorimetry studies. *Virology* **146**, 138–140 (1985).

34. Yang, C., Tian, R., Liu, T. & Liu, G. MRI reporter genes for noninvasive molecular imaging. *Molecules* **21**, 580 (2016).

35. King, N. P. *et al.* Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103–108 (2014).

36. Thelander, L., Gräslund, A. & Thelander, M. Continual presence of oxygen and iron required for mammalian ribonucleotide reduction: possible regulation mechanism. *Biochem. Biophys. Res. Commun.* **110**, 859–65 (1983).

37. Marden, M. C., Griffon, N. & Poyart, C. Oxygen delivery and autoxidation of hemoglobin. *Transfus. Clin. Biol.* **2**, 473–480 (1995).

38. Muchowska, K. B., Varma, S. J. & Moran, J. Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature* **569**, 104–107 (2019).

39. Ding, W.-X., Ni, H.-M., DiFrancesca, D., Stolz, D. B. & Yin, X.-M. Bid-dependent generation of oxygen radicals promotes death receptor activation-induced apoptosis in murine hepatocytes. *Hepatology* **40**, 403–413 (2004).

40. Phiri, K. S. *et al.* New cut-off values for ferritin and soluble transferrin receptor for the assessment of iron deficiency in children in a high infection pressure area. *J. Clin. Pathol.* **62**, 1103–1106 (2009).

41. Grossmann, J. G. *et al.* Metal-induced conformational changes in transferrins. *J. Mol. Biol.* **229**, 585–590 (1993).

42. Mack, U., Cooksley, W. G. E., Ferris, R. A., Powell, L. W. & Halliday, J. W. Regulation

of Plasma Ferritin by the Isolated Perfused Rat Liver. *Br. J. Haematol.* **47**, 403–412 (1981).

43. Dorner, M. *et al.* Ferritin synthesis by human T lymphocytes. *Science.* **209**, 1019–1021 (1980).

44. Silvestre, O. M. *et al.* Ferritin levels and risk of heart failure—the Atherosclerosis Risk in Communities Study. *Eur. J. Heart Fail.* **19**, 340–347 (2017).

45. Worwood, M., Dawkins, S., Wagstaff, M. & Jacobs, A. The purification and properties of ferritin from human serum. *Biochem. J.* **157**, 97–103 (1976).

46. Kell, D. B. & Pretorius, E. Serum ferritin is an important inflammatory disease marker, as it is mainly a leakage product from damaged cells. *Metallomics* **6**, 748–773 (2014).

47. Olthof, A. W. *et al.* Correlation between serum ferritin levels and liver iron concentration determined by MR imaging: impact of hematologic disease and inflammation. *Magn. Reson. Imaging* **25**, 228–231 (2007).

48. Li, J., Wang, R., Luo, D., Li, S. & Xiao, C. Association between Serum Ferritin Levels and Risk of the Metabolic Syndrome in Chinese Adults: A Population Study. *PLoS One* **8**, e74168 (2013).

49. Kell, D. B. & Pretorius, E. Serum ferritin is an important inflammatory disease marker, as it is mainly a leakage product from damaged cells. *Metallomics* **6**, 748–773 (2014).

50. De Martino, M., Ershov, D., van den Berg, P. J., Tans, S. J. & Meyer, A. S. Single-cell analysis of the Dps response to oxidative stress. *J. Bacteriol.* **198**, 1662–1674 (2016).

51. Sunanda Margrett Williams, A. *et al.* A Mutation Directs the Structural Switch of DNA Binding Proteins under Starvation to a Ferritin-like Protein Cage The findings of Williams et al. show how diversification of proteins can occur with minor changes in the primary sequence. A Mutation Directs the Structural Switch of DNA Binding Proteins under Starvation to a Ferritin-like Protein Cage. *Struct. Des.* **25**, 1449-1454.e3 (2017).

52. Kasyutich, O. *et al.* Silver Ion Incorporation and Nanoparticle Formation inside the Cavity of Pyrococcus furiosus Ferritin: Structural and Size-Distribution Analyses. *J. Am. Chem. Soc.* **132**, 3621–3627 (2010).

53. Jacobs, J. F., Nahid Hasan, M., Paik, K. H., Hagen, W. R. & Van Loosdrecht, M. C. M. Development of a bionanotechnological phosphate removal system with thermostable ferritin. *Biotechnol. Bioeng.* **105**, 918–923 (2010).

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

55. Sana, B., Johnson, E. & Lim, S. The unique self-assembly/disassembly property of Archaeoglobus fulgidus ferritin and its implications on molecular release from the protein cage. *Biochim. Biophys. Acta - Gen. Subj.* **1850**, 2544–2551 (2015).

56. Tosi, G., Belletti, D., Pederzoli, F. & Ruozi, B. Apoferritin nanocage as drug reservoir: is it a reliable drug delivery system? *Expert Opin. Drug Deliv.* **13**, 1341–1343 (2016).

57. Uchida, M. *et al.* A human ferritin iron oxide nano-composite magnetic resonance contrast agent. *Magn. Reson. Med.* **60**, 1073–1081 (2008).

58. Zhao, Y. *et al.* Bioengineered Magnetoferritin Nanoprobes for Single-Dose Nuclear-Magnetic Resonance Tumor Imaging. *ACS Nano* **10**, 4184–4191 (2016).

59. Matsumura, S., Aoki, I., Saga, T. & Shiba, K. A Tumor-Environment-Responsive Nanocarrier That Evolves Its Surface Properties upon Sensing Matrix Metalloproteinase-2 and Initiates Agglomeration to Enhance T_2 Relaxivity for Magnetic Resonance Imaging. *Mol. Pharm.* **8**, 1970–1974 (2011).

60. Masaki Uchida. *et al.* Targeting of Cancer Cells with Ferrimagnetic Ferritin Cage Nanoparticles. *J. Am. Chem. Soc.* **128**, 16626-16633 (2006).

61. Lin, X. *et al.* Chimeric Ferritin Nanocages for Multiple Function Loading and Multimodal Imaging. *Nano Lett.* **11**, 814–819 (2011).

62. Sánchez, P. *et al.* MRI relaxation properties of water-soluble apoferritin-encapsulated gadolinium oxide-hydroxide nanoparticles. *Dalton. Trans.* **5**, 800–804 (2009).

63. Winter, P. M. *et al.* Molecular imaging of angiogenesis in nascent Vx-2 rabbit tumors using a novel alpha(nu)beta3-targeted nanoparticle and 1.5 tesla magnetic resonance imaging. *Cancer Res.* **63**, 5838–43 (2003).

64. Li, L. *et al.* Ferritin-mediated siRNA delivery and gene silencing in human tumor and primary cells. *Biomaterials* **98**, 143–151 (2016).

65. Yang, Z. *et al.* Encapsulation of platinum anticancer drugs by apoferritin. *Chem. Commun.* **33**, 3453-3455 (2007).

66. Xing, R. *et al.* Characterization and cellular uptake of platinum anticancer drugs encapsulated in apoferritin. *J. Inorg. Biochem.* **103**, 1039–1044 (2009).

67. Falvo, E. *et al.* Antibody–drug conjugates: targeting melanoma with cisplatin encapsulated in protein-cage nanoparticles based on human ferritin. *Nanoscale* **5**, 12278 (2013).

68. Bellini, M. *et al.* Protein nanocages for self-triggered nuclear delivery of DNA-targeted chemotherapeutics in Cancer Cells. *J. Control. Release* **196**, 184–196 (2014).

69. Liang, M. *et al.* H-ferritin-nanocaged doxorubicin nanoparticles specifically target and

kill tumors with a single-dose injection. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14900–5 (2014).

70. Kilic, M. A., Ozlu, E. & Calis, S. A novel protein-based anticancer drug encapsulating nanosphere: apoferritin-doxorubicin complex. *J. Biomed. Nanotechnol.* **8**, 508–14 (2012).

71. Chen, L. *et al.* Encapsulation of curcumin in recombinant human H-chain ferritin increases its water-solubility and stability. *Food Res. Int.* **62**, 1147–1153 (2014).

72. Zhen, Z. *et al.* Ferritin nanocages to encapsulate and deliver photosensitizers for efficient photodynamic therapy against cancer. *ACS Nano* **7**, 6988–96 (2013).

73. Tetter, S. & Hilvert, D. Enzyme Encapsulation by a Ferritin Cage. *Angew. Chemie Int. Ed.* **56**, 14933–14936 (2017).

74. Carmona, U., Li, L., Zhang, L. & Knez, M. Ferritin light-chain subunits: Key elements for the electron transfer across the protein cage. *Chem. Commun.* **50**, 15358–15361 (2014).

75. Santambrogio, Paolo, *et al*. Evidence that a salt bridge in the light chain contributes to the physical stability difference between heavy and light human ferritins. *J. Biol. Chem.* **267**, 14077-14083 (1992)

76. Masuda, T., Goto, F., Yoshihara, T. & Mikami, B. Crystal structure of plant ferritin reveals a novel metal binding site that functions as a transit site for metal transfer in ferritin. *J. Biol. Chem.* **285**, 4049–59 (2010).

77. Honarmand Ebrahimi, K., Bill, E., Hagedoorn, P.-L. & Hagen, W. R. The catalytic center of ferritin regulates iron storage via Fe(II)-Fe(III) displacement. *Nat. Chem. Biol.* **8**, 941–948 (2012).

78. Levi, S. *et al.* Evidence that H-and L-chains have co-operative roles in the iron-uptake mechanism of human ferritin. *Biochem. J* **288**, 591-596 (1992).

79. Zhao, G., Arosio, P. & Chasteen, N. D. Iron(II) and Hydrogen Peroxide Detoxification by Human H-Chain Ferritin. An EPR Spin-Trapping Study [†]. *Biochemistry* **45**, 3429–3436 (2006).

80. Chen, T. T. *et al.* TIM-2 is expressed on B cells and in liver and kidney and is a receptor for H-ferritin endocytosis. *J. Exp. Med.* **202**, 955–65 (2005).

81. Lawrence, C. M. *et al.* Crystal structure of the ectodomain of human transferrin receptor. *Science.* **286**, 779–782 (1999).

82. Qian, Z. M., Xiao, D. S., Tang, P. L., Yao, F. Y. D. & Liao, Q. K. Increased expression of transferrin receptor on membrane of erythroblasts in strenuously exercised rats. *J.*

Appl. Physiol. **87**, 523–529 (1999).

83. Bierings, M. B., Baert, M. R. M., van Eijk, H. G. & van Dijk, J. P. Transferrin receptor expression and the regulation of placental iron uptake. *Mol. Cell. Biochem.* **100**, 31–38 (1991).

84. Habashy, H. O. *et al.* Transferrin receptor (CD71) is a marker of poor prognosis in breast cancer and can predict response to tamoxifen. *Breast Cancer Res. Treat.* **119**, 283–293 (2010).

85. Prutki, M. *et al.* Altered iron metabolism, transferrin receptor 1 and ferritin in patients with colon cancer. *Cancer Lett.* **238**, 188–196 (2006).

86. Kondo, K. *et al.* Transferrin receptor expression in adenocarcinoma of the lung as a histopathologic indicator of prognosis. *Chest* **97**, 1367–1371 (1990).

87. Magro, G. *et al.* Aberrant expression of TfR1/CD71 in thyroid carcinomas identifies a novel potential diagnostic marker and therapeutic target. *Thyroid* **21**, 267–277 (2011).

88. Ponka, P. & Lok, C. N. The transferrin receptor: role in health and disease. *Int. J. Biochem. Cell Biol.* **31**, 1111–1137 (1999).

89. Greene, C. J. *et al.* Transferrin receptor 1 upregulation in primary tumor and downregulation in benign kidney is associated with progression and mortality in renal cell carcinoma patients. *Oncotarget* **8**, 107052–107075 (2017).

90. Moura, I. C. *et al.* A neutralizing monoclonal antibody (mAb A24) directed against the transferrin receptor induces apoptosis of tumor T lymphocytes from ATL patients. *Retrovirology* **8**, A60 (2011).

91. Sakamoto, S. *et al.* H-Ferritin Is Preferentially Incorporated by Human Erythroid Cells through Transferrin Receptor 1 in a Threshold-Dependent Manner. *PLoS One* **10**, e0139915 (2015).

92. Wallace, D. F., McDonald, C. & Subramaniam, V. N. Transferrin Receptor 1: A Ferritin Receptor as Well? *Gastroenterology* **139**, 1052–1053 (2010).

93. de Turris, V. *et al.* Humanized archaeal ferritin as a tool for cell targeted delivery. *Nanoscale* **9**, 647–655 (2017).

94. Montemiglio, L. C. *et al.* Cryo-EM structure of the human ferritin–transferrin receptor 1 complex. *Nat. Commun.* **10**, 1–8 (2019).

95. Antileo, E. *et al.* Endocytic pathway of exogenous iron-loaded ferritin in intestinal epithelial (Caco-2) cells. *Am. J. Physiol. Liver Physiol.* **304**, G655–G661 (2013).

96. Mancias, J. D., Wang, X., Gygi, S. P., Harper, J. W. & Kimmelman, A. C. Quantitative proteomics identifies NCOA4 as the cargo receptor mediating ferritinophagy. *Nature*

508, 105–109 (2014).

97. Baraibar, M. A., Muhoberac, B. B., Garringer, H. J., Hurley, T. D. & Vidal, R. Unraveling of the E-helices and disruption of 4-fold pores are associated with iron mishandling in a mutant ferritin causing neurodegeneration. *J. Biol. Chem.* **285**, 1950–1956 (2010).

98. Andrews, S. C. *et al.* Structure, function, and evolution of ferritins. *J. Inorg. Biochem.* **47**, 161–174 (1992).

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

100. Treffry, A. *et al.* Defining the roles of the threefold channels in iron uptake, iron oxidation and iron-core formation in ferritin: A study aided by site-directed mutagenesis. *Biochem. J.* **296**, 721–728 (1993).

101. Yablonski, M. J. & Theil, E. C. A possible role for the conserved trimer interface of ferritin in iron incorporation. *Biochemistry* **31**, 9680–9684 (1992).

102. Takahashi, T. & Kuyucak, S. Functional Properties of Threefold and Fourfold Channels in Ferritin Deduced from Electrostatic Calculations. *Biophys. J.* **84**, 2256–2263 (2003).

103. Hille, B. Ionic channels in excitable membranes. Current problems and biophysical approaches. *Biophys. J.* **22**, 283–294 (1978).

104. Levi, S. *et al.* Mutational analysis of the channel and loop sequences of human ferritin H-chain. *Biochem. J.* **264**, 381–388 (1989).

105. Jappelli, R. & Cesareni, G. Loop mutations affect ferritin solubility causing non-native aggregation of subunits or precipitation of fully assembled polymers. *FEBS Lett.* **394**, 311–315 (1996).

106. Chen, H. *et al.* Engineering protein interfaces yields ferritin disassembly and reassembly under benign experimental conditions. *Chem. Commun.* **52**, 7402–7405 (2016).

107. Ahn, B. *et al.* Four-fold Channel-Nicked Human Ferritin Nanocages for Active Drug Loading and pH-Responsive Drug Release. *Angew. Chemie Int. Ed.* **57**, 2909–2913 (2018).

108. Zhang, S. *et al.* "Silent" Amino Acid Residues at Key Subunit Interfaces Regulate the Geometry of Protein Nanocages. *ACS Nano* **10**, 10382–10388 (2016).

109. Zhang, S. *et al.* Conversion of the Native 24-mer Ferritin Nanocage into Its Non-Native 16-mer Analogue by Insertion of Extra Amino Acid Residues. *Angew. Chemie Int. Ed.* **55**, 16064–16070 (2016).

110. Kim, K. K., Kim, R. & Kim, S. H. Crystal structure of a small heat-shock protein. *Nature*

394, 595–599 (1998).

111. Bova, M. P., Huang, Q., Ding, L. & Horwitz, J. Subunit exchange, conformational stability, and chaperone-like function of the small heat shock protein 16.5 from Methanococcus jannaschii. *J. Biol. Chem.* **277**, 38468–38475 (2002).

112. Riki, T. *et al.* Biological evaluation of protein nanocapsules containing doxorubicin. *International Journal of Nanomedicine* **8**, 1989–1999 (2013).

113. Flenniken, M. L. *et al.* Selective attachment and release of a chemotherapeutic agent from the interior of a protein cage architecture. *Chem. Commun.* **4**, 447–449 (2005).

114. Murata, M. *et al.* Design and function of engineered protein nanocages as a drug delivery system for targeting pancreatic cancer cells via neuropilin-1. *Mol. Pharm.* **12**, 1422–1430 (2015).

115. Berardi, A., Evans, D. J., Baldelli Bombelli, F. & Lomonossoff, G. P. Stability of plant virus-based nanocarriers in gastrointestinal fluids. *Nanoscale* **10**, 1667–1679 (2018).

116. Hu, L. *et al.* Association of Vimentin overexpression and hepatocellular carcinoma metastasis. *Oncogene* **23**, 298–302 (2004).

117. Richardson, A. M. *et al.* Vimentin is required for lung adenocarcinoma metastasis via heterotypic tumor cell–Cancer-associated fibroblast interactions during collective invasion. *Clin. Cancer Res.* **24**, 420–432 (2018).

118. Wang, C., Fiering, S. N. & Steinmetz, N. F. Cowpea Mosaic Virus Promotes Anti-Tumor Activity and Immune Memory in a Mouse Ovarian Tumor Model. *Adv. Ther.* **2**, 1900003 (2019).

119. Sutherland, A. R., Alam, M. K. & Geyer, C. R. Post-translational Assembly of Protein Parts into Complex Devices by Using SpyTag/SpyCatcher Protein Ligase. *ChemBioChem* **20**, 319–328 (2019).

120. Cristie-David, A. S. & Marsh, E. N. G. Metal-dependent assembly of a protein nano-cage. *Protein Sci.* **28**, 1620-1629 (2019).

121. Gumulec, J. *et al.* Modulation of Induced Cytotoxicity of Doxorubicin by Using Apoferritin and Liposomal Cages. *Int. J. Mol. Sci.* **15**, 22960–22977 (2014).

122. Pontillo, N., Pane, F., Messori, L., Amoresano, A. & Merlino, A. Cisplatin encapsulation within a ferritin nanocage: A high-resolution crystallographic study. *Chem. Commun.* **52**, 4136–4139 (2016).

123. Webb, B., Frame, J., Zhao, Z., Lee, M. L. & Watt, G. D. Molecular Entrapment of Small Molecules within the Interior of Horse Spleen Ferritin. *Arch. Biochem. Biophys.* **309**, 178–183 (1994).

124. Pead, S. *et al.* Metal ion binding to apo, holo, and reconstituted horse spleen ferritin. *J. Inorg. Biochem.* **59**, 15–27 (1995).

125. Kim, M. *et al.* pH-Dependent Structures of Ferritin and Apoferritin in Solution: Disassembly and Reassembly. *Biomacromolecules* **12**, 1629–1640 (2011).

126. Jutz, G. & Böker, A. Bionanoparticles as functional macromolecular building blocks - A new class of nanomaterials. *Polymer* **52**, 211–232 (2011).

127. Jutz, G., van Rijn, P., Santos Miranda, B. & Böker, A. Ferritin: A Versatile Building Block for Bionanotechnology. *Chem. Rev.* **115**, 1653–1701 (2015).

128. Jin, W., Takagi, H., Pancorbo, B. & Theil, E. C. 'Opening' the ferritin pore for iron release by mutation of conserved amino acids at interhelix and loop sites. *Biochemistry* **40**, 7525–7532 (2001).

129. Uchida, M., Kang, S., Reichhardt, C., Harlen, K. & Douglas, T. The ferritin superfamily: Supramolecular templates for materials synthesis. *Biochim. Biophys. Acta* **1800**, 834–45 (2010).

130. Breen, A. F., Wells, G., Turyanska, L. & Bradshaw, T. D. Development of novel apoferritin formulations for antitumour benzothiazoles. *Cancer Rep.* **2**, e1155 (2019).

131. Maity, B., Abe, S. & Ueno, T. Observation of gold sub-nanocluster nucleation within a crystalline protein cage. *Nat. Commun.* **8**, 1–9 (2017).

132. Zhen, Z. *et al.* Ferritin nanocages to encapsulate and deliver photosensitizers for efficient photodynamic therapy against cancer. *ACS Nano* **7**, 6988–6996 (2013).

133. Lei, Y. *et al.* Targeted tumor delivery and controlled release of neuronal drugs with ferritin nanoparticles to regulate pancreatic cancer progression. *J. Control. Release* **232**, 131–142 (2016).

134. Wang, Q. *et al.* High hydrostatic pressure encapsulation of doxorubicin in ferritin nanocages with enhanced efficiency. *J. Biotechnol.* **254**, 34–42 (2017).

135. Breen, A. F., Wells, G., Turyanska, L. & Bradshaw, T. D. Development of novel apoferritin formulations for antitumour benzothiazoles. *Cancer Rep.* **2**, e1155 (2019).

136. Yang, D. & Nagayama, K. Permeation of small molecules into the cavity of ferritin as revealed by proton nuclear magnetic resonance relaxation. *Biochem. J.* **307**, 253–256 (1995).

137. Wang, Q., Mercogliano, C. P. & Löwe, J. A ferritin-based label for cellular electron cryotomography. *Structure* **19**, 147–154 (2011).

138. Vande Velde, G. *et al.* Evaluation of the specificity and sensitivity of ferritin as an MRI reporter gene in the mouse brain using lentiviral and adeno-associated viral vectors. *Gene*

Ther. **18**, 594–605 (2011).

139. Jiang, Y. *et al.* Preparation of hypocrellin B nanocages in self-assembled apoferritin for enhanced intracellular uptake and photodynamic activity. *J. Mater. Chem. B* **5**, 1980–1987 (2017).

140. Luzzago, A. & Cesareni, G. Isolation of point mutations that affect the folding of the H-chain of human ferritin in E. coli. *EMBO J.* **8**, 569–576 (1989).

141. Kim, S. *et al.* Double-chambered ferritin platform: Dual-function payloads of cytotoxic peptides and fluorescent protein. *Biomacromolecules* **17**, 12–19 (2016).

142. Ravishankar, S. & Lim, S. Cyclodextrin conjugated ferritin nanocages reduce intracellular cholesterol level in foam cells. Nano Research, **12**, 2925-2932 (2019).

143. Králová, J. *et al.* Porphyrin - Cyclodextrin conjugates as a nanosystem for versatile drug delivery and multimodal cancer therapy. *J. Med. Chem.* **53**, 128–138 (2010).

144. Georgiev, I. S. *et al.* Two-Component Ferritin Nanoparticles for Multimerization of Diverse Trimeric Antigens. *ACS Infect. Dis.* **4**, 788–796 (2018).

145. Hamburger, A. E., West, A. P., Hamburger, Z. A., Hamburger, P. & Bjorkman, P. J. Crystal structure of a secreted insect ferritin reveals a symmetrical arrangement of heavy and light chains. *J. Mol. Biol.* **349**, 558–569 (2005).

146. Wang, Z. *et al.* Ferritin nanocage-based antigen delivery nanoplatforms: Epitope engineering for peptide vaccine design. *Biomater. Sci.* **7**, 1794–1800 (2019).

147. Chen, J. L.-Y. *et al.* Prominin-1-Specific Binding Peptide-Modified Apoferritin Nanoparticle Carrying Irinotecan as a Novel Radiosensitizer for Colorectal Cancer Stem-Like Cells. *Part. Part. Syst. Charact.* **34**, 1600424 (2017).

148. Ki, M. R., Nguyen, T. K. M., Jun, H. S. & Pack, S. P. Biosilica-enveloped ferritin cage for more efficient drug deliveries. *Process Biochem.* **68**, 182–189 (2018).

149. Lee, W. *et al.* A Double-Chambered Protein Nanocage Loaded with Thrombin Receptor Agonist Peptide (TRAP) and γ-Carboxyglutamic Acid of Protein C (PC-Gla) for Sepsis Treatment. *Adv. Mater.* **27**, 6637–6643 (2015).

150. Yoo, J. Do *et al.* Designed ferritin nanocages displaying trimeric TRAIL and tumor-targeting peptides confer superior anti-tumor efficacy. *Sci. Rep.* **10**, 1–17 (2020).

151. Hu, J. *et al.* Designed Antimicrobial and Antitumor Peptides with High Selectivity. *Biomacromolecules* **12**, 3839–3843 (2011).

152. Iegre, J. *et al.* Stapled peptides as a new technology to investigate protein–protein interactions in human platelets. *Chem. Sci.* **9**, 4638–4643 (2018).

153. Kumar, D. & Bhalla, T. C. Microbial proteases in peptide synthesis: approaches and

applications. *Appl. Microbiol. Biotechnol.* **68**, 726–736 (2005).

154. Adamina, M. *et al.* Encapsulation into sterically stabilised liposomes enhances the immunogenicity of melanoma-associated Melan-A/MART-1 epitopes. *Br. J. Cancer* **90**, 263–269 (2004).

155. Mohan, A., Rajendran, S. R. C. K., Thibodeau, J., Bazinet, L. & Udenigwe, C. C. Liposome encapsulation of anionic and cationic whey peptides: Influence of peptide net charge on properties of the nanovesicles. *LWT* **87**, 40–46 (2018).

156. Li, L., Zhang, L. & Knez, M. Comparison of two endogenous delivery agents in cancer therapy: Exosomes and ferritin. *Pharmacol. Res.* **110**, 1–9 (2016).

157. Tetter, S. & Hilvert, D. Enzyme Encapsulation by a Ferritin Cage. *Angew. Chemie Int. Ed.* **56**, 14933–14936 (2017).

158. Hestericová, M., Heinisch, T., Lenz, M. & Ward, T. R. Ferritin encapsulation of artificial metalloenzymes: Engineering a tertiary coordination sphere for an artificial transfer hydrogenase. *Dalton. Trans.* **47**, 10837–10841 (2018).

159. Llauró, A. *et al.* Cargo-shell and cargo-cargo couplings govern the mechanics of artificially loaded virus-derived cages. *Nanoscale* **8**, 9328–9336 (2016).

160. Lucon, J. *et al.* Use of the interior cavity of the P22 capsid for site-specific initiation of atom-transfer radical polymerization with high-density cargo loading. *Nat. Chem.* **4**, 781–788 (2012).

161. Patterson, D. P. *et al.* Virus-like particle nanoreactors: Programmed encapsulation of the thermostable CelB glycosidase inside the P22 capsid. *Soft Matter* **8**, 10158–10166 (2012).

162. Patterson, D. P., Prevelige, P. E. & Douglas, T. Nanoreactors by programmed enzyme encapsulation inside the capsid of the bacteriophage P22. *ACS Nano* **6**, 5000–5009 (2012).

163. Kim, S. *et al.* Double-Chambered Ferritin Platform: Dual-Function Payloads of Cytotoxic Peptides and Fluorescent Protein. *Biomacromolecules* **17**, 12–19 (2016).

164. Lange, C. & Hunte, C. Crystal structure of the yeast cytochrome bc1 complex with its bound substrate cytochrome c. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 2800–2805 (2002).

165. Li, K. *et al.* Cytochrome c deficiency causes embryonic lethality and attenuates stress-induced apoptosis. *Cell* **101**, 389–399 (2000).

166. Margoliash, E. Primary Structure and Evolution of Cytochrome C. *Proceedings of the Proc. Natl. Acad. Sci. U. S. A.* **50**, 672–679

167. Wang, L., Rivera, E. V., Benavides-Garcia, M. G. & Nall, B. T. Loop entropy and

cytochrome c stability. *J. Mol. Biol.* **353**, 719–729 (2005).

168. Acehan, D. *et al.* Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation. *Mol. Cell.* **9**, 423-432 (2002)

169. Jain, R., Kumar, S., Chhabra, R., Agarwal, M. C. & Kumar, R. Analysis of the pH-dependent stability and millisecond folding kinetics of horse cytochrome c. *Arch. Biochem. Biophys.* **585**, 52–63 (2015).

170. Bian, L. & Ji, X. Distribution, transition and thermodynamic stability of protein conformations in the denaturant-induced unfolding of proteins. *PLoS One* **9**, e91129 (2014).

171. Schafer, Z. T. *et al.* Enhanced sensitivity to cytochrome c-induced apoptosis mediated by PHAPI in breast cancer cells. *Cancer Res.* **66**, 2210–2218 (2006).

172. Saxena, M. *et al.* Inducing cell death in vitro in cancer cells by targeted delivery of cytochrome c via a transferrin conjugate. *PLoS One* **13**, e0195542 (2018).

173. Santra, S., Kaittanis, C. & Perez, J. M. Cytochrome c encapsulating theranostic nanoparticles: A novel bifunctional system for targeted delivery of therapeutic membrane-impermeable proteins to tumors and imaging of cancer therapy. *Mol. Pharm.* **7**, 1209–1222 (2010).

174. Dai, X. *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.* **5**, 2929–2943 (2015).

175. Hao, Y., Yu, X., Bai, Y., McBride, H. J. & Huang, X. Cryo-EM Structure of HER2-trastuzumab-pertuzumab complex. *PLoS One* **14**, e0216095 (2019).

176. Muthuswamy, S. K., Gilman, M. & Brugge, J. S. Controlled Dimerization of ErbB Receptors Provides Evidence for Differential Signaling by Homo- and Heterodimers. *Mol. Cell. Biol.* **19**, 6845–6857 (1999).

177. Garrett, T. P. J. *et al.* The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. *Mol. Cell* **11**, 495–505 (2003).

178. Kurokawa, Hirokazu, et al. Inhibition of HER2/neu (erbB-2) and Mitogen-activated Protein Kinases Enhances Tamoxifen Action against HER2-overexpressing, tamoxifen-resistant breast cancer cells1. *Cancer Res.* **60**, 5887-5894 (2000).

179. Clark, A. S., West, K., Streicher, S. & Dennis, P. A. Constitutive and inducible Akt activity promotes resistance to chemotherapy, trastuzumab, or tamoxifen in breast cancer cells. *Mol. Cancer Ther.* **1**, 707–717 (2002).

180. Niu, G. & Carter, W. B. Human epidermal growth factor receptor 2 regulates

angiopoietin-2 expression in breast cancer via AKT and mitogen-activated protein kinase pathways. *Cancer Res.* **67**, 1487–1493 (2007).

181. Petricevic, B. *et al.* Trastuzumab mediates antibody-dependent cell-mediated cytotoxicity and phagocytosis to the same extent in both adjuvant and metastatic HER2/neu breast cancer patients. *J. Transl. Med.* **11**, 1-11 (2013).

182. Junttila, T. T. *et al.* Ligand-Independent HER2/HER3/PI3K Complex Is Disrupted by Trastuzumab and Is Effectively Inhibited by the PI3K Inhibitor GDC-0941. *Cancer Cell* **15**, 429–440 (2009).

183. Lange, T., Nentwich, M. F., Lüth, M., Yekebas, E. & Schumacher, U. Trastuzumab has anti-metastatic and anti-angiogenic activity in a spontaneous metastasis xenograft model of esophageal adenocarcinoma. *Cancer Lett.* **308**, 54–61 (2011).

184. Izumi, Y., Xu, L., Di Tomaso, E., Fukumura, D. & Jain, R. K. Herceptin acts as an anti-angiogenic cocktail. *Nature* **416**, 279–280 (2002).

185. Narayan, M. *et al.* Trastuzumab-lnduced HER reprogramming in 'Resistant' breast carcinoma cells. *Cancer Res.* **69**, 2191–2194 (2009).

186. Vogel, C. L. *et al.* Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J. Clin. Oncol.* **20**, 719–726 (2002).

187. Lewis Phillips, G. D. *et al.* Targeting HER2-positive breast cancer with trastuzumab-DM1, an antibody-cytotoxic drug conjugate. *Cancer Res.* **68**, 9280–9290 (2008).

188. Junttila, T. T., Li, G., Parsons, K., Phillips, G. L. & Sliwkowski, M. X. Trastuzumab-DM1 (T-DM1) retains all the mechanisms of action of trastuzumab and efficiently inhibits growth of lapatinib insensitive breast cancer. *Breast Cancer Res. Treat.* **128**, 347–356 (2011).

189. Barok, M., Tanner, M., Köninki, K. & Isola, J. Trastuzumab-DM1 is highly effective in preclinical models of HER2-positive gastric cancer. *Cancer Lett.* **306**, 171–179 (2011).

190. Scaltriti, M. *et al.* Lapatinib, a HER2 tyrosine kinase inhibitor, induces stabilization and accumulation of HER2 and potentiates trastuzumab-dependent cell cytotoxicity. *Oncogene* **28**, 803–814 (2009).

191. Eigenbrot, C., Ultsch, M., Dubnovitsky, A., Abrahmsén, L. & Härd, T. Structural basis for high-affinity HER2 receptor binding by an engineered protein. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15039–15044 (2010).

192. Nord, K., Nilsson, J., Nilsson, B., Uhlén, M. & Nygren, P. Å. A combinatorial library of an α-helical bacterial receptor domain. *Protein Eng. Des. Sel.* **8**, 601–608 (1995).

193. Nord, K. *et al.* Binding proteins selected from combinatorial libraries of an α-helical bacterial receptor domain. *Nat. Biotechnol.* **15**, 772–777 (1997).

194. Friedman, M. *et al.* Directed Evolution to Low Nanomolar Affinity of a Tumor-Targeting Epidermal Growth Factor Receptor-Binding Affibody Molecule. *J. Mol. Biol.* **376**, 1388–1402 (2008).

195. Wikman, M. *et al.* Selection and characterization of HER2/neu-binding affibody ligands. *Protein Eng. Des. Sel.* **17**, 455–462 (2004).

196. Case, B. A., Kruziki, M. A., Stern, L. A. & Hackel, B. J. Evaluation of affibody charge modification identified by synthetic consensus design in molecular PET imaging of epidermal growth factor receptor. *Mol. Syst. Des. Eng.* **3**, 171–182 (2018).

197. Feldwisch, J. *et al.* Design of an Optimized Scaffold for Affibody Molecules. *J. Mol. Biol* **398**, 232–247 (2010).

198. Orlova, A. *et al.* Synthetic affibody molecules: A novel class of affinity ligands for molecular imaging of HER2-expressing malignant tumors. *Cancer Res.* **67**, 2178–2186 (2007).

199. Sörensen, J. *et al.* First-in-human molecular imaging of HER2 expression in breast cancer metastases using the 111In-ABY-025 affibody molecule. *J. Nucl. Med.* **55**, 730–735 (2014).

200. Zielinski, R. *et al.* Affitoxin-A novel recombinant, HER2-specific, anticancer agent for targeted therapy of HER2-positive tumors. *J. Immunother.* **32**, 817–825 (2009).

201. Tolmachev, V. *et al.* In-Benzyl-DTPA-Z HER2:342 , an Affibody-Based Conjugate for In Vivo Imaging of HER2 Expression in Malignant Tumors. *J. Nucl. Med.* **47**, 846-853 (2006)

202. Teodoro, J. G., Heilman, D. W., Parker, A. E. & Green, M. R. The viral protein Apoptin associates with the anaphase-promoting complex to induce G2/M arrest and apoptosis in the absence of p53. *Genes Dev.* **18**, 1952–1957 (2004).

203. Tolmachev, V. *et al.* Radionuclide therapy of HER2-positive microxenografts using a 177Lu-labeled HER2-specific affibody molecule. *Cancer Res.* **67**, 2773–2782 (2007).

204. Schardt, J. S. *et al.* Engineered Multivalency Enhances Affibody-Based HER3 Inhibition and Downregulation in Cancer Cells. *Mol. Pharm.* **14**, 1047-1056 (2017).

205. Cheng, J. *et al.* Molecular Mechanism of HER2 Rapid Internalization and Redirected Trafficking Induced by Anti-HER2 Biparatopic Antibody. *Antibodies* **9**, 49 (2020).

206. Lee, N. K. *et al.* Ferritin nanocage with intrinsically disordered proteins and affibody: A platform for tumor targeting with extended pharmacokinetics. *J. Control. Release* **267**,

172–180 (2017).

207. Greene, C. J. *et al.* Transferrin receptor 1 upregulation in primary tumor and downregulation in benign kidney is associated with progression and mortality in renal cell carcinoma patients. *Oncotarget* **8**, 107052–107075 (2017).

208. Liang, M. *et al.* H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14900–14905 (2014).

209. Heidari, A., Righetto, I. & Filippini, F. Electrostatic Variation of Haemagglutinin as a Hallmark of the Evolution of Avian Influenza Viruses. *Sci. Rep.* **8**, 1–8 (2018).

210. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**, 125–142 (1984).

211. Brezovsky, J., Kozlikova, B. & Damborsky, J. Computational analysis of protein tunnels and channels. *Methods mol. biol.* **1685**, 25–42 (2018).

212. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, 5.6.1-5.6.37 (2016).

213. Caesar, C. E. B., Esbjörner, E. K., Lincoln, P. & Nordén, B. Assigning membrane binding geometry of cytochrome c by polarized light spectroscopy. *Biophys. J.* **96**, 3399–3411 (2009).

214. Li, L. *et al.* Binding and uptake of H-ferritin are mediated by human transferrin receptor-1. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3505-3510 (2010).

215. Wallace, D. F., McDonald, C. & Subramaniam, V. N. Transferrin Receptor 1: A Ferritin Receptor as Well? *Gastroenterology* **139**, 1052–1053 (2010).

216. Corte-Rodríguez, M., Blanco-González, E., Bettmer, J. & Montes-Bayón, M. Quantitative Analysis of Transferrin Receptor 1 (TfR1) in Individual Breast Cancer Cells by Means of Labeled Antibodies and Elemental (ICP-MS) Detection. *Anal. Chem.* **91**, 15532–15538 (2019).

217. Marques, O. *et al.* Local iron homeostasis in the breast ductal carcinoma microenvironment. *BMC Cancer* **16**, 187 (2016).

218. Lipinski, B. Hydroxyl radical and its scavengers in health and disease. *Oxid. Med. Cell. Longev.* **2011**, 809696 (2011).

219. Liu, H. W. *et al.* A mitochondrial-targeted prodrug for NIR imaging guided and synergetic NIR photodynamic-chemo cancer therapy. *Chem. Sci.* **8**, 7689–7695 (2017).

220. Orino, K., Kamura, S., Natsuhori, M., Yamamoto, S. & Watanabe, K. Two pathways of

iron uptake in bovine spleen apoferritin dependent on iron concentration. *BioMetals* **15**, 59–63 (2002).

221. Radoshitzky, S. R. *et al.* Machupo virus glycoprotein determinants for human transferrin receptor 1 binding and cell entry. *PLoS One* **6**, e21398 (2011).

222. Eckenroth, B. E., Steere, A. N., Chasteen, N. D., Everse, S. J. & Mason, A. B. How the binding of human transferrin primes the transferrin receptor potentiating iron release at endosomal pH. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 13089–13094 (2011).

223. Lei, Y. *et al.* Targeted tumor delivery and controlled release of neuronal drugs with ferritin nanoparticles to regulate pancreatic cancer progression. *J. Control. Release* **232**, 131–142 (2016).

224. Watt, G. D., Jacobs, D. & Frankel, R. B. Redox reactivity of bacterial and mammalian ferritin: Is reductant entry into the ferritin interior a necessary step for iron release? *Proc. Natl. Acad. Sci. U. S. A.* **85**, 7457–7461 (1988).

225. Flick, T. G., Cassou, C. A., Chang, T. M. & Williams, E. R. Solution additives that desalt protein ions in native mass spectrometry. *Anal. Chem.* **84**, 7511–7517 (2012).

226. Hu-Lieskovan, S., Heidel, J. D., Bartlett, D. W., Davis, M. E. & Triche, T. J. Sequence-specific knockdown of EWS-FLI1 by targeted, nonviral delivery of small interfering RNA inhibits tumor growth in a murine model of metastatic Ewing's sarcoma. *Cancer Res.* **65**, 8984–8992 (2005).

227. Vinogradov, S., Batrakova, E., Li, S. & Kabanov, A. Polyion complex micelles with protein-modified corona for receptor-mediated delivery of oligonucleotides into cells. *Bioconjug. Chem.* **10**, 851–860 (1999).

228. Pirollo, K. F. *et al.* Tumor-Targeting Nanoimmunoliposome Complex for Short Interfering RNA Delivery. *Hum. Gene Ther.* **17**, 117-124 (2005).

229. Cui, C. *et al.* Downregulation of TfR1 promotes progression of colorectal cancer via the JAK/STAT pathway. *Cancer Manag. Res.* **11**, 6323–6341 (2019).

230. Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science* **357**, eaan2507 (2017).

231. Holliday, D. L. & Speirs, V. Choosing the right cell line for breast cancer research. *Breast Cancer Res.* **13**, 215 (2011).

232. DeFazio-Eli, L. *et al.* Quantitative assays for the measurement of HER1-HER2 heterodimerization and phosphorylation in cell lines and breast tumors: Applications for diagnostics and targeted drug mechanism of action. *Breast Cancer Res.* **13**, 1-18 (2011).

233. Mazor, Y. *et al.* Enhancement of Immune Effector Functions by Modulating IgG's

Intrinsic Affinity for Target Antigen. *PLoS One* **11**, e0157788 (2016).

234. Hendriks, B. S. *et al.* Impact of Tumor HER2/ERBB2 Expression Level on HER2-Targeted Liposomal Doxorubicin-Mediated Drug Delivery: Multiple Low-Affinity Interactions Lead to a Threshold Effect. *Mol. Cancer Ther.* **12**, 1816-1828 (2013).

235. Rodallec, A. *et al.* Docetaxel–trastuzumab stealth immunoliposome: Development and in vitro proof of concept studies in breast cancer. *Int. J. Nanomedicine* **13**, 3451–3465 (2018).

236. Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515–527 (2006).

237. Wen, C. C. *et al.* SV40 T/t-common polypeptide specifically induces apoptosis in human cancer cells that overexpress HER2/neu. *Cancer Res.* **66**, 5847–5857 (2006).

238. Takagi, S. *et al.* HER2 and HER3 cooperatively regulate cancer cell growth and determine sensitivity to the novel investigational EGFR/HER2 kinase inhibitor TAK-285. *Oncoscience* **1**, 196–204 (2014).

239. Subik, K. *et al.* The Expression Patterns of ER, PR, HER2, CK5/6, EGFR, Ki-67 and AR by Immunohistochemical Analysis in Breast Cancer Cell Lines. *Breast Cancer Basic Clin. Res.* **4**, 117822341000400 (2010).

240. McDermott, M. S. J. *et al.* CDK7 Inhibition is Effective in all the Subtypes of Breast Cancer: Determinants of Response and Synergy with EGFR Inhibition. *Cells* **9**, (2020).

241. Jeong, H. *et al.* Neuregulin-1 induces cancer stem cell characteristics in breast cancer cell lines. *Oncol. Rep.* **32**, 1218–1224 (2014).

242. Breen, A. F. *et al.* Protein Encapsulation of Experimental Anticancer Agents 5F 203 and Phortress: Towards Precision Drug Delivery. *Int. J. Nanomedicine* **14**, 9525–9534 (2019).

243. Jhaveri, M. S., Rait, A. S., Chung, K.-N., Trepel, J. B. & Chang, E. H. Antisense oligonucleotides targeted to the human alpha folate receptor inhibit breast cancer cell growth and sensitize the cells to doxorubicin treatment. *Mol. Cancer Ther.* **3**, 1505–12 (2004).

244. Kawamoto, M., Horibe, T., Kohno, M. & Kawakami, K. A novel transferrin receptor-targeted hybrid peptide disintegrates cancer cell membrane to induce rapid killing of cancer cells. *BMC Cancer* **11**, 359 (2011).

245. Hathaway, H. J. *et al.* Detection of breast cancer cells using targeted magnetic nanoparticles and ultra-sensitive magnetic field sensors. *Breast Cancer Res.* **13**, 1-13 (2011).

246. Brockhoff, G., Hei, P., Schlegel, J., Hofstaedter, F. & Knuechel, R. Epidermal growth factor receptor, c-erbB2 and c-erbB3 receptor interaction, and related cell cycle kinetics of SK-BR-3 and BT474 breast carcinoma cells. *Cytometry* **44**, 338–348 (2001).

247. Rodríguez, C. E. *et al.* Breast cancer stem cells are involved in Trastuzumab resistance through the HER2 modulation in 3D culture. *J. Cell. Biochem.* **119**, 1381–1391 (2018).

248. Mazzucchelli, S. *et al.* H-Ferritin-nanocaged olaparib: A promising choice for both BRCA-mutated and sporadic triple negative breast cancer. *Sci. Rep.* **7**, 7505 (2017).

249. Bonizzi, A. *et al.* Everolimus Nanoformulation in Biological Nanoparticles Increases Drug Responsiveness in Resistant and Low-Responsive Breast Cancer Cell Lines. *Pharmaceutics* **11**, 384 (2019).

250. Demogines, A., Abraham, J., Choe, H., Farzan, M. & Sawyer, S. L. Dual Host-Virus Arms Races Shape an Essential Housekeeping Protein. *PLoS Biol.* **11**, e1001571 (2013).

251. Wikman, M. *et al.* Selection and characterization of HER2/neu-binding affibody ligands. *Protein Eng. Des. Sel.* **17**, 455–462 (2004).

252. Altai, M. *et al.* Affibody-derived drug conjugates: Potent cytotoxic molecules for treatment of HER2 over-expressing tumors. *J. Control. Release* **288**, 84–95 (2018).

253. Mojia, M. *et al.* Extracellular iron diminishes anticancer effects of vitamin C: An in vitro study. *Sci. Rep.* **4**, 1–8 (2014).

254. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: A summary and pharmacological classification. *Nature Reviews Drug Discovery* **7**, 21–39 (2008).

255. Guo, M., Bhaskar, B., Li, H., Barrows, T. P. & Poulos, T. L. Crystal structure and characterization of a cytochrome c peroxidase-cytochrome c site-specific cross-link. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 5940–5945 (2004).

256. Iwashita, S. & Kanegasaki, S. Enzymic and molecular properties of base-plate parts of bacteriophage P22. *Eur. J. Biochem.* **65**, 87–94 (1976).

257. Chen, H., Zhang, S., Xu, C. & Zhao, G. Engineering protein interfaces yields ferritin disassembly and reassembly under benign experimental conditions. *Chem. Commun.* **52**, 7402–7405 (2016).

258. Kim, M. *et al.* PH-dependent structures of ferritin and apoferritin in solution: Disassembly and reassembly. *Biomacromolecules* **12**, 1629–1640 (2011).

259. Shelake, R. M., Ito, Y., Masumoto, J., Morita, E. H. & Hayashi, H. A novel mechanism of metal gel-shift by histidine-rich Ni2+-binding Hpn protein from Helicobacter pylori strain SS1. *PLoS One* **12**, e0172182 (2017).

260. Brezovsky, J. *et al.* Software tools for identification, visualization and analysis of protein

tunnels and channels. *Biotechnol. Adv.* **31**, 38–49 (2013).

261. Santambrogio, P. *et al.* Effects of modifications near the 2-, 3- and 4-fold symmetry axes on human ferritin renaturation. *Biochem. J.* **322**, 461–468 (1997).

262. Gaowa, S. *et al.* Possible role of p53/Mieap-regulated mitochondrial quality control as a tumor suppressor in human breast cancer. *Cancer Sci.* **109**, 3910–3920 (2018).

263. Jones, S., Holm, T., Mäger, I., Langel, Ü. & Howl, J. Characterization of bioactive cell penetrating peptides from human cytochrome c: Protein mimicry and the development of a novel apoptogenic agent. *Chem. Biol.* **17**, 735–744 (2010).

264. Breast cancer | McMaster Pathophysiology Review. Available at: http://www.pathophys.org/breast-cancer/. (Accessed: 18th June 2020)

265. Kaddis, C. S. *et al.* Sizing Large Proteins and Protein Complexes by Electrospray Ionization Mass Spectrometry and Ion Mobility. *J. Am. Soc. Mass Spectrom.* **18**, 1206–1216 (2007).

266. Bresters, Tjarda W., et al. The Pyruvate-Dehydrogenase Complex from Azotobacter vinelandii. 1. Purification and Properties. *Eur. J. Biochem.* **59**, 335–345 (1975).

267. Waldmann, T. *et al.* The Thermosome of Thermoplasma acidophilum and Its Relationship to the Eukaryotic Chaperonin TRiC. *Eur. J. Biochem.* **227**, 848–856 (2008).

268. Wang, Q., Kaltgrad, E., Lin, T., Johnson, J. E. & Finn, M. . Natural Supramolecular Building Blocks: Wild-Type Cowpea Mosaic Virus. *Chem. Biol.* **9**, 805–811 (2002).

269. Bockstahler, L. E. & Kaesberg, P. The Molecular Weight and Other Biophysical Properties of Bromegrass Mosaic Virus. *Biophys. J.* **2**, 1–9 (1962).

270. Vriend, G., Hemminga, M. A., Verduin, B. J. M. & Schaafsma, T. J. Swelling of cowpea chlorotic mottle virus studied by proton nuclear magnetic resonance. *FEBS Lett.* **146**, 319–321 (1982).

271. Rhee, J.-K. *et al.* Colorful virus-like particles: fluorescent protein packaging by the Qβ capsid. *Biomacromolecules* **12**, 3977–81 (2011).

272. Kuzmanovic, D. A., Elashvili, I., Wick, C., O'Connell, C. & Krueger, S. Bacteriophage MS2: Molecular Weight and Spatial Distribution of the Protein and RNA Components by Small-Angle Neutron Scattering and Virus Counting. *Structure* **11**, 1339–1348 (2003).

273. Ruigrok, R. W., Nermut, M. V & Andree, P. J. The molecular mass of adenovirus type 5 as determined by means of scanning transmission electron microscopy (STEM). *J. Virol. Methods* **9**, 69–78 (1984).

274. Tanaka, H. & Tsukihara, T. Structural studies of large nucleoprotein particles, vaults.

Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci. **88**, 416–33 (2012).

275. Rother, M., Nussbaumer, M. G., Renggli, K. & Bruns, N. Protein cages and synthetic polymers: a fruitful symbiosis for drug delivery applications, bionanotechnology and materials science. *Chem. Soc. Rev.* **45**, 6213–6249 (2016).

276. Bou-Abdallah, F. The iron redox and hydrolysis chemistry of the ferritins. *Biochimica et Biophysica Acta - General Subjects* **1800**, 719–731 (2010).

277. Fletcher, J. M. *et al.* A basis set of de novo coiled-Coil peptide oligomers for rational protein design and synthetic biology. *ACS Synth. Biol.* **1**, 240–250 (2012).

278. Liang, M. *et al.* H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection. *Proc. Natl. Acad. Sci.* **111**, 14900–14905 (2014).

7.0 Supplementary

7.1 Chapter 3 supplementary figures
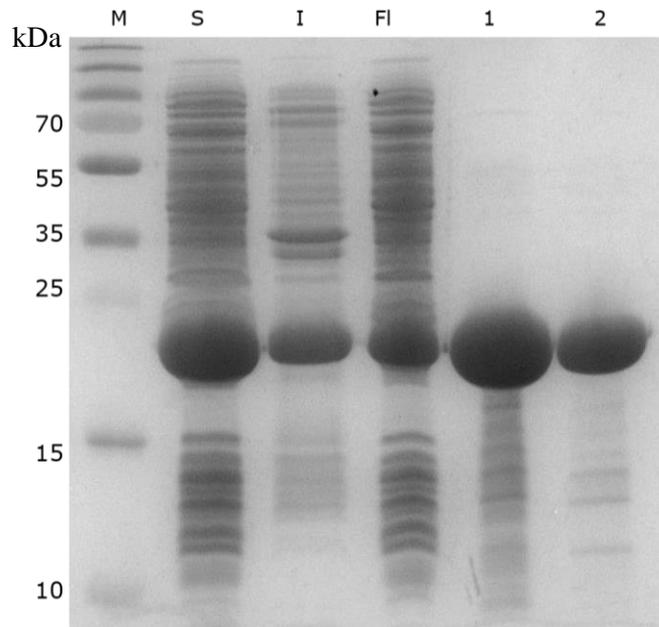
SFigure 3.1. SDS page of purified huHC-.HUHC- expression at 37 °C for 3 hrs M: Marker, S: Soluble, I: Insoluble, Fl; Flow through and 1-2: Elution.

FeO-

MHHHHHHGLN DIFEAQKIEW HELVPRGSMT TASTSQVRQN YHQDSEAAIN RQINLELYAS YVYLSMSYYF DRDDVALKNF AKYFLHQSHA KRAGAAKLMK LQNQRGGRIF LQDIKKPDRD DWESGLNAME AALHLEKNVN QSLLELHKLA TDKNDPHLSD FIETHYLNEQ VKAIKELGDH VTNLRKMGAP ESGLAEYLFD KHTLGDSDNE S

Molecular weight: 24350   **Molar Extinction coefficient:** 24410   **Theoretical pI:** 6.13

-

TfR1-

MHHHHHHGLN DIFEAQKIEW HELVPRGSMT TASTSQVRQN YHQDSEAAIN RQINLELYAS
YVYLSMSYYF DRDDVALKNF AKYFLHQSHA KRAGAAKLMK LQNQRGGRAL FQDIKKPDRD
DWESGLNAME AALHLEKNVN QSLLELHKLA TDKNDPHLSD FIETHYLNEQ VKAIKELGDH
VTNLRKMGAP ESGLAEYLFD KHTLGDSDNE S

Molecular weight: 24308 **Molar Extinction coefficient:** 24410 **Theoretical pI:** 6.13

TfR1- K68C

MHHHHHHGLN DIFEAQKIEW HELVPRGSMT TASTSQVRQN YHQDSEAAIN RQINLELYAS
YVYLSMSYYF DRDDVALKNF AKYFLHQSHA KRAGAACLMK LQNQRGGRAL FQDIKKPDRD
DWESGLNAME AALHLEKNVN QSLLELHKLA TDKNDPHLSD FIETHYLNEQ VKAIKELGDH
VTNLRKMGAP ESGLAEYLFD KHTLGDSDNE S

Molecular weight: 24283 **Molar Extinction coefficient:** 24410 **Theoretical pI:** 6.03

TfR1- K68C (-HIS-tag)

MTTASTSQVRQN YHQDSEAAIN RQINLELYAS YVYLSMSYYF DRDDVALKNF AKYFLHQSHA
KRAGAACLMK LQNQRGGRAL FQDIKKPDRD DWESGLNAME AALHLEKNVN QSLLELHKLA
TDKNDPHLSD FIETHYLNEQ VKAIKELGDH VTNLRKMGAP ESGLAEYLFD KHTLGDSDNE S

Molecular weight: 20908 **Molar Extinction coefficient:** 18910 **Theoretical pI:** 5.82

----

huHC- K68C

MHHHHHHGLN DIFEAQKIEW HELVPRGSMT TASTSQVRQN YHQDSEAAIN RQINLELYAS
YVYLSMSYYF DRDDVALKNF AKYFLHQSHE EREHAECLMK LQNQRGGRIF LQDIKKPDRD
DWESGLNAME AALHLEKNVN QSLLELHKLA TDKNDPHLSD FIETHYLNEQ VKAIKELGDH
VTNLRKMGAP ESGLAEYLFD KHTLGDSDNE S

Molecular weight: 24580 **Molar Extinction coefficient:** 24410 **Theoretical pI:** 5.66

MTTASTSQVRQN YHQDSEAAIN RQINLELYAS YVYLSMSYYF DRDDVALKNF AKYFLHQSHE

EREHAECLMK LQNQRGGRIF LQDIKKPDRD DWESGLNAME AALHLEKNVN QSLLELHKLA

TDKNDPHLSD FIETHYLNEQ VKAIKELGDH VTNLRKMGAP ESGLAEYLFD KHTLGDSDNE S

Molecular weight: 21205   **Molar Extinction coefficient:** 18910   **Theoretical pI:** 5.31

SFigure 3.2. Plasmid construct and protein properties of ferritin variants. ORANGE BOLD: His-tag, ORANGE: AVI-tag, RED: thrombin cleavage, BLUE: Ferroxidase mutation huH 222, GREEN: K68C mutation and YELLOW: Transferrin TfR- mutation Data calculated using Protein Identification and Analysis Tools on the ExPASy Server.

Method	% of MW Species (Da) present after MDCC conjugation			
	24283	24328	24666	Other species
Urea dialysis	0	43	14	43
Ammonium sulphate precipitation	27	0	49	24

SFigure 3.3. MW of predominant species present in each conjugation trial. Urea dialysis: TfR- K68C was incubated in EDTA overnight at 4 °C. TfR- K68C was then dialyzed against 8 M urea 20 mM Tris pH 7.2 overnight at 4˚C. TfR- K68C was incubated with a 10x excess of Tris(2-carboxyethyl)phosphine (TCEP) for 1 hr at room temperature and then 10x 7-Diethylamino-3-[*N*-(2-maleimidoethyl)carbamoyl]coumarin MDCC was added and incubate overnight at 4˚C. Ammonium sulphate cut conjugation: TfR- K68C was preceiptated with ammonium sulphate and resuspended in 20 mM Tris pH 7.5 to remove 8 M urea and (MDCC).

a	TfR1		µg/ml of R-PE in solution (1)	nM of R-PE (2)	nM of IgG1 (3)	Ratio of R-PE/IgG1
Test 1	IgG1 mAb (µg/ml)	1.0				
	Average abs	36357.3	1.3	5.6	6.7	0.8
Test 2	IgG1 mAb (µg/ml)	0.1				
	Average abs	4841.7	0.1	0.6	0.7	0.9

b	HER2		µg/ml of R-PE in solution (1)	nM of R-PE (2)	nM of IgG1 (3)	Ratio of R-PE/IgG1
Test 1	IgG1 mAb (µg/ml)	1.0				
	Average abs	16571.3	1.5	6.3	6.7	0.9
Test 2	IgG1 mAb (µg/ml)	0.1				
	Average abs	1558.0	0.1	0.4	0.7	0.6

SFigure 3.4. Calculating R-PE: Igg1 ratio. **a** TfR1 and **b** HER2 (1): Equation y= (x-1146.8)/26355, where x = abs and y = µg/ mL of R-PE, (2): nM of R-PE in mAb calculated using MW of R-PE 240000 Da, (3): nM of IgG1 mAb calculated using MW of IgG1 150000 Da.
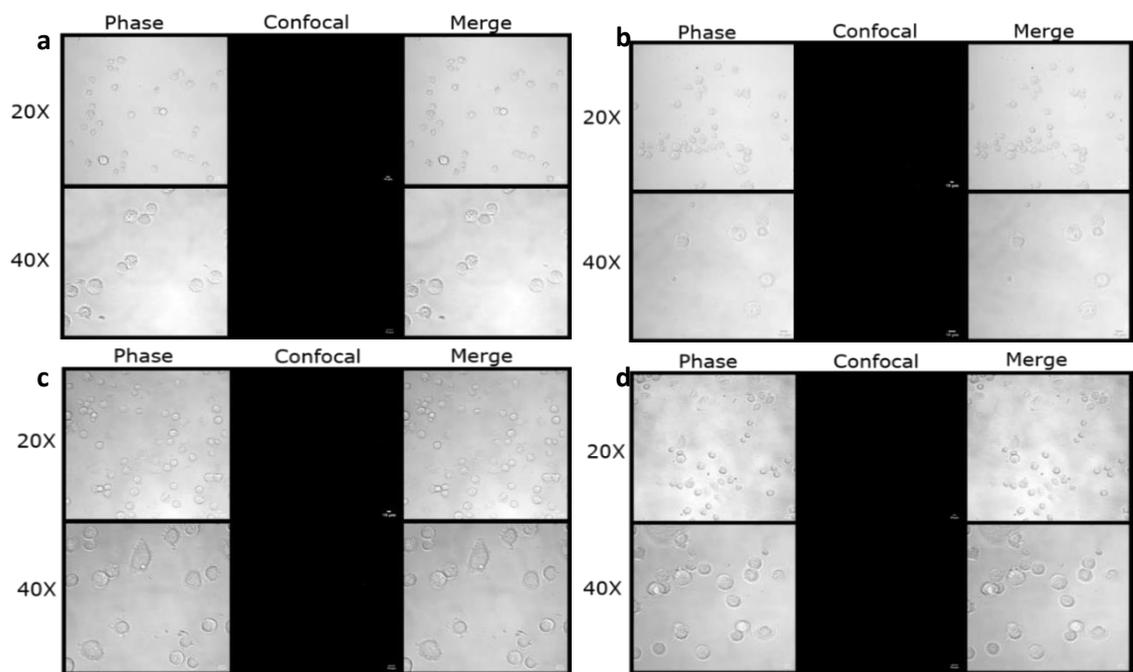
SFigure 3.5 Control internalisation experiment of apoferritin variants by SKBR3 cells. Ferritin variants did not have were **did** contain the K68C mutation but **did not** undergo conjugation. SKBR3 cells were incubated 60 μg/ mL of apoferritin for 1 hr. Images were taken at X20 and X40 under confocal (488 nm channel) and phase-contrast microscopy. Scale bar 10 μm. **a:** huHC-, **b:** TfR-, **c:** huHC-(-HIS) and
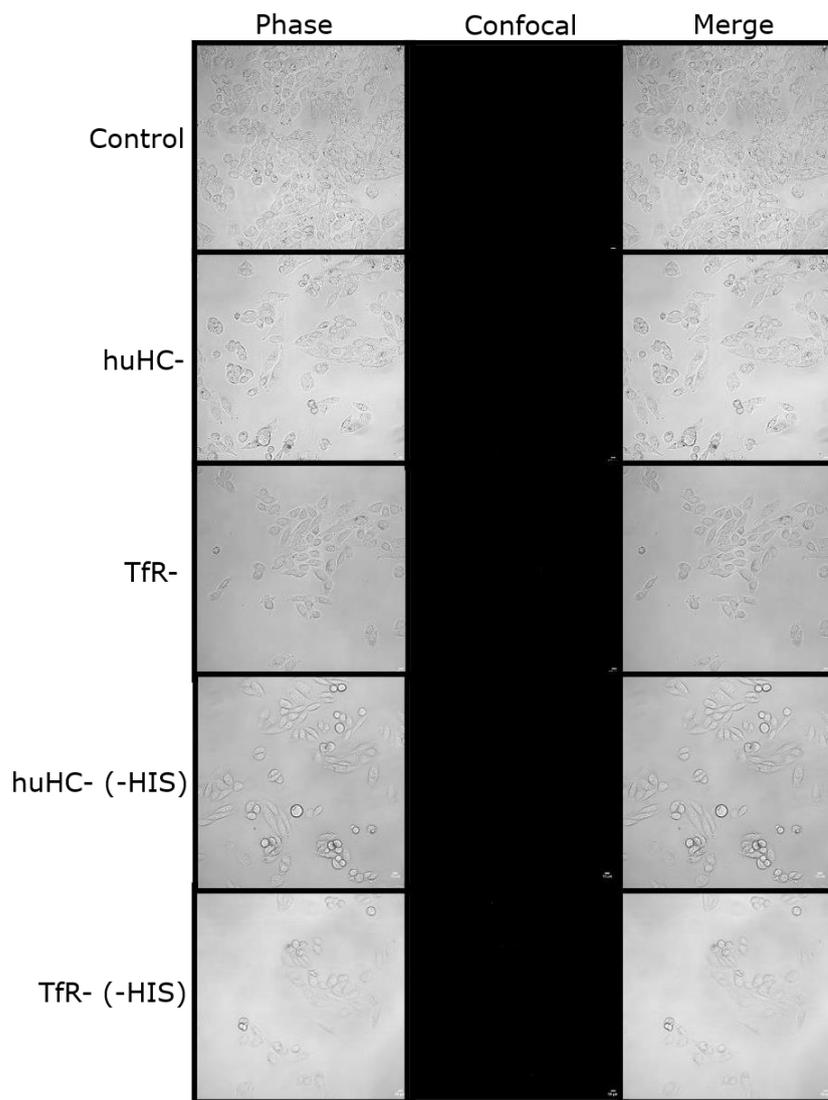
SFigure 3.6 Control internalisation experiment of apoFt variants by MRC-5 cells. apoFt variants **did** contain the K68C mutation and **did** undergo conjugation. MRC-5 cells were incubated 60 μg/ mL of apoFt for 1 hr. Images were taken at X20 and under confocal (488 nm channel) and phase-contrast microscopy. Scale bar 10 μm.

huHC- Afb A102C

MHHHHHHSSDE DEGENLYFQS TGSEAENEKF NKEMRNAYW EIALLPNLNNQ QKRAFIRSLY
DDPSQSANLLAE AKKLNDAQA PKVDGGGGSG GSMTTASTS QVRQNYHQDSE AAINRQINLE
LYASYVYLSMS YYFDRDDVAL KNFAKYFLHQ SHEEREHAE KLMKLQNQRGG RIFLQDIKK
PDRDDWESGLN AMECALHLEK NVNQSLLELH KLATDKNDP HLSDFIETHYLN EQVKAIKELG
DHVTNLRKMG APESGLAEYLFD KHTLGDSDNE S

Molecular weight: 31725   **Molar Extinction coefficient:** 28880   **Theoretical pI:** 5.41

Afb S42C

MHHHHHHSSDE DEGENLYFQS TGSEAENEKF NKEMRNAYW EIALLPNLNNQ QKRAFIRSLY
DDPSQCANLLAE AKKLNDAQA PK

Molecular weight: 9767   **Molar Extinction coefficient:** 9970   **Theoretical pI:** 5.59

TfR1- Afb A102C

MHHHHHHSSDE DEGENLYFQS TGSEAENEKF NKEMRNAYW EIALLPNLNNQ QKRAFIRSLY
DDPSQSANLLAE AKKLNDAQA PKVDGGGGSG GSMTTASTS QVRQNYHQDSE AAINRQINLE
LYASYVYLSMS YYFDRDDVAL KNFAKYFLHQ SHEEREHAE KLMKLQNQRGG RALFQDIKK
PDRDDWESGLN AMECALHLEK NVNQSLLELH KLATDKNDP HLSDFIETHYLN EQVKAIKELG
DHVTNLRKMG APESGLAEYLFD KHTLGDSDNE S

Molecular weight: 31684   **Molar Extinction coefficient:** 28880   **Theoretical pI:** 5.41

SFigure 3.7. Plasmid construct and protein properties of affibody containing variants. ORANGE BOLD: His-tag, Grey/White: TEV cleavage, Green: $Z_{HER2:342}$ affibody, GREEN: A102C mutation and YELLOW: Transferrin TfR- mutation. Data calculated using Protein Identification and Analysis Tools on the ExPASy Serve

7.2   Chapter 4 supplementary figures

huHC-

MHHHHHHHGLN DIFEAQKIEW HELVPRGSMT TASTSQVRQN YHQDSEAAIN RQINLELYAS YVYLSMSYYF DRDDVALKNF AKYFLHQSHE EREHAEKLMK LQNQRGGRIF LQDIKKPDRD DWESGLNAME AALHLEKNVN QSLLELHKLA TDKNDPHLSD FIETHYLNEQ VKAIKELGDH VTNLRKMGAP ESGLAEYLFD KHTLGDSDNE S

Molecular weight: 24605   **Molar Extinction coefficient:** 24410   **Theoretical pI:** 5.74

----

huHC- (-HIS)

MT TASTSQVRQN YHQDSEAAIN RQINLELYAS YVYLSMSYYF DRDDVALKNF AKYFLHQSHE EREHAEKLMK LQNQRGGRIF LQDIKKPDRD DWESGLNAME AALHLEKNVN QSLLELHKLA TDKNDPHLSD FIETHYLNEQ VKAIKELGDH VTNLRKMGAP ESGLAEYLFD KHTLGDSDNE S

Molecular weight: 21230   **Molar Extinction coefficient:** 18910   **Theoretical pI:** 5.42

Δ160

MHHHHHHHGLN DIFEAQKIEW HELVPRGSMT TASTSQVRQN YHQDSEAAIN RQINLELYAS YVYLSMSYYF DRDDVALKNF AKYFLHQSHE EREHAEKLMK LQNQRGGRIF LQDIKKPDRD DWESGLNAME AALHLEKNVN QSLLELHKLA TDKNDPHLSD FIETHYLNEQ VKAIKELGDH VTNLRKMG

Molecular weight: 22127   **Molar Extinction coefficient:** 22920   **Theoretical pI:** 6.12

ΔDE

MT TASTSQVRQN YHQDSEAAIN RQINLELYAS YVYLSMSYYF DRDDVALKNF AKYFLHQSHA KRAGAAKLMK LQNQRGGR ALFQDIKKPD RDDWESGLNA MEAALHLEKN VNQSLLELHK LATDKNDPHL SDFIETHYLN EQVKAIKELG DHVTNLRKMG

Molecular weight: 18455   **Molar Extinction coefficient:** 17420   **Theoretical pI:** 7.11

Figure 4.1 Physio-chemical properties of apoFt variants. Histidine-tag (Orange -Bold), Avitidine-tag (Orange), Thrombin cleavage site (Red), Ferroxidase mutation (Blue-Underline) and Transferrin receptor-1 mutations(Yellow-Underline)

8.0 PIP reflective Statement

As part PIP training requirement, I undertook a ~3-month placement between September-December 2020. The PIP was carried out in a spin off company called 'Hy2care'. The Hy2care PIP placement was advertised by Futform, a research and Innovation Staff Exchange project by the EU. Hy2care is a spin off from Twente university and is also based in the University of Twente, Enschede, Netherlands.

The company's primary product is '...gelating bioresorbable hydrogels for the arthroscopic treatment of cartilage defects, aiming at functional cartilage repair.' Due to the IP nature of the company, a basic description of the aims and outcomes will be included in the statement below. compound: polymer ratio. In terms of the outcome, I did not manage to achieve the exact compound/polymer ratio, but I did manage to identify an alternative route that could possibly be efficient for achieving

this ratio in a controlled manner compared to the standard route. A personal objective I wanted to achieve was to get a better insight into the dynamics of a spin-off company including the relationship between the company and the university, as well as how they strategically market their product. As the main office was located at a different site and I was a temporary member of the company, this objective wasn't really achieved.

Generally speaking, moving to a new country with no prior contacts was initially very stressful. However, because the company was based in a university, it was much easier to meet and connect with people of a similar age demographic. As such it become a more enjoyable and rewarding experience. Even though this experience is not comparable to an individual who is relocating to another country for work, I believe it has given me a good insight. In terms of career path, I feel I am more comfortable with pursuing a biotech job in Europe.