



THE UNIVERSITY
of ADELAIDE

**ENGINEERED HUMAN HEAVY-CHAIN FERRITIN TO IMPROVE
ANTI-TUMOUR DRUG DELIVERY PERFORMANCE**

by

Shuang Yin (a1714202)

School of Chemical Engineering & Advanced Materials
Faculty of Engineering, Computer and Mathematical Sciences
The University of Adelaide

A thesis submitted for the degree of Doctor of Philosophy

February 2021

DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signature:



Date: 11-02-2021

EXECUTIVE SUMMARY

Ferritin exists ubiquitously in nature. It has been used in medical application for decades because of its unique and promising plastic structure, together with high biocompatibility and stability. The most widely adopted mammalian ferritin, is human heavy-chain ferritin (HF_n), which has been employed as an anti-tumour drug loading platform. HF_n comprises 24 subunits to form a 12 nm hollow sphere. Various kinds of anti-tumour drugs have been explored to be loaded onto HF_n by diverse drug loading approaches. However, protein loss and undesirable drug loading ratio are bottlenecks to practical drug loading. In addition, HF_n has shortcomings such as short half-life in circulation. Different half-life extenders or other functional characteristics such as enhanced tumour targeting ability, have been equipped to HF_n to improve its efficacy and broaden its application.

In this research thesis, focus is on improvement of HF_n performance as an anti-tumour drug delivery platform through 1) functionalisation by fusion, and 2) drug loading approach. Functionalisation and drug loading investigation were initially conducted in 2 strategies and compared. The first strategy tries to extend the half-life of HF_n, improve tumour targeting ability, and enhance the drug loading performance. It fuses 2 functional peptides, PAS and RGDK, to HF_n. PAS is a 40 aa peptide composed of repetitive P, A, and S residues. It can bind to water molecules and enlarge the hydrodynamic volume to extent half-life in circulation. RGDK is a four-residue tumour targeting peptide. A thermally induced passive diffusion approach was explored to load small molecule doxorubicin hydrochloride (DOX) to HF_n to improve drug loading performance. The second strategy fuses HF_n with a peptide drug called pro-apoptotic peptide (P). The sequence of P is KLAKLAKKLAKLAK. It can selectively kill tumour cells by disrupting mitochondrion membrane. This strategy makes use of the alternative fusion approach to load drug to guarantee loading efficiency and simplify loading process. RGDK was also employed in this strategy to improve HF_n tumour targeting ability.

Eight HF_n based proteins were designed for these 2 strategies. Two of them are controls, HF_n and E-helix truncated HF_n (sHF_n). Three PAS functionalised HF_ns are HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK. GFLG and PLGLAG are two enzyme-cleavable sites leading to the removal of PAS-RGDK from HF_n after reaching tumour tissues. They are 4 aa and 6 aa long, respectively. sHF_n is not employed in strategy 1 and HF_n is the only control.

Another 3 proteins P-HFn, sHFn-P, and sHFn-P-RGDK. HFn and sHFn are used and they are both controls.

All proteins were successfully expressed in *Escherichia coli* (*E. coli*). Transmission electron microscopy (TEM) and size-exclusion chromatography (SEC) results showed that 2 controls and 3 PAS functionalised HFns were expressed in soluble assembly. However, 3 P functionalised HFns failed to self-assemble. P-HFn was expressed in inclusion bodies (IBs), and sHFn-P and sHFn-P-RGDK in soluble monomers. Without the correct structure, 3 P functionalised HFns could not be applied as anti-tumour drugs. Therefore, strategy 2 failed and there was no need for the 3 P functionalised HFns to undergo further bioactivity evaluations. Molecular docking and molecular dynamic (MD) simulation were conducted to investigate the impacts of P peptide on HFn assembling. Results demonstrate that the high density of positive charge and the hydrophobicity of P were the main causes for the failure of P functionalisation.

In strategy 1, the 3 PAS functionalised HFns and their control HFn underwent further experimental investigations. Following the expression, HFn and 3 PAS functionalised HFns were purified using heat-acid precipitation and chromatography. Purification of functionalised HFns were optimised and it is worth to note that the inserted functional peptides resulted in the conformation change and therefore impacted the purification process. Fusion of PAS peptide has led to an E-helix turnover phenomenon, resulting in a significant decline of the stability. In structural comparisons of HFn and 3 PAS functionalised HFns, PAS has enlarged the hydrodynamic volumes of assemblies due to its hydration ability. However, RGDK peptide did not make any significant impact on HFn conformation.

Following preparation of high-quality HFn-based proteins, a thermally induced passive diffusion approach was introduced to load drug efficiently. Conditions in thermally induced passive diffusion approach was optimised for both HFn and HFn-GFLG-PAS-RGDK through an orthogonal trial. Variables were heating temperature, heating time and buffer pH. Condition of pH 7.5, 50 °C, 6 h was found to be the optimal for all 4 HFn-based proteins. It achieved a significantly improved protein recovery yield and loading ratio in comparison with previous loading approaches. Protein recovery yield of functionalised HFns were lower in contrast with HFn, because of the stability decrease after fusion. Molecular docking combined with MD simulation revealed the mechanisms behind DOX loading process.

Finally, the anti-tumour performances of DOX loaded 4 HFns (i.e. DOX loaded HFn and 3 PAS functionalised HFns) were assessed in detail *in vitro* and *in vivo* tests. Intracellular distribution assay demonstrated DOX carried by these 4 HFn-based proteins could accumulate in tumour cell nucleus to exert function. MTT assay showed that RGDK in 2 PAS-RGDK functionalised HFns significantly improved the cytotoxicity. Cellular uptake test showed that RGDK had significantly enhanced tumour cell uptake efficiency. In animal tests, biodistribution results proved the improvement of tumour targeting ability by RGDK in HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK, compared with HFn and HFn-PAS. Pharmacokinetic test showed an approximate 5 times extension of half-life in circulation resulting from PAS peptide fusion. In anti-tumour growth test, 2 PAS-RGDK functionalised HFns achieved the highest anti-tumour efficacy.

Overall, this work has significantly improved HFn as an anti-tumour drug delivery platform. Functionalisation with PAS and RGDK peptide can improve anti-tumour drug delivery performance through half-life extension and improvement of tumour targeting. Thermally induced drug loading approach can be used in both HFn and functionalised HFns to obtain a desirable loading performance. Mechanism of HFn drug loading has been revealed by experiments aided with computation tools including molecular docking and MD simulation.

The findings will be of immediate benefit and interest to a wide range of researchers and manufacturers for drug delivery using ferritin.

ACKNOWLEDGEMENTS

Throughout this PhD candidature, as a joint program student, I have received countless academic and life supports and enlightenments from supervisors, professors and team members from The University of Adelaide and Institute of Process Engineering (IPE), Chinese Academy of Sciences (CAS). I am very grateful that I had the chance to study here.

My principal Supervisor Associate Prof Jingxiu Bi has always been encouraging and responsive. During this four-year study, our pharmaceutical laboratory has gone through many changes and upgrades. Jingxiu always tries her best to provide a better platform for students to conduct research, such as our study trip to IPE and Shanxi. Her view of research has helped me understand more about what research is and how I should conduct it. She also helped me in life, hosting me at home when I did not find a place to live. My external supervisor Prof Yongdong Liu has been supportive throughout my candidature. She has offered constructive suggestions on my research project task design and along the progression. She also kindly provided accommodation and experiment facilities when I was in IPE. Co-supervisor Prof Sheng Dai helped me with my academic writing and from which I learned a lot.

Associate Prof Kenneth Davey, FIChemE, CEng, CSci, has been a great help to my academic writing, especially in terms of the logical structure of drafts and the selection of academic words. His wisdom, advice, together with passion and strictness on work have enlightened me. Great appreciations to Prof Anton Middelberg and Prof Woo-Seok Choe, who have given me valuable advice in the progress of my research project. Many thanks to the kind facility help from Prof Yingli Wang and Associate Prof Yan Wang from Shanxi Traditional Chinese Medicine University.

I am grateful to the all the staff members in School of Chemical Engineering & Advanced Materials for both Administration and Laboratory services, and Adelaide Graduation Centre staff for their support in every milestone.

My gratitude to my team members here, Bingyang Zhang, Yiran Qu, Thai Thao Ly, Nhat Hoang Huynh, Afshin Karami, Lukas Gerstweiler, Yechuan Zhang, Seonho Yun. Team members in IPE, CAS, Yao Zhang, Chun Zhang, Qi Wang, Fangxia Guo, Tong Ye, Xue Liu,

Weiying Wang, plus team members in Shanxi, Hong Luo, Yafei Guo, Jiaming Wang, Yue Wu and Jing Zhao. I am blessed to have such warm and diverse groups to work in. They have inspired me with good qualities, such as resilience, self-confidence, perseverance, self-motivation and communicational skills.

I also greatly appreciate all the emotional support and kind advice from my sincere friends. Liping Liu and Tianshu Zhang accompanied me to explore the Australian culture and countless beautiful sceneries, which greatly energised me. We also exchanged our PhD research and life experiences, which supported me during hard times. Thanks to Beiha Yanez, my best Australian friend, who helped me a lot in English speaking and understanding local culture, invited me over to home parties and made me feel at home. Thanks for the company of Jun Xie when I was in IPE, it means a lot to me. Thanks to Phyllis Luo for sharing your wonderful and unusual life experience that encouraged me.

Lastly, I would like to express my gratefulness to my family members. Thanks to my parents Shuping Wang and Guoqing Yin, my boyfriend Kaiwen Dang, my brother Teng Yin, sister-in-law Xiaohui Ying and my lovely niece Pengpeng for your unconditional supports during this PhD candidature, that allowed me to focus on study without worrying too much.

TABLE OF CONTENTS

DECLARATION	ii
EXECUTIVE SUMMARY	iii
ACKNOWLEDGEMENTS	vi
TABLE OF CONTENTS	viii
LIST OF FIGURES	xiv
LIST OF TABLES	xviii
ABBREVIATIONS	xx
Chapter 1	1
INTRODUCTION	1
1.1 Background	2
1.2 Aim and objectives	3
1.3 Thesis outline	5
1.4 References	7
Chapter 2	9
LITERATURE REVIEW	9
2.1 Introduction	10
2.2 Ferritin structure and properties: the basis of a promising drug nanocarrier	11
2.2.1 Structure	11
2.2.2 Chemical and biological properties	14
2.3 Ferritin drug-loading approaches and mechanisms	15
2.3.1 Passive-loading	17
2.3.2 pH-induced disassembly-reassembly	19
2.3.3 8 M Urea-induced disassembly-reassembly	20
2.3.4 Alternative drug-loading approaches	22

2.3.5 Selection and comparison of drug-loading approaches	22
2.3.6 Improvement of drug loading performance	25
2.4 Ferritin functionalisation.....	26
2.4.1 Chemical conjugation	27
2.4.2 Gene engineering technology	28
2.4.3 Hybridisation.....	32
2.5 Computational tools used in prediction of protein conformation, property and protein-ligand interaction	42
2.5.1 New protein conformation prediction by MD simulation.....	42
2.5.2 Protein-ligand interaction study by molecular docking	46
2.6 Chapter summary	48
2.7 References.....	50
Chapter 3.....	61
MOLECULE DESIGN, EXPRESSION OF FUNCTIONALISED HUMAN HEAVY-CHAIN FERRITINS AND COMPUTATIONAL INVESTIGATION OF SELF-ASSEMBLING	61
3.1 Introduction.....	62
3.2 Materials and methods	64
3.2.1 Materials	64
3.2.2 HF _n -based protein sequence design.....	65
3.2.3 Plasmid construction for sHF _n , sHF _n -P and sHF _n -P-RGDK.....	67
3.2.4 Heat-shock transformation and plasmid sequencing	69
3.2.5 Shake flask and fermenter fermentation of different <i>E. coli</i> strains	70
3.2.6 Transmission electron microscopy (TEM) and size exclusion chromatography (SEC) analysis of bacterial lysate supernatants	72
3.2.7 Computational analysis of the impact of pro-apoptotic peptide on HF _n and sHF _n subunit property and structure	73
3.3 Results and discussion	76
3.3.1 Characterisations of plasmid of sHF _n , sHF _n -P and sHF _n -P-RGDK.....	76

3.3.2 Confirmation of plasmid transformation into <i>E. coli</i>	77
3.3.3 Expression comparison of 8 HF _n -based proteins	78
3.3.4 Expression conformation comparison of soluble expression of HF _n -based proteins	81
3.3.5 Impact of pro-apoptotic peptide on HF _n and sHF _n biochemical property and structure by computational analysis.....	83
3.4 Conclusions.....	91
3.5 References.....	92
Chapter 4.....	95
PURIFICATION PROCESS DEVELOPMENT AND STRUCTURE COMPARISON OF HUMAN HEAVY-CHAIN FERRITIN AND FUNCTIONALISED HUMAN HEAVY-CHAIN FERRITINS.....	95
4.1 Introduction.....	96
4.2 Materials and methods	98
4.2.1 Materials	98
4.2.2 Comparison of host cell proteins (HCPs) removal by heat-acid precipitation	98
4.2.3 Nucleic acid removal by Q FF IEC.....	99
4.2.4 Nucleic acid removal by HIC	99
4.2.5 Scale up purification of HF _n	100
4.2.6 Mono Q purification of functionalised HF _n s after heat-acid precipitation.....	100
4.2.7 Secondary and tertiary structural comparison.....	101
4.2.8 Nanoparticle structural comparison	101
4.3 Results and discussion	102
4.3.1 Optimisation of HCPs removal by heat-acid precipitation.....	102
4.3.2 Nucleic acid removal by Q FF IEC.....	104
4.3.3 Nucleic acid removal by HIC	107
4.3.4 Scalable HCPs and nucleic acid removal pathway of HF _n	111
4.3.5 Functionalised HF _n s purification by Mono Q IEC after heat-acid precipitation .	111

4.3.6 Functionalisation effect on HF _n secondary and tertiary structure	112
4.3.7 Functionalisation effect on HF _n nanoparticle structure.....	113
4.4 Conclusions.....	118
4.5 References.....	119
Chapter 5.....	122
THERMALLY INDUCED DOXORUBICIN LOADING OF HUMAN HEAVY-CHAIN FERRITIN AND FUNCTIONALISED HUMAN HEAVY-CHAIN FERRITIN	122
5.1 Introduction.....	123
5.2 Materials and methods	125
5.2.1 Materials	125
5.2.2 Thermally induced passive loading of DOX into HF _n and HF _n -GFLG-PAS-RGDK	125
5.2.3 Thermally induced passive loading of DOX into HF _n -PLGLAG-PAS-RGDK and HF _n -PAS.....	126
5.2.4 TEM analysis of protein/DOX.....	127
5.2.5 Stability of DOX loaded HF _n and PAS functionalised HF _n s.....	127
5.2.6 Computational study of interactions of HF _n and DOX in thermally induced drug loading.....	127
5.3 Results and Discussion	128
5.3.1 Optimisation of HF _n thermally induced passive loading to increase drug loading	128
5.3.2 Optimisation of HF _n -GFLG-PAS-RGDK thermally induced passive loading	132
5.3.3 Conformation characterisation of optimal protein/DOX	135
5.3.4 Stability study of DOX loaded proteins.....	136
5.3.5 Interactions between HF _n and DOX in thermally induced drug loading by computational analysis.....	137
5.4 Conclusion	142
5.5 References.....	143

Chapter 6.....	145
ENGINEERED HUMAN HEAVY-CHAIN FERRITINS PERFORMANCE AS AN ANTI-TUMOUR DRUG DELIVERY PLATFORM <i>IN VITRO</i> AND <i>IN VIVO</i>	145
6.1 Introduction.....	146
6.2 Materials and methods	148
6.2.1 Materials	148
6.2.2 Intracellular distribution analysis in tumour cells.....	148
6.2.3 Cytotoxicity against tumour cells	149
6.2.4 Cellular uptake assay of 4T1 cell line.....	150
6.2.5 Pharmacokinetic study	151
6.2.6 Biodistribution study in 4T1 tumour bearing mice.....	151
6.2.7 Anti-tumour study in 4T1 tumour bearing mice	152
6.2.8 Statistical analysis.....	152
6.3 Results and discussion	152
6.3.1 DOX intracellular distribution after cellular internalisation in 3 cell lines	152
6.3.2 Functionalisation impact on anti-proliferation efficacy of protein/DOX <i>in vitro</i> . 153	
6.3.3 RGDK functionalisation impact on 4T1 uptake efficiency of protein/DOX <i>in vitro</i>	155
6.3.4 PAS Functionalisation impact on pharmacokinetic profile <i>in vivo</i>	156
6.3.5 Functionalisation impact on protein biodistribution <i>in vivo</i>	158
6.3.6 Functionalisation impact on protein/DOX anti-tumour efficacy <i>in vivo</i>	160
6.4 Conclusion	163
6.5 References.....	164
Chapter 7.....	167
CONCLUSION AND FUTURE RESEARCH.....	167
7.1 Conclusions.....	168
7.2 Future research directions	170

Appendix A-- Supporting tables	171
Appendix B – Supporting figures	177
PUBLICATIONS DURING CANDIDATURE	192

LIST OF FIGURES

Figure 2.1 Hollow spherical structures of three types of ferritin.	12
Figure 2.2 Structure of mammalian ferritin subunit and assembly.	13
Figure 2.3 Schematic of ferritin drug-loading approaches.	16
Figure 2.4 Locations of reactive residues for chemical conjugation on the outer surface of HF _n subunit and nanocage.	28
Figure 2.5 Theoretical conformations of HF _n subunits and nanocages with peptide fusion at different sites.	30
Figure 2.6 Procedures of hybridisation of two types of ferritin subunits through two different means.	33
Figure 3.1 Schematic of expression cassettes of the pET30a plasmids harbouring genes for all HF _n -based proteins.	66
Figure 3.2 Schematic of <i>E. coli</i> construction.	67
Figure 3.3 Schematics of theoretical P-HF _n and sHF _n -P subunits with fully stretched linkers.	74
Figure 3.4 Procedure for computational analysis of P impact on HF _n assembling.	75
Figure 3.5 1 % agarose gel electrophoresis results of pET30a vector restriction enzyme treatment.	76
Figure 3.6 1.5 % agarose gel electrophoresis results of PCR for sHF _n , sHF _n -P and sHF _n -P-RGDK.	77
Figure 3.7 Pictures of <i>E. coli</i> BL21 (DE3) colonies grown on the agar LB plates.	78
Figure 3.8 Bacterial growth curve in batch fermenter fermentation.	78
Figure 3.9 Expression results of HF _n -based proteins except for P-HF _n	79
Figure 3.10 Expression results for standard and low temperature induction of P-HF _n	80
Figure 3.11 TEM images of bacterial lysate supernatants.	81
Figure 3.12 SEC of bacterial lysate supernatants of sHF _n -P and sHF _n -P-RGDK and reducing SDS-PAGE results of SEC fractions.	82
Figure 3.13 RMSD of structures in 5 ns MD simulation.	83
Figure 3.14 Radius of gyration of structures in 5 ns MD simulation.	84
Figure 3.15 3D structure of top 4 complexes of HF _n subunit and P obtained from HADDOCK 2.4 docking analysis.	84

Figure 3.16 3D structure of top 4 complexes of sHFn subunit and P obtained from HADDOCK 2.4 docking analysis.....	86
Figure 3.17 Surface hydrophobicity and charge comparison of top 1 complexes with HFn and sHFn subunit.	87
Figure 3.18 Comparison of sHFn3P with HFn and sHFn subunit after Gromacs MD simulation.	88
Figure 3.19 Comparison of P3HFn with HFn and sHFn subunit N-terminal after Gromacs MD simulation.....	89
Figure 3.20 Structures of HFn, sHFn, sHFn3P and P3HFn subunit after MD simulation, front view.....	90
Figure 3.21 Surface charge property of P.	90
Figure 4.1 Heat-acid precipitation of HFn and functionalised HFns.	103
Figure 4.2 12 % reducing SDS-PAGE images of peaks collected from Q FF column.	105
Figure 4.3 Q FF chromatography purification of HFn, HFn-PAS and HFn-GFLG-PAS-RGDK.	106
Figure 4.4 SDS-PAGE images of HFn and 3 PAS functionalised HFns after HIC.....	108
Figure 4.5 HIC chromatography purification for HFn and functionalised HFns.....	109
Figure 4.6 Large-scale butyl chromatogram and SDS-PAGE image.	111
Figure 4.7 Mono Q chromatogram and SDS-PAGE image.....	112
Figure 4.8 Circular dichroism spectra (A) and intrinsic fluorescence spectra of four HFn-based proteins (B).	113
Figure 4.9 Size distributions of proteins from DLS.....	114
Figure 4.10 5 % Native-PAGE analysis results of purified proteins.	114
Figure 4.11 HPSEC-MALLS chromatograms of proteins.....	115
Figure 4.12 TEM images of purified proteins.....	116
Figure 4.13 Schematic illustration of assumed protein structures.	117
Figure 5.1 Thermally induced DOX loading to HFn under different experimental conditions	130
Figure 5.2 Thermally induced DOX loading to HFn-GFLG-PAS-RGDK under different experimental conditions.	133
Figure 5.3 Size-exclusion chromatograms of HFn and functionalised HFns/DOX.	135
Figure 5.4 TEM images of protein/DOX.	136
Figure 5.5 DOX leakage over time at different conditions.....	137

Figure 5.6 3D structures of 9 complexes after 10 ns MD simulation.	138
Figure 5.7 RMSD of HF _n subunit and DOX in complexes during MD simulation	139
Figure 5.8 Short-range non-bonded interaction energy of HF _n subunit and DOX in complexes during MD simulation.....	139
Figure 5.9 TEM image of soluble HF _n -DOX aggregates (A) and schematic of conformation of soluble HF _n -DOX aggregates (B).	142
Figure 6.1 Intracellular distribution of DOX in tumour cells.	153
Figure 6.2 Proliferation inhibition effects on tumour cells.....	154
Figure 6.3 Mean DOX fluorescence intensity in cellular uptake test.	155
Figure 6.4 Plasma concentrations of protein/DOX and free DOX in SD rats of different groups.	157
Figure 6.5 Biodistribution results of HF _n and PAS functionalised HF _n s in 4T1 tumour bearing mice.....	159
Figure 6.6 <i>In vivo</i> tumour inhibitory effects on 4T1 tumour-bearing mice.	161
Figure 6.7 Schematics of different tumour cellular internalization mechanisms of DOX and protein/DOX.	162
Figure B1 Standard curve for Bradford assay.....	177
Figure B2 Standard curve for nucleic acid concentration determination.....	178
Figure B3 Standard linear curves of correlations between DOX or HF _n -based protein nanocages concentrations and optical densities.....	179
Figure B4 Size-exclusion chromatograms of 45 °C HF _n /DOX samples.....	180
Figure B5 Size-exclusion chromatograms of 50 °C HF _n /DOX samples.....	181
Figure B6 Size-exclusion chromatograms of 60 °C HF _n /DOX samples.....	182
Figure B7 Size-exclusion chromatograms of 45 °C HF _n -GFLG-PAS-RGDK/DOX samples.	183
Figure B8 Size-exclusion chromatograms of 50 °C HF _n -GFLG-PAS-RGDK/DOX samples.	184
Figure B9 Size-exclusion chromatograms of 60 °C HF _n -GFLG-PAS-RGDK/DOX samples.	185

Figure B10 2D diagrams of 9 complexes of HFn subunit with DOX after 10 ns MD simulation.
..... 190

Figure B11 Standard curve of fluorescence intensity-doxorubicin concentration in SD rat
plasma. 191

LIST OF TABLES

Table 2.1 Comparative summary of doxorubicin loading performances with different loading approaches.....	24
Table 2.2 Comparative summary of ferritin drug-loading and functionalisation over the past 20 years.	35
Table 3.1 Primer sequences for PCR.	67
Table 3.2 Reagents for PCR.....	68
Table 3.3 Interacted residues of HF _n and interaction types in top 1 complex.....	85
Table 3.4 Interacted residues of sHF _n and interaction types in top 1 complex.	86
Table 4.1 Purity and protein recovery yield (step) in heat-acid precipitation.....	103
Table 4.2 Purity, recovery yield (step) and nucleic acid removal of all peaks in Q FF chromatography.	106
Table 4.3 Purity, protein recovery yield (step) and nucleic acid removal in HIC.	109
Table 4.4 Protein particle sizes and zeta potentials measured from DLS.	114
Table 4.5 High performance size exclusion chromatography and multi-angle laser light scattering (HPSEC-MALLS) analysis result.	115
Table 5.1 Levels of variables used in the orthogonal tests for optimisation of thermally induced DOX loading to HF _n and HF _n -GFLG-PAS-RGDK.....	125
Table 5.2 Drug loading ratios (Ns), proportions of DOX loaded in nanocage and protein recovery yields in HF _n thermally induced drug loading optimisation.	130
Table 5.3 Comparison on DOX loading to HF _n in this work and previous studies.....	132
Table 5.4 Ns, proportions of DOX loaded in nanocage and protein recovery yields in HF _n -GFLG-PAS-RGDK thermally induced drug loading optimisation	134
Table 5.5 DOX loading efficiency of HF _n and functionalised HF _n s.....	135
Table 5.6 List of residues in HF _n subunit that interacts with DOX in complexes and the interaction type.....	140
Table 5.7 Five (5) residue average hydrophobicity of residues in DOX binding area in complex 1, 4, 5 and 3.....	141
Table 6.1 IC ₅₀ values of all groups.....	154
Table 6.2 Half-life in circulation of each protein/DOX in SD Rats (n = 3).....	157
Table 7.1 Experiment procedure and brief result summary of all HF _n -based proteins in this thesis.	169

Table A1 Primary sequences of 8 HF _n -based protein subunits.	171
Table A2 Sequencing results from BGI Company (China)	173

ABBREVIATIONS

In the order of first appearance in thesis.

L-chain	Light-chain
H-chain	Heavy-chain
HF _n	Human heavy-chain ferritin
sHF _n	Short human heavy-chain ferritin, E-helix truncated human heavy-chain ferritin
<i>E. coli</i>	<i>Escherichia coli</i>
PAS	Polypeptide made up of P, A and S residues
RGDK	Tetrapeptide made up of R, G, D and K residues
P	Pro-apoptotic peptide
EPR	Enhanced permeation and retention effect
MD	Molecular Dynamic
Bfr	Bacterioferritin
DPS	DNA-binding protein from starved cells
TfR I	Transferrin receptor I
DMSO	Dimethyl sulfoxide
DMF	Dimethylformamide
UV/vis	Ultraviolet light/ visible light
GdHCl	Guanidium chloride
SDS	Sodium dodecyl sulphate
TEM	Transmission Electron Microscope
PLAA	Poly-L-aspartic acid
SMCC	Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate
EDC	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide
NHS	N-Hydroxysuccinimide
PEG	Polyethylene glycol
MAL	4-Methylalloxy-3,5-dimethoxyphenethylamine
GFP	Green fluorescent protein
PDB	Protein data bank
NMR	Nuclear magnetic resonance
RMSD	Root mean square deviation

RMSF	Root mean square fluctuation
R _g	Radius of gyration
VLP	Virus-like particle
TPP	Tumour penetration peptide
MMP-2	matrix metalloproteinase-2
PAGE	Polyacrylamide gel electrophoresis
IPTG	Isopropyl β-D-thiogalactoside
PCR	Polymerase chain reaction
EDTA	Ethylenediaminetetraacetic acid
LB medium	Luria-Bertani medium
BSA	Bovine Serum Albumin
DTT	Dithiothreitol
SEC	Size-exclusion chromatography
OD	Optical density
PB	Phosphate buffer
H-bond	Hydrogen-bond
HCPs	Host cell proteins
IEC	Ion-exchange chromatography
FT	Flow through
AS	Ammonium sulphate
CD	Circular Dichroism
IF	Intrinsic Fluorescence
DLS	Dynamic Light Scattering
HPSEC-MALLS	High-Performance Size Exclusion Chromatography coupled with Multiple Angle Laser Light Scattering
HIC	Hydrophobic interaction chromatography
CV	Column volume
HPLC	High Performance Liquid Chromatography
RI	Refractive index
DOX	Doxorubicin hydrochloride
Protein/DOX	Doxorubicin loaded protein
N	Drug loading ratio i.e., the number of DOX per HF _n or functionalised HF _n nanocage

NVT	Temperature Equilibration
NPT	Pressure equilibration
PS	Penicillin-streptomycin solution
FBS	Fetal bovine serum
PBS	Phosphate-buffered saline
SD	Standard deviation
IC ₅₀	Half-maximal inhibitory concentration
T _{1/2}	Half-life in blood circulation
NRP1	Neuropilin-1

Chapter 1

INTRODUCTION

1.1 Background

Drug delivery systems have been extensively used in tumour treatment to enhance drug efficacy and reduce side effects. Ferritin, a type of protein nanocage, is widely regarded as a promising anti-tumour drug delivery platform. In mammalian cells, ferritin stores iron ions and protects cells from Fenton reactions.¹ A typical ferritin assembly includes 24 of 2 types of subunits, heavy-chain (H-chain) and light-chain (L-chain). Both subunits consist of 5 helices (helix A, B, C, D, and E) and loops and turns. H-chain subunit has a ferroxidase centre that can oxidise Fe^{2+} to Fe^{3+} , while L-chain subunit facilitates iron nucleation.²⁻⁴ Ferritin assembly has a 12 nm outer diameter and 8 nm inner diameter. Eight hydrophilic channels and 6 hydrophobic channels on ferritin shell are the entrance and exit for iron ions and protons.⁵ Because of the hollow sphere structure, ferritin assembly is a near-ideal container for drugs. Its protein nature and mammal source make it biodegradable and highly biocompatible. In addition, ferritin is economical to produce and store. Ferritin can be expressed in assembly form in *Escherichia coli* (*E. coli*), and its thermal stability and chemical stability are advantageous in preparation and storage.^{6, 7} Moreover, its nanometre scale size endows it with enhanced permeation and retention effect. Inherent affinity of human heavy-chain subunit (HFn) to human transferrin receptor 1 (TfR1) can be made use of to achieve active tumour targeting.⁸

In order to load various types of drugs into ferritin, researchers have developed various drug loading approaches, including passive loading, pH-induced and 8 M urea induced disassembly-reassembly. Small molecules can take advantage of the channels on ferritin shell and be loaded by incubation with ferritin under suitable conditions, which is called passive loading approach.⁹ In passive loading approach, ferritin nanocage remain as assembly in the whole process. Additives i.e., low concentrations of denaturants, and stressors, particularly, high pressure, have been applied to expand channels to facilitate drug loading.¹⁰⁻¹²

In contrast with passive loading approach, in pH-induced and 8 M urea-induced disassembly-reassembly approaches, ferritin dissociates into individual subunits because of the disruption of inter-subunit interactions caused by extreme pH or 8 M urea.^{13, 14} Except for the two approaches mentioned above, in terms of peptide drugs, drug loading by fusion and hybridisation of different foreign peptide inserted ferritin subunits are alternatives.¹⁵ However, due to the conformational change in the drug loading process, by-product aggregates are likely

to form and the problem of protein loss remains as a challenge. For most cases of ferritin drug loading, regardless of the drug loading approach, drug loading ratios are below 100 molecules per HF_n, which is undesirable, considering ferritin cavity volume and 4500 iron ion storage ability.¹ Overall, in loading small molecule chemotherapeutics, drug loading performance of ferritin is not always desirable, because of the low protein recovery yield and loading ratio.

Functionalisation is a common method to improve ferritin drug delivery performance through equipping ferritin with other advantageous properties. Ferritin has a short half-life in circulation and its tumour targeting ability is limited. Chemical conjugation and gene technology are 2 common means. Chemical conjugation makes use of the –SH and –NH₂ groups of residues on ferritin outer surface. Folic acid was conjugated with ferritin for tumour targeting.¹⁶ Gene technology includes mutation of ferritin and fusion of functional peptides. Researchers have mainly used mutations of residues in ferritin subunit to study residue function in ferritin assembly.¹⁷ Fusion is the most frequently used functionalisation means for ferritin because of the site-specific, stable modification and well-established technique. Subunit N-terminal, subunit C-terminal, and the flexible turn connecting helix D and E are the main available insertion sites.¹⁸ In some cases, foreign peptide fusion has affected negatively the soluble expression level and self-assembly of ferritin.¹⁹ Therefore, the study of selection of functional peptides and its effects on conformation of functionalised ferritin can guide to develop efficient protein-based drug carriers.

1.2 Aim and objectives

The overarching aim of this work is to improve human heavy-chain ferritin (HF_n) performance as an anti-tumour drug delivery platform. Functionalisation of HF_n is conducted through fusion of functional peptides, and drug loading approach is investigated in detail to improve protein yield and drug loading efficiency.

PAS functionalisation and pro-apoptotic peptide (P) functionalisation are explored in this thesis. In total, 8 HF_n-based proteins were designed. HF_n and E-helix truncated HF_n (sHF_n) are controls. Three PAS fused HF_ns are named as HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGK, respectively. PAS peptide is a 40 aa half-life extender and RGDK peptide is a tumour penetration tetrapeptide.^{20, 21} After guiding HF_n to reach tumour cells, functional peptides PAS and RGDK are supposed to be removed. GFLG (4 aa) and PLGLAG

(6 aa) are two cleavable sites responsive to tumour cell overexpressed enzymes cathepsin B and matrix metalloproteinase 2/9, respectively.^{22, 23} Doxorubicin hydrochloride (DOX) was loaded to 3 PAS functionalised HFns and control HFn in a study of thermally induced passive diffusion approach. This approach introduced high temperature as a stressor and the conditions were investigated in detail to optimise loading performance. The rest 3 proteins are an anti-tumour peptide drug, P functionalised HFn or sHFn, named P-HFn, sHFn-P and sHFn-P-RGDK. A significant impact of P on HFn assembling was found in expression and was investigated in detail. The objectives of this thesis are as follows:

1. Design functionalised HFns by fusing and express these in *E. coli*. Compare expression of HFns inserted with different functional peptides, including PAS, RGDK and P, and investigate P impact on HFn assembling by molecular characterisations and computational tools.
2. Develop purification pathways for recombinant HFn and 3 PAS functionalised HFns to remove host cell proteins (HCPs) and nucleic acids, and study impacts of functional peptides on conformation changes and consequent purification pathway changes.
3. Explore a new passive diffusion drug loading approach, thermally induced passive drug loading approach, for both HFn and 3 PAS functionalised HFns to improve drug loading ratio and reduce protein loss. Study the interactions between HFn and loaded DOX by computational methods to understand ferritin drug loading mechanisms.
4. Assess how functional peptide PAS and RGDK improve anti-tumour performance by comparing anti-tumour performances in both *in vitro* and *in vivo* tests.

The methodology is illustrated as **Figure 1.1**.

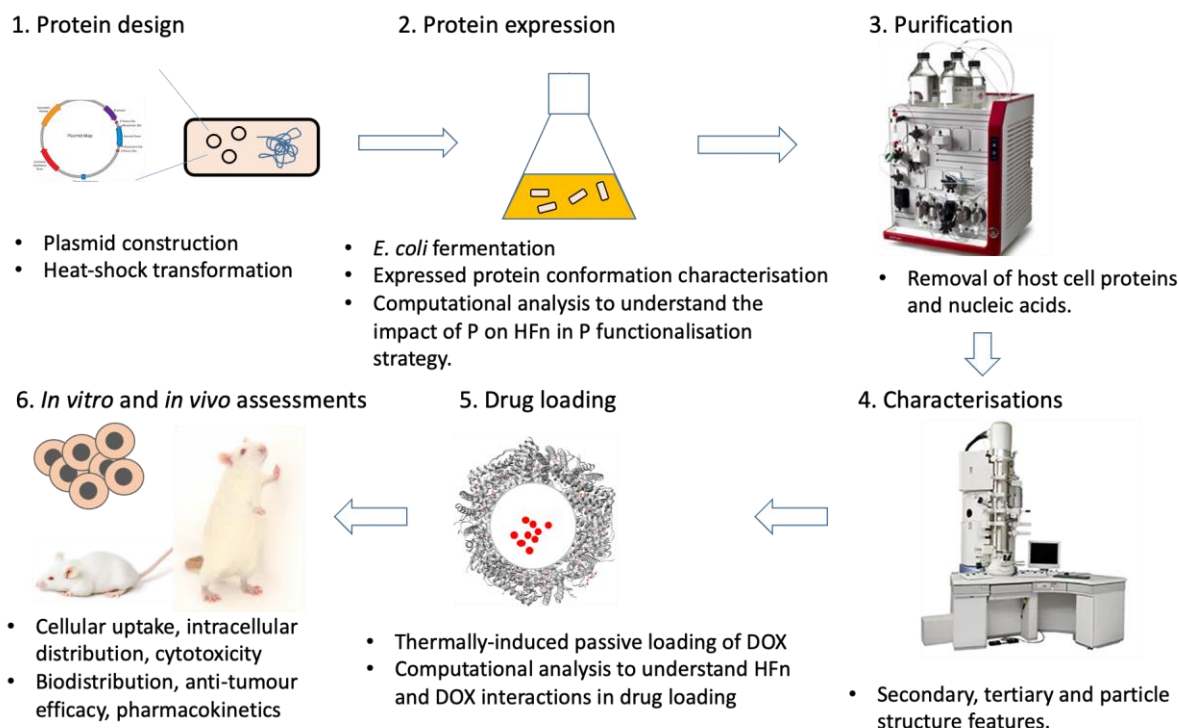


Figure 1.2 Flow chart of research methodology in this thesis.

In strategy 1, PAS functionalisation, all steps have been performed, while in strategy 2, P functionalisation, steps 1 and 2 were conducted.

1.3 Thesis outline

This thesis has 7 chapters, and the general outline is as follows:

Chapter 1: Introduction, a brief introduction of this thesis, the background information of using ferritin as a drug delivery system, thesis aim and objectives, together with thesis outline.

Chapter 2: Literature review. This chapter describes ferritin structure features and properties that make ferritin a desirable drug delivery platform. Multiple drug loading approaches and functionalisation means adopted in ferritin are then evaluated in detail. Main research problems within ferritin drug loading approaches and functionalisation methods are included in corresponding sections. Literature about computation approaches, especially molecular dynamic simulation and molecular docking for protein structure prediction and protein-ligand interaction analysis are also briefly reviewed.

Chapter 3: Molecule design, expression of functionalised HFns and computational investigation of self-assembling. In this chapter, plasmids harbouring genes encoding designed proteins were constructed and transformed into *E. coli* BL21 (DE3). Protein molecule

conformation were characterised. Computational analysis was adopted to understand the possible reasons why some of the designed HF_n-based proteins could not be expressed in soluble and self-assembly form.

Chapter 4: Purification process development and structure comparison of HF_n and functionalised HF_ns. A heat-acid precipitation followed by chromatography process was developed to obtain high purities for all 4 HF_n-based proteins. Differences in structure of assembled HF_n-based proteins have demonstrated the impact of inserted functional peptides (PAS and RGDK) on HF_n conformation and the consequent purification pathway changes.

Chapter 5: Thermally induced doxorubicin loading of HF_n and functionalised HF_ns. A thermally induced passive diffusion approach was introduced to load DOX to HF_n and PAS-functionalised HF_ns. Process factors such as heating time, temperature and buffer pH were optimised for HF_n-GFLG-PAS-RGDK and HF_n DOX loading. HF_n and PAS functionalised HF_ns drug loading performances were compared. Molecular dynamic simulation and docking software was used to simulate the interaction sites and types of HF_n and DOX in drug loading to help explain the mechanism behind drug loading and the formation of by-products.

Chapter 6: Engineered HF_ns performance as an anti-tumour drug delivery platform were tested *in vitro* and *in vivo*. *In vitro*, cellular uptake efficiency was investigated to confirm RGDK peptide function. Intracellular distribution was conducted to monitor if drug could enter nucleus to kill tumour cells. Cytotoxicity was performed to compare IC₅₀ values. *In vivo*, pharmacokinetic study determined the half-lives in blood circulation. Biodistribution study used *in vivo* imaging to prove the function of RGDK peptide. Anti-tumour test aimed to confirm the improvement of anti-tumour performance after functionalisation.

Chapter 7: Conclusions and future directions. Conclusions derived from key findings in previous chapters and future research directions for ferritin as a drug delivery platform.

Appendix A includes Supporting Tables for **Chapter 3**. **Appendix B** lists Supporting Figures for **Chapter 3, 4, 5** and **6**.

1.4 References

1. Arosio, P.; Ingrassia, R.; Cavadini, P., Ferritins: a family of molecules for iron storage, antioxidation and more. *Biochimica et Biophysica Acta* 2009, 1790 (7), 589-599.
2. Zang, J.; Chen, H.; Zhao, G.; Wang, F.; Ren, F., Ferritin cage for encapsulation and delivery of bioactive nutrients: From structure, property to applications. *Critical Reviews in Food Science and Nutrition* 2017, 57 (17), 3673-3683.
3. Arosio, P.; Elia, L.; Poli, M., Ferritin, cellular iron storage and regulation. *IUBMB Life* 2017, 69 (6), 414-422.
4. Levi, S.; Cesareni, L. A.; Cozzi, A.; Franceschinelli, F.; Albertini, A.; Arosio, P., Mechanism of ferritin iron uptaking and deletion mapping of ferroxidase center in H-chain and deletion mapping of the ferro-oxidase Site. *The Journal of Biological Chemistry* 1988, 263 (34), 18086-18092.
5. Theil, E. C.; Liu, X. S.; Tosha, T., Gated pores in the ferritin protein nanocage. *Inorganica Chimica Acta* 2008, 361 (4), 868-874.
6. Kim, S. W.; Kim, Y. H.; Lee, J., Thermal stability of human ferritin: concentration dependence and enhanced stability of an N-terminal fusion mutant. *Biochemical and Biophysical Research Communications* 2001, 289 (1), 125-129.
7. Santambrogio, P.; Levi S.; Arosio, P.; Palagi, L.; Vecchio, G.; Lawson, D. M.; Yewdall, S. J.; Artymiuk, P. J.; Harrison, P. M.; Jappelli, R.; Cesareni, G., Evidence that a salt bridge in the light chain ferritin contributes to the slight difference of physical stability between heavy and light chain. *The Journal of Biological Chemistry* 1992, 267 (20), 14077-14083.
8. Li, L.; Fang, C. J.; Ryan, J. C.; Niemi, E. C.; Lebron, J. A.; Bjorkman, P. J.; Arase, H.; Torti, F. M.; Torti, S. V.; Nakamura, M. C.; Seaman, W. E., Binding and uptake of H-ferritin are mediated by human transferrin receptor-1. *Proceedings of the National Academy of Sciences of the United States of America* 2010, 107 (8), 3505-3510.
9. Liu, X.; Wei, W.; Huang, S.; Lin, S. S.; Zhang, X.; Zhang, C.; Du, Y.; Ma, G.; Li, M.; Mann, S.; Ma, D., Bio-inspired protein-gold nanoconstruct with core-void-shell structure: beyond a chemo drug carrier. *Journal of Materials Chemistry B* 2013, 1 (25), 3136-3143.
10. Liu, X.; Jin, W.; Theil, E. C., Opening protein pores with chaotropes enhances Fe reduction and chelation of Fe from the ferritin biomineral. *Proceedings of National Academy of Sciences of the United States of America* 2003, 100 (7), 3653-3658.
11. Wang, Q.; Zhang, C.; Liu, L.; Li, Z.; Guo, F.; Li, X.; Luo, J.; Zhao, D.; Liu, Y.; Su, Z., High hydrostatic pressure encapsulation of doxorubicin in ferritin nanocages with enhanced efficiency. *Journal of Biotechnology* 2017, 254, 34-42.
12. Yang, R.; Tian, J.; Liu, Y.; Yang, Z.; Wu, D.; Zhou, Z., Thermally induced encapsulation of food nutrients into phytoferritin through the flexible channels without additives. *Journal of Agricultural and Food Chemistry* 2017, 65 (46), 9950-9955.
13. Kuruppu, A. I.; Zhang, L.; Collins, H.; Turyanska, L.; Thomas, N. R.; Bradshaw, T. D., An apoferritin-based drug delivery system for the tyrosine kinase inhibitor Gefitinib. *Advanced Healthcare Materials* 2015, 4 (18), 2816-2821.
14. Yang, Z.; Wang, X.; Diao, H.; Zhang, J.; Li, H.; Sun, H.; Guo, Z., Encapsulation of platinum anticancer drugs by apoferritin. *Chemical Communications (Camb)* 2007, (33), 3453-3455.
15. Shuvaev, V. V.; Khoshnejad, M.; Pulsipher, K. W.; Kiseleva, R. Y.; Arguiri, E.; Cheung-Lau, J. C.; LeFort, K. M.; Christofidou-Solomidou, M.; Stan, R. V.;

- Dmochowski, I. J., Spatially controlled assembly of affinity ligand and enzyme cargo enables targeting ferritin nanocarriers to caveolae. *Biomaterials* 2018, 185, 348-359.
16. Zhen, Z.; Tang, W.; Zhang, W.; Xie, J., Folic acid conjugated ferritins as photosensitizer carriers for photodynamic therapy. *Nanoscale* 2015, 7 (23), 10330-10333.
 17. A.Luzzago, G. C., Isolation of point mutations that affect the folding of the H chain of human ferritin in E.coli. *the EMBO Journal* 1989, 8 (2), 569-576.
 18. Lee, W.; Seo, J.; Kwak, S.; Park, E. J.; Na, D. H.; Kim, S.; Lee, Y. M.; Kim, I. S.; Bae, J. S., A double-chambered protein nanocage loaded with thrombin receptor Agonist Peptide (TRAP) and gamma-carboxyglutamic acid of protein C (PC-Gla) for sepsis treatment. *Adv Mater* 2015, 27 (42), 6637-6643.
 19. Uchida, M.; Flenniken, M. L.; Allen, M.; Willits, D. A.; Crowley, B. E.; Brumfield, S.; Willis, A. F.; Jackiw, L.; Jutila, M.; Young, M. J., Targeting of cancer cells with ferrimagnetic ferritin cage nanoparticles. *Journal of the American Chemical Society* 2006, 128 (51), 16626-16633.
 20. Schlapschy, M.; Binder, U.; Borger, C.; Theobald, I.; Wachinger, K.; Kisling, S.; Haller, D.; Skerra, A., PASylation: a biological alternative to PEGylation for extending the plasma half-life of pharmaceutically active proteins. *Protein Engineering Design & Selection* 2013, 26(8), 489-501.
 21. Barui, S.; Saha, S.; Mondal, G.; Haseena, S.; Chaudhuri, A., Simultaneous delivery of doxorubicin and curcumin encapsulated in liposomes of pegylated RGDK-lipopeptide to tumor vasculature. *Biomaterials* 2014, 35(5), 1643-1656.
 22. Aggarwal, N.; Sloane, B. F., Cathepsin B: multiple roles in cancer. *Proteomics Clinical Applications* 2014, 8 (5-6), 427-437.
 23. Olson, E. S.; Jiang, T.; Aguilera, T. A.; Nguyen, Q. T.; Ellies, L. G.; Scadeng, M.; Tsien, R. Y., Activatable cell penetrating peptides linked to nanoparticles as dual probes for in vivo fluorescence and MR imaging of proteases. *Proceedings of Natational Academy of Sciences of the United States of America* 2010, 107 (9), 4311-4316.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

Ferritin is produced by almost all living organisms, including archaea, bacteria, plants and animals. In mammals, it acts as a buffer against iron-deficiency and iron-overload by storing and releasing iron. It has been used as a drug nanocarrier and in medical-related areas for some decades.¹⁻⁴ Mammalian ferritin structure and properties have been well-investigated. A wide range of drugs targeting different diseases have been loaded onto ferritin, including anti-tumour chemotherapeutics, neutral drugs, antibiotics and genes. Various approaches of drug-loading into ferritin have been studied, and condition optimisation have been used to achieve desirable loading performance. As a drug delivery platform, ferritin has several advantageous properties such as uniform hollow sphere structure, high biocompatibility and low toxicity, but it cannot be perfect. Its poor drug loading performance is one of its practical bottlenecks. In addition, researchers have worked on functionalisation means to incorporate other functional molecules with it to broaden its application. Functionalisation could also impact on ferritin structure and drug delivery performance.

However, despite an increasing global interest, there has not been a substantial review summarising and evaluating ferritin as a drug nanocarrier in detail, for example, comparing and explaining drug-loading approaches and ferritin functionalisation means. Besides that, the investigation of functionalisation impact on ferritin and drug loading were mainly performed using experiments, which focus on result comparison to verify the functionalisation design aim or select the optimal drug loading condition, lacking a detailed mechanism study. Computational analysis is a powerful tool to complement experiments and help interpret experiment findings on an atomic and molecule level and has the potentials to be employed in ferritin drug-loading mechanism and functional peptide impact of ferritin study.^{5, 6}

In this chapter, 1) ferritin as an anti-tumour drug delivery platform, and 2) computational techniques that could be used in functionalisation impact and drug loading investigation are critically reviewed.

This review consists of 4 sections: 1) Introduction of ferritin basic structure and property, focusing on those paving the way for ferritin as a drug nanocarrier; 2) Contrast of different means to load drugs in detail by investigating mechanisms and comparing these in terms of the

crucial evaluation indicator of loading performance. This is because evaluation of indicators and understanding of drug-loading mechanisms guide selection of judicious loading approach; and 3) Discussion of multiple ferritin functionalisation approaches, aims, advantages and potential risks. Findings will be of immediate interest to a wide range of researchers and benefit to manufacturers for drug delivery using ferritin. 4) Brief introduction of two currently used computational techniques for prediction of structure, property, and protein-ligand interaction of ferritin, and related proteins, molecular dynamic (MD) simulation and molecular docking. These have high potentials to be used in investigation of functional peptide impact and drug loading mechanism.

2.2 Ferritin structure and properties: the basis of a promising drug nanocarrier

A growing knowledge of ferritin structure and its properties has been revealed in continuing research since it was first identified in 1937.⁷ This understanding in growth has contributed to its potential development for drug delivery because the unique structure and properties are the solid foundation of ferritin to be a drug nanocarrier.

2.2.1 Structure

Ferritin is a protein family, and it is ubiquitous in nature. Its primary sequence differs in different species but shows a highly conservative conformational structure. The whole family has been divided into three main types according to structural features: 1) classical ferritin; 2) bacterioferritin (Bfr); and 3) DNA-binding protein from starved cells (DPS).^{8, 9} These main types are shown schematically as **Figure 2.1**. Classical ferritin is dominant in nature. It differs in primary sequence, cytological location, and regulation model in different species.¹⁰ 24 subunits form a hollow octahedron with a 12 nm outer and 8 nm inner, diameter (**Figure 2.1A**). This hollow octahedron has 2-fold, 3-fold and 4-fold symmetry axes (**Figure 2.1 D-F**). In eukaryotes and some bacteria, classical ferritin is constructed by one type of subunit called H-type subunit whilst in mammals, classical ferritin is heteropolymers made up of heavy-chain subunits (H-chain subunit, 182 aa, 21 kDa) and light-chain subunits (L-chain subunit, 174 aa, 19 kDa). H-type subunit and H-chain subunit are believed to derive from a common ancestor. The ratio of H- and L-chains in mammalian ferritin varies from tissue to tissue.¹¹ Naturally produced classical ferritin has an iron-core and plays an important role in iron metabolism by storing iron hydrides in its hollow-cavity and protecting oxidative damage from Fenton

reaction caused by excessive Fe(II).¹² Bfr is found in some bacteria and archaea. Compared with classical ferritin, it has 12 heme groups between subunits in a twofold symmetric binding site (**Figure 2.1B**). The function of heme groups is not completely understood. The former two kinds of ferritin are called maxi-ferritins. This third, which consists of only 12 subunits is called mini-ferritin. It is 3 nm less in diameter than the former two (**Figure 2.1C**). It is produced during stationary phase, and its main role is to protect bacterial chromosomes. The unique and uniform nanocage of ferritin provides a natural and near-ideal container for drugs. The abundant sources and diverse types of ferritin also offer researchers a wide range of choices when using ferritin as a drug nanocarrier. Because mammalian ferritin, such as human ferritin and commercialized horse spleen ferritin, are the most used types, this review therefore primarily focuses on the structure and properties of these.

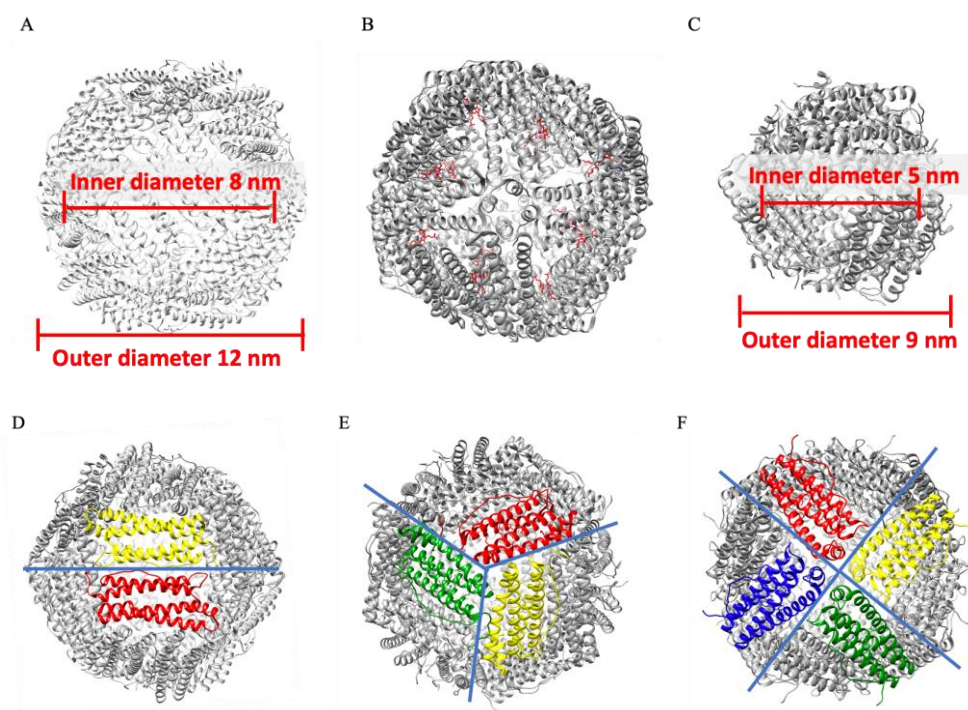


Figure 2.1 Hollow spherical structures of three types of ferritin. **A**, Classical ferritin (PDB: 2fha). **B**, bacterioferritin (Bfr) (PDB: 4am2). **C**, DNA-binding protein from starved cells (DPS) (PDB: 2iy4). **D**, 2-fold symmetry axis of classical ferritin. **E**, 3-fold symmetry axis of classical ferritin. **F**, 4-fold symmetry axis of classical ferritin. Molecules in red in **B** are the heme groups. Structures were obtained from PDB and visualized using Chimera.¹³⁻¹⁵

The H-chain and L-chain subunits of mammalian ferritins share approximately 55 % of primary structure. The conformation is similar. Both consist of five helices (helix A, B, C, D and E) together with BC loop, AB turn, CD turn and DE turn connecting helices (as is shown in **Figure 2.2B** and **C**). At each terminus of subunit, there is one short non-helical extension. Four antiparallel long helices form a bundle, and the short helix E lies at a 60° angle to bundle axis.

In H-chain subunit four helix bundle, there is a ferroxidase centre that can oxidize Fe^{2+} to Fe^{3+} (**Figure 2.2A**). L-chain subunit lacks this centre but facilitates iron nucleation. Non-covalent bonds between residues within subunits form ferritin subunit conformation and non-covalent bonds between subunits lead to the ferritin nanocage. These interactions include hydrogen bonds, salt bridges and hydrophobic interactions, and are well-documented.^{16, 17}

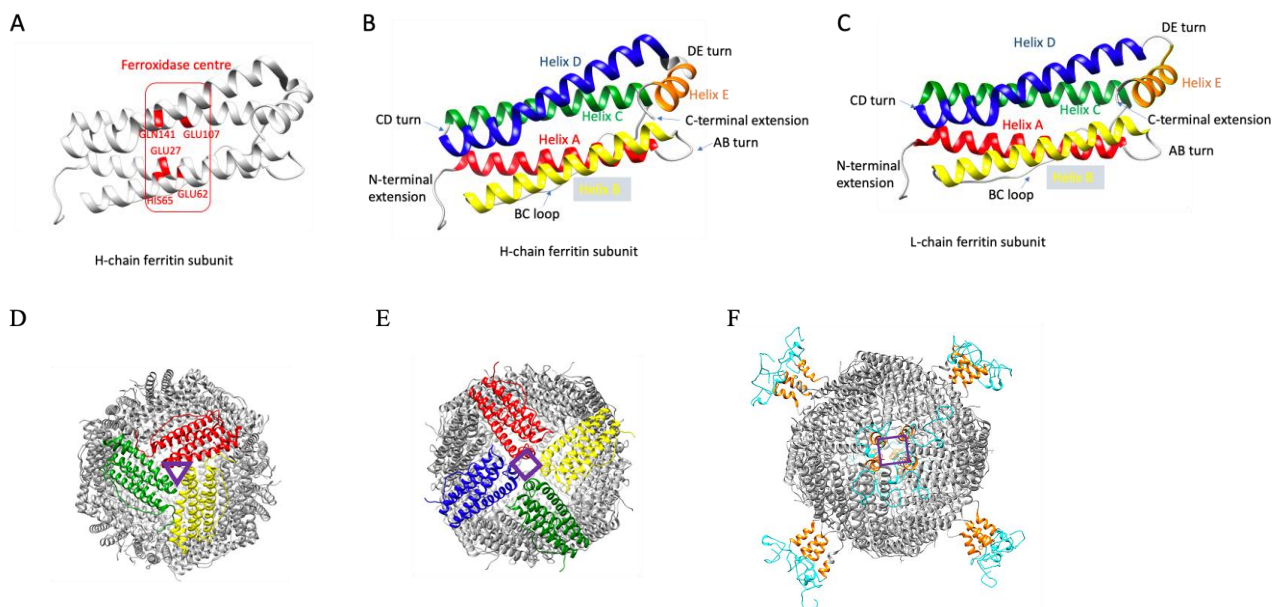


Figure 2.2 Structure of mammalian ferritin subunit and assembly.

A, Ferroxidase centre in H-chain subunit. **B**, H-chain subunit, **C**, L-chain subunit. **D**, ferritin nanocage, hydrophilic channel (in purple) formed by three subunit N-terminals of ferritin at 3-fold symmetry axis. **E**, ferritin nanocage, hydrophobic channel (purple rectangle) surrounded by 4 helix E from 4 subunits at 4-fold symmetry axis. **F**, C-terminal fused E-helix extruded ferritin (flipped ferritin), helix E is in orange and the foreign peptide cyan, hydrophobic channel is in purple. Structures A, B, C and D were originally obtained from PDB 2fha and modified and visualized using Chimera.¹⁵

When 24 subunits assemble into one ferritin cage, 14 traversing pores are formed in the ferritin shell. End of helix C and start of helix D from three subunits forms a 0.3-0.4 nm triangle-shaped channel at ferritin 3-fold symmetry axis (**Figure 2.2D**). Residues around the channels are mainly hydrophilic and are responsible for iron ions to enter the ferritin cavity. Eight hydrophilic channels can be opened and closed to respond to biological iron need.¹⁸ Each of 6 hydrophobic channels at ferritin 4-fold symmetry axis is surrounded by 4 inwardly pointing E helix (**Figure 2.2E**). It is suggested that these provide an exit pathway for protons during iron mineralisation.^{19, 20} In drug loading studies, the 14 channels have been made use of for transporting various drugs into ferritin cavity.

Ferritin structure also lays the foundation of functionalisation. Natural mammalian ferritin has all 24 N-terminal extensions outside the nanocage, which are readily available for displaying

foreign functional peptides. Despite the fact that C-terminal together with the E helix of subunit are buried inside natural ferritin nanocage, an interesting turnover phenomenon of E helix has been found by Luzzago and Cesareni after fusing a large peptide (90 aa), α -peptide of β -galactosidase to subunit C-terminus.²¹ They speculated that 24 foreign peptides together with 24 helix E were pushed out of the ferritin cavity and exposed to solvent because substrate was digested by the enzyme-fused ferritin. A theoretical schematic illustration of ferritin with extruded helix E and C-terminal inserted peptides (flipped ferritin) is shown in **Figure 2.2F**. In this flipped ferritin, the hydrophilic channels remain the same whilst hydrophobic channels expanded. This E-helix turnover has allowed subunit C-terminal available for ferritin functionalisation. In the same study, they also deleted the E helix to study its contribution to ferritin assembly and found that the truncated ferritin was still able to fold subunit and assemble into 24-mer, only with a relatively low thermal and chemical stability compared with the full-length ferritin. E helix deleted ferritin with expanded hydrophobic channels, consequently, can also be used as a drug nanocarrier.

2.2.2 Chemical and biological properties

An acknowledged attraction of ferritin is its greater stability in contrast with normal proteins. UV absorbance and circular dichroism data show that the whole nanocage of human H-chain ferritin remains stable at a temperature of 80 °C.²² Heating is therefore a good means for purification.²³ T_m values of horse-spleen ferritin and recombinant human H-chain ferritin measured in differential scanning calorimetry are also significantly high at pH 7, respectively, 93 and 77 °C.²⁴ A study of long-term stability of doxorubicin loaded ferritin revealed that at 4 °C in water, no significant change is detectable in structure and activity, even after 12 weeks.²⁵ This stability has therefore made it convenient, and largely inexpensive to manufacture and store. Ferritin is however prone to form aggregates below freezing. L-chain subunits possess better thermal and chemical stability than H-chain subunits do, in both dissociated and associated forms. This is largely due to a salt bridge formed by Lys62 and Glu107 located inside L-chain 4-helix bundle.^{26, 27}

The ferritin nanocage can resist a more extreme pH range than can normal proteins. It has a unique pH-responsivity. When solution pH is < 3.4 or > 10 , non-covalent interactions between ferritin subunits are disrupted and subunits dissociate.²⁸ With pH back to neutral, subunits can

re-associate to ferritin nanocages. This unique property is the cornerstone for the pH-induced disassembly-reassembly drug-loading approach.

Ferritin has an intrinsic targeting ability as a drug nanocarrier. Human H-chain ferritin subunit has been shown to selectively bind to human transferrin receptor I (TfR I), that is overexpressed in a range of human tumour cells and blood-brain barrier.^{23, 29-31} Mice L-chain ferritin and horse-spleen ferritin (mostly L-chain) has an innate affinity to scavenger receptor class A member 5.³²⁻³⁴ Targeting tissues or cells that do not overexpress these two receptors is also feasible through functionalisation. The plasticity and stability of ferritin structure has allowed for functional peptide insertion without disrupting ferritin conformation. Additionally, mammalian ferritin is non-toxic, has low immunogenicity, high biocompatibility and biodegradability. Similar to all drug nanocarriers, ferritin can improve drug stability *in vitro* and *in vivo*, prevent unwanted interactions, and lengthen small molecule half-life in circulation.

Overall, on the basis of its advantageous structure and properties, ferritin is a highly promising drug nanocarrier. Its uniform shell structure and stability lay the foundation of drug nanocarrier application. Its ubiquity in various species offers researchers an abundant and stable resource, together with a wide range of options in drug vector selection. Its traversing channels and pH-responsivity pave the way for drug-loading approach development. Its structural plasticity allows for incorporation of other functional molecules. Its high biocompatibility and biodegradability guarantee the safety in usage to a certain degree. As a result, researchers took advantages of all these characteristics and developed diverse drug loading approaches and functionalisation means. A great number of drugs have been loaded to different mammalian ferritins in research.

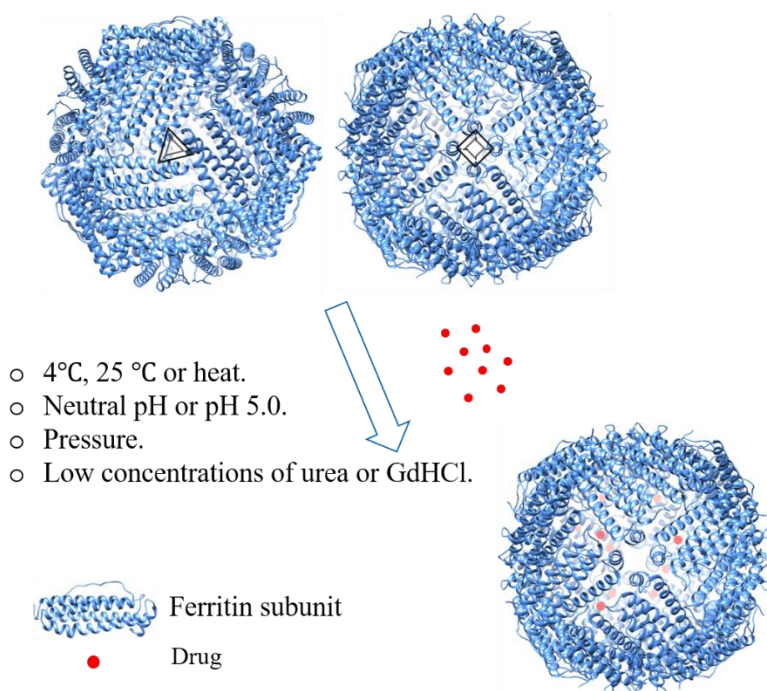
2.3 Ferritin drug-loading approaches and mechanisms

When ferritin was initially used as a drug nanocarrier, researchers drew on its natural function, storing metal ion Fe. This is the reason why various transition metal ions and metal-ion containing molecules were encapsulated into the ferritin cavity.³⁵ This cavity was also explored as a size-constrained synthesis vessel for organic and inorganic nanoparticles.³⁶ As more detailed information about ferritin structure and properties was accumulated, especially the channels in the shell, pH-responsivity, satisfactory safety and stability, various means of drug-

loading were established and a wider range of drugs, and even nutrients, have been investigated.^{37, 38}

Theoretically, drugs can be encapsulated in ferritin or tightly bound to ferritin outer surface. However, because of the ability of preventing drug degradation, drug encapsulation is preferred, and the following drug-loading approaches focus on drug encapsulation. Drug-loading comprises two steps: the entering of drugs into cavity and the remaining of drugs inside. Present drug-loading approaches use two distinct strategies in the first step: 1) Passive-loading, in which drugs diffuse inside cavity through ferritin channels; 2) Disassembly-reassembly, which opens ferritin cage and restores the structure after encapsulating drugs (**Figure 2.3**).

A. Passive diffusion



B. Disassembly-reassembly

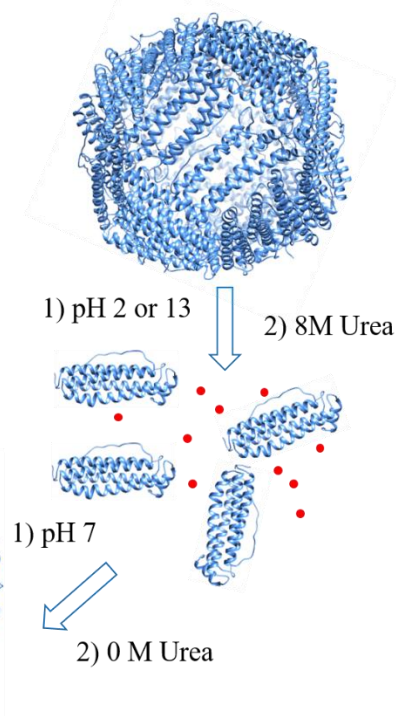


Figure 2.3 Schematic of ferritin drug-loading approaches.

A, passive-loading through hydrophilic or hydrophobic channels. **B**, disassembly-reassembly, 1) pH-induced, 2) 8 M Urea-induced. Ferritin structures were obtained from PDB and visualized using Chimera.¹⁵

In theory, drugs remain encapsulated after entering cavity because of physical entrapment and/or chemical interactions with ferritin inner surface. Physical entrapment is basically a mixing equilibrium of ferritin and drugs. External factors such as mixing time, mixing speed, temperature, initial ferritin and drug molar ratio play an important role in drug-loading ratio (number of drugs per ferritin nanocage). Property factors including drug sizes compared with ferritin channels and ferritin nanocage integrity after drug encapsulation impact the stability of

drug loaded ferritin (ferritin/drug). Chemical interactions mainly include hydrophobic interactions, electrostatic interactions, hydrogen bonds and covalent bonds. In chemical interactions, all factors that impact interaction equilibrium matter for loading ratio whilst strength of interactions is positively related to ferritin/drug stability. For example, external factors, molar ratio and mixing time impact interaction equilibrium. Properties such as, charges and hydrophobicity of drugs, exposed interaction sites of ferritin under the conditions of drug-loading, and ferritin nanocage integrity after drug-loading, will decide the drug-loading ratio and ferritin/drug stability.

Although ferritin is robust, conditions such as non-neutral pH, additives and increased temperature used in drug-loading approaches can cause different degrees of alterations to ferritin exposed interaction sites, channel areas, and its nanocage integrity after drug-loading. In addition, due to the structural alterations and possible drug-ferritin cross-linking in drug-loading, aggregation of ferritin or ferritin-drug can happen and cause precipitation. This can lead to problems such as, drugs loaded in aggregates and ferritin loss. Loading ratio is related to dosage in practice. Drug loaded in aggregates, protein loss and low ferritin/drug stability increase preparation and storage costs. Additionally, these impact ferritin/drug efficacy *in vitro* and *in vivo*. Therefore, drug-loading ratio, protein recovery, aggregation level and ferritin/drug stability are crucial indicators in evaluating the performance of drug-loading approaches. Apart from the external and property factors mentioned above, ferritin modification for functionalisation can also cause changes in drug loading performance but we this will be discussed in the functionalisation section. In the following, different types of drug-loading approaches are introduced, analyses of the mechanisms of drug entrance and remaining are described, the drug-loading performance indicators of each approach are compared, and the means to enhance performance are discussed.

2.3.1 Passive-loading

Passive-loading is completed by co-incubating drugs and ferritin. This process takes advantage of the 14 channels on the ferritin shell to allow drug entrance. The size of drugs applicable are small because of the meaningfully small channel size of 0.3-0.5 nm diameter. Theoretically at least, hydrophobic drugs such as tyrosine kinase inhibitor Gefitinib diffuse through 6 hydrophobic channels whilst hydrophilic drugs such as metal ions and doxorubicin

hydrochloride pass through 8 hydrophilic channels.³⁹ Drugs remain encapsulated primarily due to chemical interactions.

Initial passive-loading means involves zero chemical or physical stress on ferritin structure. Incubation time ranges from 1 h to overnight (> 12 h) for mixtures to reach equilibrium. Normally temperature is 4 °C and pH is neutral for the sake of stabilities of drugs and ferritin. For drugs that are soluble in aqueous solution, usual buffers such as phosphate buffer are chosen, but for hydrophobic drugs, co-solvents such as DMSO and DMF are added to increase drug solubility. Amongst selected drugs, metal ions and those bearing positive net charges, tend to have greater loading potential. This is because ferritin inner surface bears negative charges at neutral pH. For example, 145 AgNO₃ and 104 AuCl₃ were reportedly trapped in one ferritin cavity, this is significantly greater compared with ~ 12 molecules per cavity with Gefitinib.⁴⁰ In passive-loading cases, most of the drugs are not physically trapped inside ferritin cavity because of the small sizes of drugs, but there are exceptions. Reactive precursors can aggregate into large molecules in ferritin cavity to achieve physical entrapment. This also results in an improved drug loading-ratio and ferritin/drug stability due to a reduced possibility of leakage via channels during loading and storage. Fan et al. made use of the oxidation of Fe ions to form a 4.7 nm iron core inside recombinant human H-chain ferritin.⁴¹ Copper and platinum nanoparticle have also been encapsulated in ferritin via a facile mixing procedure.⁴²

43

Although the ferritin nanocage is widely acknowledged as being highly stable, it does not mean it remains rigid. Recently, ferritin channels have been demonstrated to be flexible. Phosphotungstic acid, a molecule with a slightly greater size than channels, can be encapsulated via stress-free diffusion approach.⁴⁴ Additionally, stressors that can selectively expanding channels without disrupting the ferritin nanocage have been found to make a measurable difference in loading performance. Investigated stressors include acidic pH, low concentration denaturants, and an increase of either temperature or pressure. Acidic acetic buffer (pH 5.0) was speculated to swell ferritin channels and enhance ferritin electrostatic interaction with low pKa drug.⁴⁵ In terms of denaturants, Liu et al. and Yang et al., concluded that urea, guanidine (10 mM to 1 M), and surfactant Triton X-100 (1 % and 10 %) could disrupt the hydrogen bonds around hydrophilic channels.^{22, 46} Hydrophobic channels have also been reported to be expanded by 20 mM urea.⁴⁷ Following the discovery of the different effects of temperature on

the global structure of ferritin cage and channel areas, a thermally induced drug-loading approach was introduced. Yang et al. heated soybean ferritin with a nutrient for 1 h at 60 °C and successfully loaded an average 10.5 molecules per ferritin.⁴⁸ Within a certain pressure range, the ferritin nanocage can be maintained whilst the more sensitive channel area might undergo conformational change. For example, Wang et al. have shown that with 500 MPa pressure at pH 5.5, 32 doxorubicin can be loaded into each ferritin cavity, and that the ferritin recovery percentage reaches almost 100 % in the presence of 20 mM L-arginine.⁴⁹

Passive-loading is advantageously a most ready means that presents least potential harm to the ferritin cage when compared with other means. It, however, applies only to small molecules, and it prefers reactive precursors, metal ions and positively charged molecules. Usually, a satisfactory protein recovery and low level of aggregation are obtained because no, or minor, structural alterations occur, but the stability of ferritin/drug in storage tends to be undesirable.

2.3.2 pH-induced disassembly-reassembly

pH-induced disassembly-reassembly is the most commonly used means. With this approach, drugs of either small or large size, or nanoparticles smaller than the ferritin cavity, can be loaded. It is centred on the special pH-responsivity of ferritin nanocage. Very low or high pH induces significant subunit aggregation and causes irreversible damage to ferritin. In consequence, the normally used pH range is 2 - 7.5 or 13 - 7. Acidic pH-induced approach has been investigated and used more frequently than the alkaline counterpart. A practical prerequisite for this approach is that the applicable drugs must be stable in these pH ranges. Both physical entrapment and chemical interactions lead to drug remaining in this approach. Drug property affects loading performance. Metal containing drugs for example, such as cisplatin, 56 molecules were capsulated per cage, whilst for doxorubicin, 28.3 molecules were trapped in one ferritin in similar pH ranges.^{40, 50}

pH-induced disassembly-reassembly had been regarded as a reversible process for some time, but in 2011, Kim et al. refuted this. This is because they found two holes on equine spleen ferritin after reassembly.²⁸ This structural damage in pH-induced disassembly-reassembly process is likely to cause further aggregation and precipitation. In two separate trials, ferritin recovery was reported as 25 % and 20 %.^{49, 51} Researchers, therefore, slowed down the pH adjustment course by applying 2-step adjustment back to neutral and increased ferritin recovery

by more than 3 times in the case of loading doxorubicin into horse-spleen ferritin.⁵² Except for the main pH-induced disassembly-assembly cause, there are other factors facilitating aggregation and precipitation. For example, phosphoric acid, OH⁻ and Cl⁻ anions have been shown to facilitate aggregation.⁵³ Using glycine-actate and tris-base buffer instead of phosphoric acid-acetate and glycine-HCl buffer was found to result in greater ferritin recovery. Additionally, drugs that interact with ferritin as a cross-linker contribute to aggregation and precipitation. Understanding the interaction ways can help reduce corresponding aggregation and precipitation. Ruozi et al. found the mixture of ferritin and streptomycin sulphate (pKa is *ca.* 8) became turbid when pH was around ferritin pI. This is because every streptomycin possesses two positive charges in that condition.⁵⁴ To mitigate protein loss, they increased the final pH to 8 to remove one positive charge on streptomycin molecule and found the precipitates could be partially recovered into soluble forms, and significantly improved protein recovery to 70 %. With alkaline disassociation-re-association, Pontillo et al. argued that it did not cause holes on horse-spleen ferritin based on their UV/vis and circular dichroism results, but it still had protein loss.⁵⁵

To conclude, the pH-induced disassembly-reassembly remains an appropriate choice to load larger molecules despite protein loss and aggregation. The approach is facile and does not require special equipment or reagents. Adopting stepwise pH changes and judiciously choosing buffer and adjusting other conditions in pH drug-loading procedure aid improved loading.

2.3.3 8 M Urea-induced disassembly-reassembly

High concentration of chaotropic agents (Urea, Guanidinium chloride) have been shown to disrupt secondary forces within proteins, and consequently have been adopted to denature a wide range of proteins. Ferritin was believed to be resistant to common denaturing approaches because of its robust and stable structure. In last century, ferritin stability against urea, guanidinium chloride and sodium dodecyl sulphate was explored, mostly with native ferritin such as horse-spleen ferritin, in which the majority of subunits are L-chain. Sedimentation-velocity demonstrated horse-spleen apoferritin retained sedimentation-coefficient after 24 h incubation with 8 M urea at room temperature, and 48 h incubation with 1 % SDS (w/v) at 37 °C.⁵⁶ Another ultracentrifugation study using guanidinium chloride showed native horse-spleen apoferritin existed in aggregated form after 4 h incubation with 7 M guanidinium chloride pH 7.5, and that 48 h incubation induced formation of an oligomer intermediate-like

structure.⁵⁷ Consequently, a combination of acidic pH and high concentration of denaturant was employed to ensure complete ferritin dissociation.⁵⁸ However, recently, studies have reported 8 M urea solution to successfully dissociate recombinant ferritin. The process is relatively facile. Ferritin is pre-incubated in 8 M urea solution for 0.5 h, mixed with the drug, and treated by dialysis against decreasing concentration gradient of urea buffer to slowly remove urea. Two studies employing human H-chain ferritin and a third using *Pyrococcus furiosus* ferritin were conducted by the same researchers.^{23, 29, 59} The *P. furiosus* ferritin study reported a record high doxorubicin loading-ratio (404 doxorubicin per ferritin cavity). The reason for this high loading-ratio has not been thoroughly investigated, but it is believed to be due to the special nature of archaea ferritin. In 2016, Lei et al. directly characterised dissociated ferritin after 8 M urea treatment under transmission electron microscopy (TEM).⁶⁰ A number of intact ferritin spheres were observed, but most of H-chain ferritin disassembled into smaller, fragment-like structures.

These contrasting findings obtained from earlier and recent studies can be reliably ascribed to different characterisation approaches and incubation conditions, together with different ferritin types and sources. Ferritin subunits are likely to undergo dissociation, and then gradually and randomly, aggregate during treatment. TEM directly visualizes the structure after 8 M urea treatment. However, ultracentrifuge and sedimentation-velocity can measure only density and gross size. Although the sequence of wild-type and recombinant ferritin can be exactly the same, the relatively poor folding ability of *E. coli* can make a difference in protein chemical stability. However, structural changes and mechanisms during this process have not been fully elucidated, and therefore need further investigation.

One prerequisite for this approach is that the drugs used should be urea resistant. Protein recovery is relatively low compared with passive-loading. This is primarily because of the involvement of disassembly-reassembly. Despite the fact that 8 M urea approach is not as thoroughly investigated as passive diffusion and pH-induced disassembly-reassembly, it remains a reasonable alternative for drugs that cannot stand harsh pH conditions, and that are not small enough to directly penetrate ferritin channels.

2.3.4 Alternative drug-loading approaches

Alternative loading approaches have been explored on the basis of drug-ferritin binding strategies different from the above approaches. Emulsification, a common way to mix immiscible bioactive compounds with encapsulating materials, was utilized to combine hydrophobic ferroptosis activator and hydrophobic rapamycin with iron core containing ferritin into a 78 nm particle.⁶¹ Strong coordination interactions helped radio-sensitizer Bi₂S₃ nanocrystals conjugate to the ferritin outer surface.⁶²

In some cases, drugs are connected to ferritin through covalent bonds. As a result, binding sites of drugs tend to be predictable and precise. 24 tumour targeting pro-apoptotic peptide (CGKRRK (KLAKLAK)₂) and Glucagon-like peptide-1 were loaded to ferritin by N-terminal fusion, and demonstrated that fusion did not affect ferritin nanocage self-assembly.^{63, 64} Kang et al. introduced a cysteine to ferritin subunit C-terminal and made use of it to conjugate with a fluorescent label, by forming a di-sulphate bond.⁶⁵ An enzyme drug superoxide dismutase was also encapsulated through the same way, and in the same study, the enzyme was successfully loaded through conjugation with the outer surface of horse-spleen ferritin.⁶⁶

Theoretically, because of the stronger bonds, products from these loading approaches tend to have a comparatively greater ferritin/drug stability in storage. Because of the fact that drug release inside target cells is all done by lysosome enzyme digestion of the ferritin nanocage, drug release profile will not be affected by drug-loading approaches.

2.3.5 Selection and comparison of drug-loading approaches

Drug-loading approaches are facile, effective and suitable for a range of drugs according to drug properties. In selection of drug-loading approaches, the prerequisites for each approach, such as drug size and pH or 8 M urea resistance should be taken into consideration. For example, small molecules (normally < 600 Da) can enter cavity through ferritin channels, therefore, passive-loading and disassembly-reassembly approaches both apply. Larger molecules are however, blocked and channel expanding measures or disassembly-reassembly approaches need to be adopted. Because proteins cannot stand harsh pH or 8 M urea conditions, a special *Archaeoglobus fulgidus* ferritin that can disassemble-reassemble at physiological condition (pH 7) is used as the nanocarrier.⁶⁷ It is because of the range of drug-loading approaches and

diverse ferritin types that it is always likely that, at least one means can be found for a particular payload. For drugs that theoretically can be loaded through more than one approach, selection depends on the 4 evaluation indicators, namely, loading-ratio, protein recovery, aggregation and ferritin/drug stability. Doxorubicin is the most commonly loaded drug and has been loaded in ferritin through passive diffusion, pH-induced disassembly-reassembly, and 8 M urea-induced disassembly-reassembly in previous studies, shown in **Table 2.1**. It is taken as an example to compare drug loading approaches in detail.

Table 2.1 Comparative summary of doxorubicin loading performances with different loading approaches.

Ferritin type	Drug-loading approach	Initial concentrations	Loading ratio	Ferritin recovery percentage	Aggregation	Ferritin /drug stability	Reference
Recombinant human H-chain ferritin	pH-induced, pH 2-7.5	Ferritin, 0.5 mg mL ⁻¹ Drug, 0.11 mg mL ⁻¹	28.3	/	/	40 % drug loss in 72 h in PBS at 4 °C	50
Horse-spleen ferritin	pH-induced, stepwise pH adjustment, pH 2.5-4-7.4	Ferritin, 2.22 mg mL ⁻¹ Drug, 0.85 mg mL ⁻¹	24	55 ± 7	yes	20 % drug loss in 72 h in PBS at 4 °C	52
Equine-spleen ferritin	pH-induced, pH 2-7	Ferritin, 3.12 mg mL ⁻¹	22 ± 1	25	/	/	51
Recombinant human H-chain ferritin	8 M urea-induced	Ferritin, 1 mg mL ⁻¹ Drug, 1 mg mL ⁻¹	32.5	64.8	/	< 15 % of drug leaked in PBS in 60 h at 37 °C	60
Recombinant human H-chain ferritin	8 M urea-induced	Ferritin, 1 mg mL ⁻¹ Drug, 1 mg mL ⁻¹	33	/	/	< 10 % of drug leaked in mouse serum in 60 h at 37 °C	29
Recombinant human H-chain ferritin	High hydrostatic pressure (500 MPa) passive-loading	Ferritin, 1.5 mg mL ⁻¹ Drug, 1 mg mL ⁻¹	32 ± 2	100	no	< 40 % of drug leaked in PBS in 60 h at 37 °C	49

‘/’ means no data.

Doxorubicin is hydrophobic, has hydrogen bond donors and acceptors, and bears positive charge when $\text{pH} < 8.3$. It can therefore bind to ferritin through non-covalent interactions such as hydrogen bonds, ionic bonds and hydrophobic interactions. Doxorubicin is smaller than ferritin channel, chemical interactions will be dominant in affecting loading-ratio and ferritin/doxorubicin stability. Additionally, in all cases summarised in **Table 2.1**, excessive amounts of doxorubicin were used, so it can be reasonably assumed that exposed interaction sites of ferritin were saturated and mixing equilibrium for physical entrapment was achieved. In this way, ferritin exposed interaction sites in drug-loading process and nanocage integrity after drug encapsulation would mainly account for the loading performance differences.

As is seen in **Table 2.1**, high-pressure passive-loading approach has the highest protein recovery, suggesting it poses the least harm to ferritin structure. This is because subunit dissociation exposes more ferritin interface areas and has a greater chance to give rise to unwanted protein/drug-protein/drug or protein-protein interactions. In **Table 2.1**, loading-ratios of 8 M urea approaches and high-pressure passive-loading approach using recombinant human H-chain ferritin are similar. pH-induced approaches have lower loading-ratios. The 8 M urea approach has the greatest ferritin/doxorubicin stability, whilst the pH-induced approach has the least. The differences in two disassembly-reassembly approaches results suggest that numbers and locations of exposed interaction sites in drug loading, and ferritin structural integrity after encapsulation are different. In addition, solution pH differences in these two approaches can impact ferritin-doxorubicin interactions. Overall, 8 M urea approach seems to be milder than pH-induced one in terms of ferritin structure damage.

2.3.6 Improvement of drug loading performance

An ideal drug loading performance for ferritin is the achievement of all 4 satisfying indicators simultaneously. However, it is difficult to achieve in practice. Aggregation and ferritin loss data were not reported in many drug-loading studies, and a number of reported protein recoveries were unsatisfactory. Loading-ratios in most cases so far have been < 100 drug molecules per ferritin cavity. This is actually undesirable considering the cavity volume of ferritin and its iron-storage capacity (up to 4500 iron ions).⁶⁸ This is probably because of the limited number of interaction sites in the ferritin cavity, and weak interactions between ferritin and drugs.

To enhance drug loading performance, researchers have investigated different means based on improvements of physical entrapment and chemical interactions between drugs and ferritin. Condition optimization in drug loading process is one practical means for both chemical interaction dependent and physical entrapment dependent drugs. For example, stressor investigation in passive-loading, buffer selection and pH adjustment course change in pH-induced disassembly-reassembly mentioned above can help improve loading ratio and protein recovery. Prior to that, for chemical interaction dependent drugs, understanding the interactions between drugs and ferritin can be a good option to figure out the significant conditions and guide condition level selection. Computational tools, docking and molecular dynamic simulation, which are used for receptor-ligand complex analysis can be helpful. Experimental techniques such as spectroscopic characterisations may detect interaction sites and kinetics. Addition of helper reagents is another means to enhance loading performance. Through co-encapsulation with AuCl_4 and reduction agent NaBH_4 that formed an Au nanoparticle inside, 5-Fu, a 105 Da molecule has no detectable leakage in dialysis for 12 h.⁶⁹ Complexation of drugs with transition metal ions before encapsulation can improve the binding with ferritin. Doxorubicin complexed with Cu(II), Mn(II), Zn(II) and Fe(III) have been shown to have greater loading.⁷⁰ Incorporation of poly-L-aspartic acid (PLAA) has been shown to enhance electrostatic interaction and facilitate daunomycin encapsulation.^{45, 71}

In summary, various ferritin drug-loading approaches have been extensively investigated. Different approaches have particular prerequisites and are applied to different types of drugs. To select a suitable drug-loading approach, loading-ratio, aggregation level, protein recovery and ferritin/drug stability need to be considered. Understanding drug properties, ferritin structural changes and ferritin-drug interactions in drug-loading can judiciously guide selection of an appropriate drug-loading approach.

2.4 Ferritin functionalisation

Although ferritin possesses many desirable characteristics as a drug nanocarrier, it needs to be functionalised for a broader application and improvement in treatment outcomes. The two most investigated functions are half-life extension and targeting ability. A 2-3 h half-life in circulation of human H-chain ferritin is unsatisfactory.⁷² Passive targeting ability and the limited types of targeted receptors of ferritin cannot always achieve satisfactory safety.

Widely employed ferritin functionalisation approaches include 3 conventional means that are used for all proteins, namely, 1) chemical conjugation, 2) mutation and 3) fusion. These mainly introduce changes through covalent bond connections. Non-covalent interactions in hybridization, or directly adding another layer of shell outside ferritin cage, are examples of novel ways in which to combine ferritin with other moieties.⁷³ In parallel with drug-loading practices, different approaches apply to different alterations.

The addition of functional moiety and any alterations of ferritin can result in potential structural and property changes, affecting drug loading performance and treatment outcome. Most important in ferritin functionalisation is a judicious design of the functionalized molecule in order to reduce potential damage to ferritin structure and its stability, and to achieve the addition of intended functions.

2.4.1 Chemical conjugation

Chemical conjugation usually takes place after drug-loading. It is, therefore, only affect ferritin/drug stability in storage. Free primary or secondary amide groups and –SH groups in residue side chains on ferritin surface are the two main reaction sites. **Figure 2.4** shows the potential sites for chemical modification on human H-chain ferritin subunit and nanocage. Compared with the two free –SH groups from Cys residues on surface per ferritin subunit, **Figure 2.4A**, there are more primary and secondary amide groups available, **Figure 2.4C**. –SH modification is therefore more site-specific and has relatively low modification ratio. For primary or secondary amide groups, cross-linkers such as succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC), ethylene dichloride (EDC), and N-hydroxysuccinimide (NHS) are commonly in use. Stable amide bonds are generated. Targeting moieties of small molecules including folic acid and hyaluronic acid, and large compound such as antibodies and polyethylene glycol (PEG)-RGD, have been successfully conjugated to ferritin surface.^{71, 74-77} For –SH group conjugation, prior to the reaction between –SH groups and cross-linkers with 4-methylalloxy-3,5-dimethoxyphenethylamine (MAL) reactive group, amide groups that can compete with –SH groups are blocked using N-ethylmaleimide. In this way, 3 antibodies have been modified on one ferritin to target melanoma and 5 kDa PEG-MAL has been directly conjugated to increase the half-life in circulation.^{78, 79}

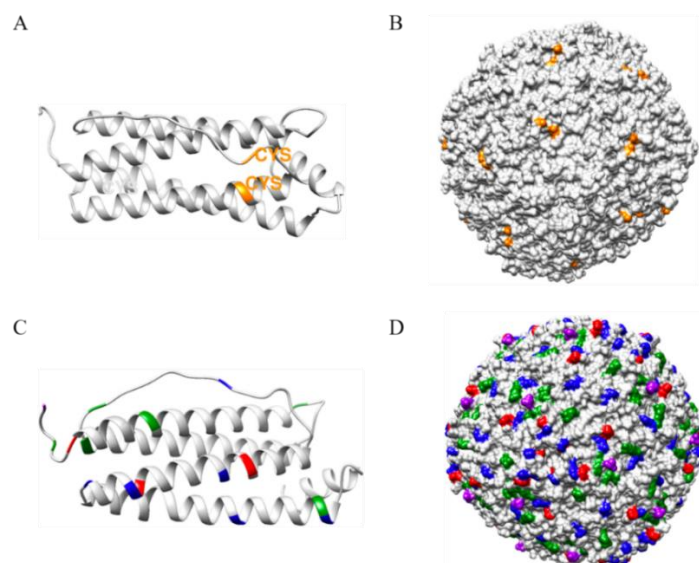


Figure 2.4 Locations of reactive residues for chemical conjugation on the outer surface of HFn subunit and nanocage.

A, two 2 Cys residues in one subunit, Cys is in orange. **B**, Cys residues on HFn assembly outer surface. **C**, 6 Lys (blue), 6 Arg (green), 3 His (red) and subunit N-terminal residues (purple) in HFn subunit. **D**, all residues containing primary or secondary amide groups on HFn assembly outer surface.

Overall, chemical conjugation is a facile functionalisation means that only impact ferritin/drug stability in storage. Equipping ferritin with multiple functions concomitantly can be achieved but success is not guaranteed. In addition, chemical conjugation may hamper ferritin's inherent abilities such as binding to TfR1, because of the blocking of ferritin outer surface.

2.4.2 Gene engineering technology

A more extensively adopted functionalisation means makes use of gene engineering technology. This uses mutation or fusion to alter ferritin primary sequence. Compared with chemical modification, this can be used to ensure a well-structured and constant pattern of intended alteration. A practicable sequence design of the modified ferritin is necessary. The length, sequence and structure of exogenous peptides and the location of alterations will affect ferritin spherical structure, stability, drug loading performance and expression in the host.

Mutation, especially site-specific mutation and/or depletion, has been used to study ferritin residue functions of iron storage and initial structural formation.^{20, 21, 80, 81} Non-covalent interactions in the ferritin interface are crucial for assembly-stability. Consequently, mutations at the ferritin surface are more likely to maintain ferritin integrity whilst mutations at the interface can impact subtle changes in self-assembly behaviour. Some ferritin mutants have

been designed for special properties. An interesting mutation of bacterial ferritin is that designed by Turrís et al. that allows it to keep its reversible disassembly-reassembly at neutral pH, but to possess binding ability with human transferrin receptor.⁸² These were achieved by substituting 12 amino acid residues on bacterial ferritin subunit for 12 residues from human ferritin subunit BC loop that are crucial for receptor binding and not essential in assembly. These researchers dissociated the humanized archaeal ferritin to dimer in low ionic strength buffer and then added MgCl₂ and cytochrome C to achieve drug-loading.⁸³ Wang et al. designed a copper-assisted ferritin mutant with strengthened stability and improved doxorubicin loading-ratio.⁸⁴ These authors introduced an artificial metal ion binding pocket for copper binding at ferritin 2-fold interface. The coordination force between copper and mutant interface residue side chains was demonstrated to be stronger than that of the original non-covalent interactions in ferritin.

Fusion accounts for the majority of genetic modifications, and in most fusions of the subunit, N-terminal, C-terminal and the flexible DE turn are selected to insert various exogenous peptides. Normally, short (< 100 aa) and non-structured peptides are more frequently adopted. This is because they pose a lower risk of adversely impacting assembly stability. Flexible linker rich in G and S are often introduced to reduce steric hindrance between foreign peptide and ferritin.

N terminal and DE turn are naturally displayed outside ferritin nanocage. Because of steric hindrance, N terminus can accommodate larger peptides than DE turn region and is more commonly used. **Figures 2.5A** and **2.5B** illustrate the N-terminal fused and DE turn inserted ferritin subunits and assemblies. Foreign peptides used are often targeting moieties or half-life extenders. IL-4R targeting peptide were fused onto DE turn to enhance localized drug delivery.⁸⁵ Falvo et al. introduced two peptides containing repetitive P, A, and S residues (40 aa and 75 aa), which are thought to be a genetically encodable substitute of PEG chain, to subunit N-terminal to prolong half-life.⁸⁶ Although in theory, PAS-rich peptide should be longer than 200 aa to be able to lengthen half-life, the well-organized and repetitive presence on ferritin surface compensated for its short-length. They also found that doxorubicin loading ratios of two PAS fused ferritins through pH-induced disassembly-reassembly approach were over 80, which is significantly greater than ferritin. Albumin binding domain fused human H-chain ferritin (HF_n) also had greater drug-loading-ratio and better ferritin/drug stability than

HF_n in pH-disassembly-reassembly approach.⁷² These foreign peptides might have helped capture drugs around it during drug-loading and blocked ferritin channels in storage.

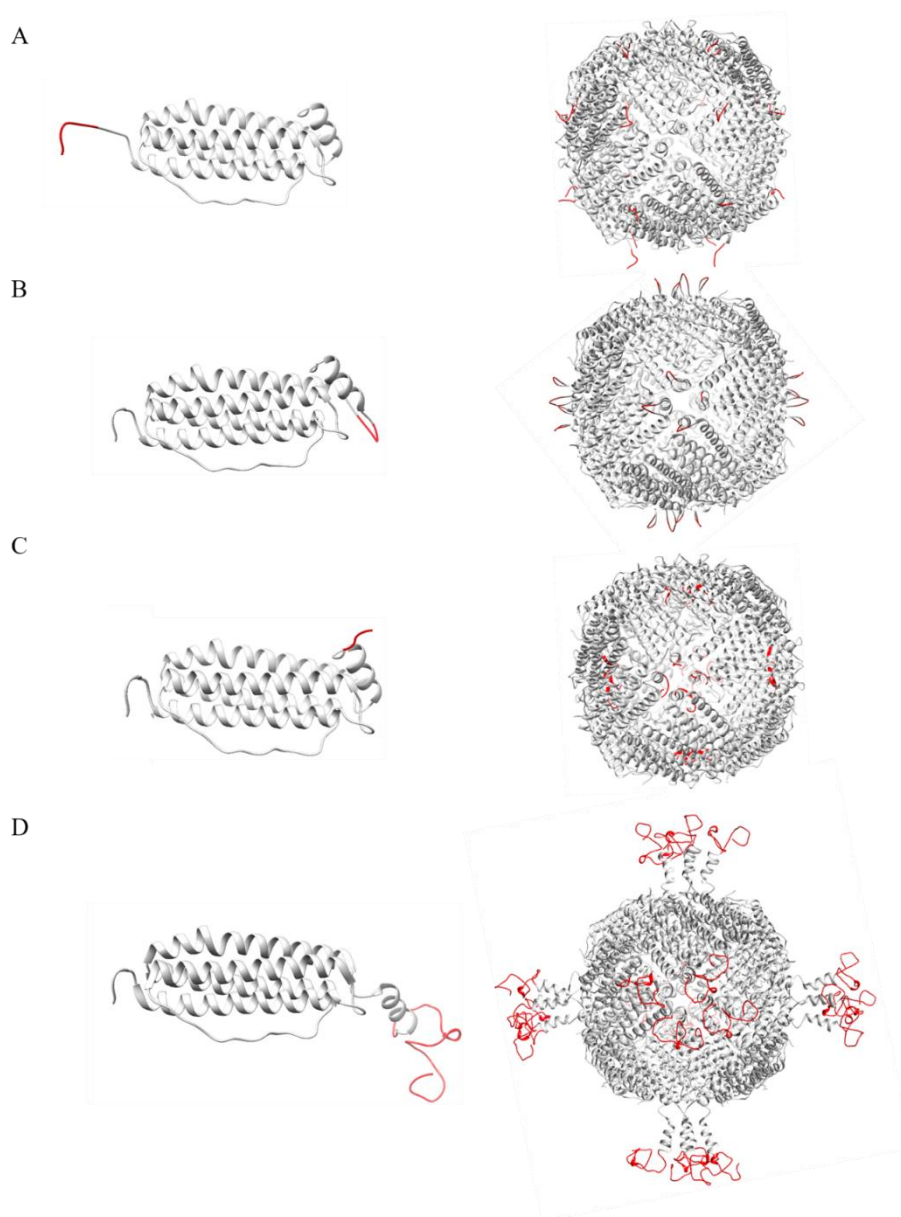


Figure 2.5 Theoretical conformations of HF_n subunits and nanocages with peptide fusion at different sites.

A, N-terminal fused HF_n conformation. **B**, DE turn fused HF_n conformation. **C**, C-terminal fused HF_n conformation. **D**, C-terminal fused HF_n conformation with extruded E helices and inserted peptides. Foreign peptides are shown in red colour. Conformations are modified using PDB 2fha and visualised using Chimera.¹⁵

In some studies, stimuli-responsive peptides are added to help ferritin possess special functions. Fracasso et al. inserted a matrix metallo-proteinases responsive site in PAS fused ferritin designed by Falvo et al. to prevent PAS from blocking the inherent targeting ability of ferritin.⁸⁷

In 2018, Falvo et al. revised the cleavable site-PAS peptide by replacing two residues with acidic residue glutamate to reduce aggregation of the modified ferritin.⁸⁸

Because of the possible E helix turnover, short, foreign peptide inserted at C-terminal will be located inside the ferritin cavity, whilst longer peptides, will be extruded with E helix (**Figure 2.5C** and **2.5D**). Assuming the average volume of every amino acid equals 100 \AA^3 , the radius of ferritin cavity is $\sim 40 \text{ \AA}$.¹⁵ In consequence, this means the cavity volume is $\sim 2.68 \times 10^5 \text{ \AA}^3$. Consequently, in maximum, twenty-four 111 aa peptides, including E helix, can be stored inside the ferritin cavity assuming negligible steric hindrance and repulsion force between molecules and structure of peptides.

Two, short, model antigenic peptides, OT-1 (8aa) and OT-2 (17 aa) were introduced on the exterior surface and inner cavity of ferritin through DE turn and C-terminal fusion. Despite the different location of epitopes, they exhibited desirable immunogenicity.⁸⁹ Red fluorescence protein with a six-consecutive His residues were fused onto human ferritin C-terminal and displayed on ferritin exterior surface.⁹⁰ In another study, enhanced green fluorescence protein was fused to human H-chain ferritin C-terminal with and without flexible linker. It was found that adding the spacer increased particle fluorescence.⁹¹ A 12 aa targeting peptide together with a 15 aa linker was connected to ferritin C-terminal and protruded from the outer surface of ferritin cage.⁹² Because E helix is not necessary for the formation of ferritin nanocage, E helix truncated ferritin (short ferritin) can be used to reduce steric hindrance caused by long and/or complex exogenous peptides. GALA peptide, the structure of which will transition from random coil to helix as solution pH decrease, was introduced to C-terminal of L-chain short ferritin subunit so that this modified cavity remained intact at physiological pH, but disassembled when pH was 6.0.⁹³ Fenobody, made up of short Archaea ferritin fused with anti H5N1 single-domain antibody at C-terminal, exhibited extended half-life and enhanced affinity over the single-domain antibody.⁹⁴

Because of the high plasticity of ferritin, it is possible to introduce more than one peptide onto the ferritin subunit. Two functional peptides can either merge into one or can be separately added to two insertion sites. Kim et al. reported two successful constructions of two bi-specifically functionalized ferritins.^{63,95} Lee et al. fused a disordered protein for extending half-life and a functional peptide for targeting.⁹⁶

Gene engineering however has several practical drawbacks. One of which is that newly introduced peptides which do not serve as targeting moieties can impede original binding ability to its receptors. This can be resolved by introducing cleavable linker between functional peptide and ferritin. Another drawback is that the protein expression level tends to decrease after genetic modification.⁹⁷ In the majority of studies, *E. coli* has been selected as the expression host. This is because it is readily abundant, inexpensive and able to ‘correctly’ fold modified ferritin. It is claimed that the expression yield of RGD-ferritin in *E. coli* is less than one tenth of that for ferritin.⁹⁸ To address this problem, Guo et al. fused PfTrx, a type of thioredoxin scaffold protein that can help increase modified protein expression level onto subunit C-terminal, and reported the expression level increased by more than 5 times.⁹⁹ Sometimes mutants, especially those introducing extra cysteine residues, can form insoluble inclusion bodies, thereby requiring additional refolding process. This further reduces final yield. Ferritins fused with BCCP, CBP and α TSHscFv were all expressed in insoluble form, and to simplify the process, researchers used re-folding to restore protein natural structures and to concomitantly load cargo inside.¹⁰⁰ Moreover, sequence change is highly likely to cause structural alterations and affect drug loading performance in an uncertain way. Multiple functionalisation is hard to realize in the light of sequence design, despite the multiple available modification sites on ferritin.

2.4.3 Hybridisation

Hybridisation is relatively less used but is a potent option to functionalise ferritin. Natural heterochain ferritin is likely the inspiration for this approach. This typically incorporates two different ferritin subunits and sometimes can achieve $1+1 > 2$ effect.

One way to achieve hybridisation involves the facile mixing of different types of modified subunits and then the reassembly of a hybrid 24-mer, **Figure 2.6A**. Several studies have reported generating chimeric ferritins using this approach.¹⁰¹ In 1997, it was reported that variants of human H-chain ferritin successfully co-assembled with wild-type H-chain ferritin through a 6 M guanidinium chloride assisted acidic pH disassembly-reassembly.¹⁰² Two types of fluorescence nanoprobes with stimuli-responsive abilities based on hybrid ferritin were constructed via acidic pH-induced approach.^{103, 104} Additionally, hybrid ferritin can work as drug nanocarriers, as does homomeric ferritin. ⁶⁴Cu was loaded into hybrid ferritin and nicked

ferritin demonstrated improved drug-loading-ratio and pH-responsive drug release.^{105, 106} Insoluble modified ferritin subunits can assemble with the aid of soluble subunits. Boumaiza et al. successfully assembled modified insoluble monomers with L-chain or H-chain monomers through denaturation and re-folding.¹⁰⁷ Co-expression of two types of ferritin subunits in the expression host is a hybridisation option free from the complex and uncontrollable denaturation-refolding (**Figure 2.6B**). Lee et al. engineered two dual-epitope ferritin through synthesizing two plasmids harbouring two ferritin-epitope genes and transformed them into the same *E. coli*.¹⁰⁸

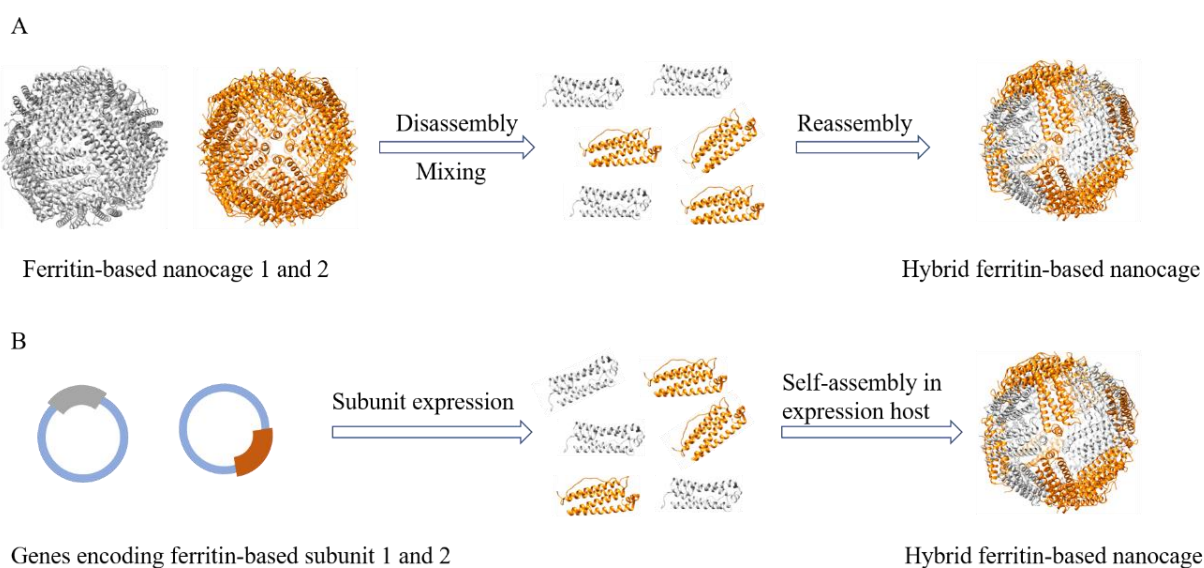


Figure 2.6 Procedures of hybridisation of two types of ferritin subunits through two different means. **A**, disassembly-reassembly means. **B**, co-expression means.

Currently, hybridisation is applicable only to mixing a few types of subunits, preferably 2 to 3. This is because, it requires condition-optimisation to precisely control the average ratio of those types of subunits in one hybrid 24-mer.

A number of functionalisation approaches have therefore been established for ferritin. Through these, different outcomes including, localized delivery, half-life extension, and special stimuli-responsivity, together with the addition of new functions are readily achieved. What functionalisation to add, and how to functionalize are necessarily selected, based on particular demands for treatment of particular disease and what benefits and risks any foreign moiety might bring.

With its robust hollow sphere structure, desirable properties, diverse drug-loading approaches and various functionalisation means, ferritin has been investigated extensively and a wide range of drugs tried. **Table 2.2** presents a comparative summary of studies using ferritin as a drug nanocarrier over the past 20 years. The table summarises the commonly loaded drug types, loading approaches, ferritin functionalisation and overall aim.

Table 2.2 Comparative summary of ferritin drug-loading and functionalisation over the past 20 years.

Cargo	Resource	Loading/hybridization	Functionalisation/Aims	References
Gefitinib	Recombinant human H-chain ferritin	Passive-loading, pH 7.2, using DMSO as co-solvent	/	39
AgNO₃, AuCl₃, Cisplatin	Recombinant human H-chain ferritin	Passive-loading: AgNO ₃ , AuCl ₃ , pH 7.8 pH-induced: Cisplatin, pH 2.0-7.5	/	40
Iron oxide nanoparticles	Recombinant human H-chain ferritin	Passive-loading of Fe ²⁺ and H ₂ O ₂	/	41
CuCO₃ nanoparticles	Recombinant human H-chain ferritin	Passive-loading	/	42
2 nm Platinum nanoparticles	Horse-spleen ferritin	Passive-loading at 37 °C	/	43
Uranyl acetate dehydrate, phosphotungstic acid, curcumin, doxorubicin	Recombinant human H-chain ferritin	Passive-loading	/	44
Daunomycin	Commercial apoferritin powder	Passive-loading, pH 5.0, Co-encapsulation of PLAA	/	45
Fe²⁺ and NADH/FMN, to form a Fe core	Recombinant frog ferritin (H type)	Passive-loading, using low concentrations of urea, GdHCl and Triton X-100 to open pores	/	22
Rutin	Soybean seed ferritin	Passive-loading using 2 mM GdHCl to open channels	/	46

Epigallocatechin Gallate	Red bean ferritin	Passive-loading using 20 mM urea to open hydrophobic channel	/	47
Rutin	Soybean seed ferritin	Passive-loading at 60 °C	/	48
Doxorubicin	Recombinant human H-chain ferritin	High hydrostatic pressure (500 MPa) passive-loading	/	49
Doxorubicin	Human H-chain ferritin	pH-induced, pH 2.0-7.5	/	50
Doxorubicin	Equine ferritin	pH-induced	Magnetic nanoparticle coating through biotin-streptavidin binding	51
Streptomycin	Equine spleen ferritin	pH-induced, pH 2.0-8	/	54
Doxorubicin	Horse-spleen ferritin	pH-induced, pH 2.5-7.4, stepwise increase of pH	/	52
Cisplatin	Horse-spleen apoferritin	Alkaline pH-induced, pH 13-7	/	55
Doxorubicin	Recombinant human heavy chain ferritin	8 M urea-induced	/	23, 29
Doxorubicin	<i>P. furiosus</i> ferritin	8 M urea-induced	N-terminal fusion: SP94 peptide (SFSIIHTPILPL), for targeting	59
Atropine, carbachol, doxorubicin	Ferritin	8 M urea-induced	/	60

Erastin and rapamycin	Equine spleen ferritin with iron core	Through emulsification	/	61
CGKRR(KLAKLAK)₂, proapoptotic peptide.	Recombinant human H-chain ferritin, without the E helix	Through fusion	N-terminal fusion: CGKRR(KLAKLAK) ₂ , peptide drug. C-terminal fusion: cleavage linker+ GFP	63
Glucagon-like peptide-1, an incretin.	Recombinant human H-chain ferritin	Through fusion	N-terminal fusion: Glucagon-like peptide-1, peptide drug, with a linker: GSFGG	64
F5M	Hyperthermophilic archaeon <i>P. furiosus</i> ferritin	Covalently conjugation through disulphide bond formation	Flexible DE loop region insertion: thrombin cleavage peptide, for enzyme responsive drug release C-terminal mutation: Cysteine replaced the terminal Glycine, for covalent drug-loading	65
Superoxide dismutase	Archaeoglobus fulgidus ferritin mutant, Horse-spleen ferritin	Covalently conjugation	Monoclonal antibody conjugation, for targeting	66
Doxorubicin and Bi₂S₃ nanocrystals	Horse-spleen ferritin	Doxorubicin: pH-induced Bi ₂ S ₃ nanocrystals: conjugation	/	62
5-FU and Au nanoparticle	Ferritin	Passive-loading, making use of the Au nanoparticle to reduce leakage	/	69
Doxorubicin	Recombinant human H-chain ferritin	pH-induced, using Cu(II), Mn(II), Zn(II) and Fe(III) as helper agents	N-terminal fusion: RGD4C, for targeting	70
Daunomycin	Horse-spleen ferritin	Passive-loading, pH 5.0, PLAA co-encapsulation	Hyaluronic acid conjugation for targeting, using EDC and NHS reagents	71
Doxorubicin	Recombinant human H-chain ferritin	pH-induced, pH 2-8	N-terminal fusion: two types of PAS peptide, for half-life extension	86

Doxorubicin	Recombinant human H-chain ferritin	pH-induced, pH 2.0-7.4	N-terminal fusion: Albumin binding domain+ (GGGS) ₄ linker for half-life extension	72
Doxorubicin	Ferritin	pH-induced	The surface of ferritin/doxorubicin is coated with around 20 nm thick Au nanoshell for multiple stimuli-sensitivities	73
ZnF16Pc	Recombinant human H-chain ferritin	Passive-loading in PBS	Folic acid conjugation, using EDC and NHS, for targeting	74
/	Recombinant mutated human H-chain ferritin	/	Mutation: replace two Cys residue on ferritin subunit by Ser Chemical modification: specific modification of β -cyclodextrin, for targeting	75
Uranium	Horse-spleen ferritin	Passive-loading, pH 7.5	Conjugation of Fab' antibody fragment for targeting	76
Resveratrol	Ferritin	pH-induced	Chemical conjugation of PEG ₂₀₀₀ -RGD for targeting	77
Cisplatin	Recombinant human H-chain ferritin	pH-induced, pH 2-7.5	mAb conjugation using NHS-PEG-MAL cross-linker for targeting	78
/	Recombinant human H-chain ferritin	/	N-terminal fusion: α -melanocyte-stimulating hormone peptide, for targeting Chemical modification: PEG	79
/	Chimeric archaeal-human ferritin	/	Mutation: substitution of 12 residues located at the loop between B and C helix, for targeting	82
Cytochrome C	Chimeric archaeal-human ferritin	Neutral pH-induced disassembly-reassembly	Mutation: substitution of 12 residues located at the loop between B and C helix, for targeting	83

Doxorubicin	Ferritin	Passive-loading, using Cu (II) as helper reagent	Mutation at ferritin 2-fold interface to strengthen ferritin assembly	84
/	Recombinant human L-chain ferritin	/	Flexible DE loop region insertion: interleukin-4 receptor (IL-4R)-targeting peptide, AP-1	85
Doxorubicin	Recombinant human H-chain ferritin	pH-induced, pH 2-7.5	N-terminal fusion: circulation prolonging peptide (CNARGDMHC) and a flexible linker GGGGS	109
Doxorubicin	Recombinant human H-chain ferritin	pH-induced, pH 2-8	N-terminal fusion: MMPs cleavable peptide linker + PAS peptide for prolonging half-life	87
Doxorubicin and mitoxantrone	Recombinant human H-chain ferritin	pH-induced, pH 2.5-7.5	N-terminal fusion: 20 aa revised PAS peptide which has two glutamate residues for prolonging half-life and reduce aggregation	88
/	Ferritin	/	OT-1 or OT-2 epitope fused on C-terminal or DE loop	89
/	Human H-chain ferritin	/	C-terminal fusion: G ₃ SG ₃ TG ₃ SG ₃ linker +red fluorescence protein and 6 His residues, for imaging	90
/	Human H-chain ferritin	/	C-terminal fusion: G ₃ SG ₃ TG ₃ SG ₃ linker+ enhanced green florescence protein +6 His residues or enhanced green florescence protein +6 His residues, for imaging	91
			Chemical modification: Aptamer, for targeting	
Iron and arsenate	Human H-chain ferritin	Passive-loading	C-terminal fusion: G ₃ SG ₃ TG ₃ SG ₃ linker + EG11 peptide (YHWYGYTPQNVI), targeting peptide	92
/	Recombinant human L-chain ferritin	/	C-terminal fusion: 3 types of GALA peptide, for pH 6-induced dissociation	93
/	<i>P. furiosus</i> ferritin	/	Substituting DE turn and E helix with anti-H5N1 nanobody through linker GGGSGGGSGGGGS	94

/	Recombinant human L-chain ferritin	/	N-terminal and Flexible loop region fusion: RGD peptides, GRDGSP peptide, AP1 peptide, for targeting	95
/	Recombinant human H-chain ferritin	/	N-terminal fusion: epidermal growth factor for targeting	97
/	Recombinant human H-chain ferritin	/	C-terminal fusion: PfTrx, for soluble expression	99
Eu³⁺	Recombinant human L-chain ferritin, in inclusion bodies form.	Conducted at the same time as refolding step, pH-induced or 8 M urea-induced	N-terminal fusion: BCCP, CBP or α TSHscFv	100
/	Recombinant frog M ferritin and its variant	Hybridisation, pH-induced	Wild-type subunit and K104E mutant	101
/	Recombinant human H-chain ferritin and its variants	Hybridisation, 6 M GdnHCl assisted and pH-induced	6 variants: ferritin without 1-13 aa, L169R, C90G-C102A, C90G-C102A-C130A, I85C, D131I-E134F Cys chemical modification: Iodoacetic acid, iodoacetamide and 5-iodoacetamidofluorescein	102
/	Recombinant H-chain ferritin	Hybridization, pH-induced	Subunit type 1: Cy5.5-MMPs cleavable peptide-ferritin, Type 2: black hole quencher-ferritin	103
/	Horse-spleen ferritin	Hybridization, pH-induced	Subunit type 1: carboxy fluorescein-tagged peptide (DRGETGPAC)-ferritin Type 2: black hole quencher-ferritin	104
⁶⁴Cu	Recombinant H-chain ferritin	Hybridization and cargo loading, pH-induced	Subunit type 1: Cy5.5-H chain Type 2: RGD4C-H chain	105
Doxorubicin, doxorubicin-metal complex, curcumin and quercetin	Recombinant human H-chain ferritin	Hybridization, alkaline pH-induced Drug-loading: passive-loading	Subunit type1: ferritin Type 2: E-helix truncated ferritin	106

/	Recombinant human ferritin	Hybridization, denaturation (6 M GdHCl) and re-folding	Subunit type 1: L-chain or H-chain Type 2: Camel hepcidin-H chain C-terminal fusion: A: ferritin subunit+ G ₃ SG ₃ TG ₃ SG ₃ linker +HCV epitope A+His ₆ B: ferritin subunit+ G ₃ SG ₃ TG ₃ SG ₃ linker +HCV epitope B+His ₆ C: ferritin subunit+ G ₃ SG ₃ TG ₃ SG ₃ linker+ HCV epitope C+His ₆ D: ferritin subunit+ G ₃ SG ₃ TG ₃ SG ₃ linker+ extracellular domain of myelin oligodendrocyte glycoprotein + His ₆	107
/	Recombinant Human H-chain ferritin	Hybridization, co-expression of subunit A and B, subunit B and C		108
/	Recombinant human H-chain ferritin	/	^{99m} Tc labelling	110
Doxorubicin	Equine spleen ferritin	pH-induced, pH 3-7	/	111
Curcumin	Recombinant human H-chain ferritin	pH-induced, pH 12.5-7.4	/	112
Doxorubicin	Horse-spleen ferritin	pH-induced, pH 2.7-6.5	Ferritin surface was modified with 1.3 nm Au particle and conjugated with antibody via a peptide (HWRGWVC) linker	113
Proanthocyanidin dimer	Horse-spleen ferritin	pH-induced, pH 2-7.5	/	114
Everolimus	Human H-chain ferritin	Passive-loading and pre-complexate with Cu(II)	/	115

2.5 Computational tools used in prediction of protein conformation, property and protein-ligand interaction

Computational technique is a potent tool from a microscopic and high-resolution perspective to study bio-macromolecules. It supports experiments by either guiding experiment design or interpreting experimental findings. It simulates macromolecule structure and interactions between molecules in real microenvironments. Currently it has been applied in a wide range of bio-macromolecule-related research fields, such as prediction of protein conformation and structural stability, protein-ligand interaction and complex conformation, drug discovery, etc.^{116, 117}

As computational technique is a significant subject, this chapter focuses only on its applications including: 1) Protein conformation prediction, validity assessment and biochemical property analysis by MD simulation. These techniques offer reliable structural and biochemical property information of a newly designed protein or modified proteins. It can be adopted in functionalised ferritin structure prediction and the analysis of the impact of inserted functional peptides on ferritin conformation and biochemical property. 2) Protein-ligand interaction analysis by molecular docking. One of the mechanisms of drug remaining in ferritin is chemical interaction. Molecular docking technique can reveal the detailed non-covalent interaction information between ferritin and drug in drug-loading process. In addition, the fused functional peptide in a functionalised ferritin may interact with ferritin subunit residues through non-covalent bonds to disrupt ferritin structure. Molecular docking can also shed light on these possible interactions.

2.5.1 New protein conformation prediction by MD simulation

MD simulation is the calculation of atom movements of a molecule in a defined environment in a period of simulation time using force field and restraint algorithms. It can be used to predict protein structure and assess the stability of the predicted structure. Extensively used MD simulation services include packages such as AMBER, CHARMM, NAMD, GROMACS and online servers.^{118, 119}

2.5.1.1 Protein conformation prediction

Protein conformation is the basis for its function. In the design of new functionalised proteins, such as a modified ferritin, integration of ferritin with functional peptides or mutations of ferritin is frequent. The structure of functionalised ferritin is uncertain but critical to the realisation of design aims. The structure comparison of the functionalised ferritin with ferritin is also of great meaning in studying the impact of the modification. Direct determination of the new functionalised and original protein structure by X-ray crystallisation or nuclear magnetic resonance (NMR) is a compelling and extensively used approach. However, crystallisation requires extremely high purity of target protein, and is time-consuming and does not work for many proteins.¹²⁰ Protein crystal structure is likely to be different from its structure in functioning environment because protein is non-rigid and dynamic. NMR is usually restricted to proteins smaller than 100 kDa and difficult to analyse proteins in a dynamic fashion.¹²¹ Compared with these approaches, MD simulation offers an edge in analysis of protein structure in a dynamic way and in a simulated real environment.¹²²

Strategies of structure prediction mechanism of MD software or webserver have been divided into 4 categories by Dorn et al, first principle without database information, first principle with database information, fold recognition and threading method, plus comparative modelling and sequence alignment method.¹²³ First principle without database information strategy is on a thermodynamic basis and the folding process is basically driven by the minimisation of structure global free energy. Currently, due to the limitations on accuracy of algorithms and intense calculation demand, this strategy is mostly restricted to peptides.¹²⁴ For example, PEP-FOLD is a facile and reliable online service to predict peptide structure. 5-36 residue linear peptide and peptide with one or two disulphide bonds can be predicted within 40 min if using 40 cores.^{125, 126} Kukic et al. implemented a CamTube force field into GROMACS and successfully obtained a folded 60 residue polyvaline peptide by a 20 μ s metadynamic simulation.¹²⁷

The other 3 strategies involve extensive searches and comparisons of the input sequence with the ones in data base. The protein 3D structure data base, protein data bank (PDB), at present has over 170 000 structures of either single biomacromolecule or complex obtained from crystallisation or NMR.¹²⁸ In contrast with the first strategy, these are more applicable to large

and complex proteins, especially modified proteins such as mutants and foreign peptide fused proteins. Services such as I-TASSER, PROSPECT and SWISS-MODEL belong with these three strategies. However, because of the different algorithms and strategies behind, structures predicted by different services can differ. In addition, the predicted structure validity is not 100 % guaranteed, due to the limitations of the algorithms and difficulties in simulating the complex real environment. A judicious selection of suitable service is indispensable. The selection mainly depends on whether the force field and restraint algorithms implemented matches with the real environment and the type of target protein. For services using the first strategy, prediction precision is higher, and the simulation takes shorter in small linear peptides. For services involving data base information search, proteins with high homogeneity can achieve a greater precision in structure prediction.

2.5.1.2 Protein conformation validity assessment

After obtaining a predicted 3D structure of the new functionalised protein, the structure validity needs to be evaluated. This can be achieved by putting the obtained structure in MD simulation to monitor its trajectory and evaluate its stability through measurement of some parameters. The source of the obtained structure can be PDB file from crystallisation or NMR, from MD simulation, complex structures from molecular docking analysis or an artificial 3D structure. An artificial 3D structure of fused proteins or mutants is obtained by directly editing the original protein 3D structure from PDB. The artificial structure is based on theoretical structures of its components, and usually is the expected structure of the new protein. The stability of this structure can be compared with that of the original protein in MD simulation to test the validity of the expectation.

For example, various types of macromolecule structure stability have been assessed by parameter measurements in GROMACS. GROMACS is a free, open-source MD simulation software. It offers a wide selection of analyses and can be applied for multiple aims, including the validity of protein conformation.¹¹⁸ Panchal et al. compared structural stability of a series of helicase mutants through analysis of parameters such as root mean square deviation (RMSD) and radius of gyration (Rg) in GROMACS to find the most stable mutant.¹²⁹ RMSD measurement in GROMACS has also been used as a fast means to evaluate the quality of structure of a 61 residue protein.¹³⁰ RMSD is the mean structure coordinate deviations of the dynamic structure in simulation from a reference structure. When the reference structure is the

input one, RMSD plot reflects if this initial structure can maintain stable throughout the simulation. The lower RMSD is in simulation, the more stable of the input structure is. When the reference structure is a crystal structure of relevant protein, such as the wild-type structure of a modified protein, RMSD measurement demonstrates the stability impact of the modification on the original protein. Rg reveals the compactness of the structure during simulation. If the structure is stably folded, the fluctuation of Rg should be small. Both of them are commonly used parameters in structure validity evaluation.¹³¹

2.5.1.3 Biochemical property analysis

In the meantime of the MD simulation validity assessment of the obtained structure, some biochemical property information of the new protein are revealed, which can help interpret differences between modified proteins and the original ones in other characterisations and figure out the impact of the modifications. USCF Chimera, PyMOL, VMD and Discovery Studio are powerful tools to visualise the surface hydrophobicity, surface charge, non-covalent bonds within the protein from the resultant structure and trajectory file of MD simulation.^{132, 133} In multiple chain protein such as ferritin and virus-like particle (VLP), MD simulation can study the complex assembly-disassembly process.¹³⁴ A 200 ns GROMACS simulation of ferritin assembly at pH 2 along with real-time high-speed atomic force microscopy revealed a detailed dynamic process of ferritin disassembly and found that it started with the opening of the 3-fold channels.¹³⁵ MD simulation of a murine polyomavirus VLP capsomere found that the hydrophobicity and charge property of capsomere and the solution conditions impacted on VLP assembling and these could explain their experiment results.¹³⁶

All the functions of MD simulation mentioned above, prediction of 3D structure, obtained structure validity evaluation and biochemical property information revelation allow researchers to compare modified proteins with the original ones, examine the impact of modification on the original protein and explain the differences in other experimental characterisations. In a lot of cases, MD simulation offers a good method to interpret and verify experimental findings. The selection of the suitable MD simulation service is crucial and needs to be judicious.

2.5.2 Protein-ligand interaction study by molecular docking

As mentioned in section 2.4, drug loaded in ferritin interacts with ferritin through non-covalent bonds in most cases. Functional peptides inserted on ferritin can affect ferritin structure and property (section 2.5). The study of the interactions between drug and ferritin can interpret the phenomena found in drug-loading process and possibly guide drug-loading experiment. Protein-ligand interaction study also has the potential to investigate the inserted peptide interactions with ferritin subunit and the way it affects assembly and structure.

Current experimental approaches for protein-ligand interaction analysis include the spectroscopic methods such as Fourier-transform infrared spectroscopy (FTIR), circular dichroism (CD), ultraviolet spectroscopy, fluorescence spectroscopy and Raman spectroscopy, together with isothermal titration calorimetry (ITC) and hydrogen/deuterium exchange mass spectrometry.¹³⁷⁻¹⁴⁰ Spectroscopic methods study the different spectra of the complex with that of the original protein and usually make use of some probes and markers to identify the interaction sites to obtain local structure information. In contrast with spectroscopic methods, experimental methods such as X-ray crystallisation and NMR give information about the whole complex 3D structure.¹⁴¹ Multiple techniques are usually used in combinations to obtain complete information and increase the result accuracy. However, almost all these experimental approaches require a dominant form of complex and a strong interaction to achieve a desirable analysis accuracy. If ligand interacts with protein in multiple ways, the complexity of the sample can overwhelm the data interpretation. If the interaction is weak and slow, these techniques may take a long time or even cannot obtain any results because of the sensitivity. What is more, ligands that are peptides are not always feasible because of the high flexibility and numbers of interactive ways to the protein.

Computational technique, molecular docking, on the other hand, has offered a rapid and powerful way to investigate multiple ways of both protein-small molecule and protein-peptide interaction. It has been well-established in the past decades and widely used alone or along with experimental techniques in research.^{142, 143} Molecular docking adopts a database-based searching algorithm to rank the affinity and possible binding model of 2 molecules.⁵ The 2 molecules are called target and ligand, respectively. Target is usually a macromolecule such as a protein. The structure of it is preferred to be known from crystallisation or NMR. In some

cases, structure based on homology model can apply. Ligand can be a specific one or a drug database as in drug discovery research. Due to the limitations on algorithm and computer resources, ligand can be small molecules, peptides and small proteins.

According to the flexibility of target and ligand, docking strategy can be generally divided into two different types. Initially, as the lock-key hypothesis implies, both target and ligand are treated as rigid body. This strategy is on a basis of shape complementarity and works well in bound complexes with strong interactions.¹⁴⁴ As the introduction of ‘dynamic’ concept in molecule movements, flexible docking strategy has become dominant. In flexible docking strategy, structure of one or both of target and ligand are regarded flexible. Ligand bonds and angles are always treated as rotatable and flexible if it is small molecules. Target and large ligand, however, because of the structural complexity, sometimes is treated as a rigid body, or in other cases, only local areas such as the active sites are treated with flexibility to a certain degree.

In terms of ferritin drug-loading mechanism analysis, protein-small molecule docking techniques is sufficient. Commonly used flexible protein-small molecule docking program include AutoDock, GOLD, Flex X, and so on. AutoDock works well with small molecule ligand but not with peptide ligand with over 4 residues. Subramanian et al. used MD simulation to obtain structures of ferritin at different simulation times and then applied AutoDock to identify the binding site of dihydroflavin on ferritin.¹⁴⁵ Shahwan et al. used AutoDock Vina, an upgraded version of Autodock, to find out the most possible human ferritin (PDB ID: 3AJO)-enzyme inhibitor Donepezil complex. Following that, the complex underwent a 100 ns MD simulation in GROMACS for stability comparison with ferritin to assess complex structure validity, and the interaction type and sites in complex were visualised using PyMOL and Discovery Studio.¹⁴⁶ To obtain a reliable complex prediction, the selection of docking program should be careful, and the combination of MD simulation structure validity assessment is beneficial.

In ferritin-functional peptide interaction analysis, protein-peptide and protein-protein docking need to be applied. However, because of the relatively large size of ligand compared with small molecules, it is time-consuming and challenging.¹⁴⁷ Protein-peptide docking is more feasible and can achieve a relatively high accuracy compared with protein-protein docking. Docking

programs in use can be classified into three types, template-based, local and global docking by different searching algorithm mechanisms. Ciemny et al.'s review has well documented these.¹⁴⁸ For example, HADDOCK is a potent local docking webserver driven by experimental data.¹⁴⁹ It is a rigid-body docking process but provides refinement of the binding complex through endowing peptide and interactive protein residues with full flexibility, and ranks binding models based on binding free energy. It applies to protein-peptide and protein-protein docking. Deplazes et al. has figured out the binding model of spider toxin peptide with ion channel protein using HADDOCK.¹⁵⁰ A benchmark protein-protein docking analysis of 37 membrane proteins has been achieved by Koukos et al.¹⁵¹ Other techniques are often used along with docking to verify complex structure in research. Cytochrome b5–cytochrome c protein to protein docking were successfully performed using HADDOCK with NMR verification.¹⁵²

Molecular docking is effective and efficient in study of possible multiple interaction ways of protein and ligand. Ligands are preferred to be small molecules and peptides. The selection of suitable program is of great importance and depends on the type of the ligand, whether it is small molecule or peptide or protein, and the searching algorithm and target ligand flexibility level of the program. The protein-ligand interaction is better to be further analysed by MD simulation or other experiment techniques to assess validity.

2.6 Chapter summary

Based on the review of relevant literature in this chapter, several key findings are made:

1. Mammalian ferritin has a unique hollow sphere structure made up of 24 subunits and 14 channels on the shell. It has desirable thermal and pH stability compared with common proteins. HFn has an inherent active tumour targeting ability, high biocompatibility and low immunogenicity in medical application. All the above features have made ferritin, especially HFn, a drug delivery platform with great potentials.
2. Passive loading, and disassembly-reassembly are 2 main ferritin drug loading approaches. Other alternatives such as hybridisation have also been explored. Drug loading mechanism is likely to be physical entrapment and chemical interactions. Currently the bottleneck of ferritin drug loading performance is its unsatisfactory

protein recovery percentage and drug loading ratio. Investigation of ferritin drug loading process is needed.

3. As a drug delivery platform, ferritin can be improved by incorporation of functional moieties. Ferritin structure holds high plasticity, can be functionalised by chemical modification and gene technology. Fusion is the most common way. Fusion of targeting moiety and half-life extender can improve ferritin performance as drug delivery platform. However, fusion may affect ferritin structure and self-assembly. The selection of functional moiety and functionalisation site are of great importance.
4. Computational techniques are powerful tools to supplement experiments. MD simulation can be used to predict 3D structure of functionalised ferritin, evaluate obtained structure validity, offer biochemical property information and reveal the impact of functional peptide on ferritin.
5. Molecular docking investigates multiple interaction ways between a ligand and a target macromolecule and ranks the possible binding models. Ligand can be small molecules or peptides or even proteins. Small molecule ligand can achieve a greater predictive accuracy compared with peptide and protein ligand. Molecule docking analysis combined with MD simulation has great potentials to analyse ferritin drug-loading mechanism and study the effects of functional peptide interactions with ferritin subunit on ferritin structure through assessment of protein-ligand complex structural stability and obtainment of complex biochemical property.

To address these current problems, in this research thesis HFn is improved by functionalisation through fusion of two foreign functional peptides (PAS and RGDK). Condition optimisation of drug loading approach has been conducted. The impacts of functionalisation on expression, purification, structural and bioactivities are studied by mainly experiments. Computational analysis using Gromacs MD simulation and HADDOCK molecular docking has been used to help understand the P impact on ferritin self-assembling. Mechanisms behind doxorubicin loading is analysed by Gromacs MD simulation and PyRx molecular docking.

In the next chapter, **Chapter 3**, the design, construction and expression of HFn and functionalised HFns are investigated. Computational techniques, MD simulation and molecular docking are used to probe the problems caused by functional peptide in some functionalised HFns.

2.7 References

1. Belletti, D.; Pederzoli, F.; Forni, F.; Vandelli, M. A.; Tosi, G.; Ruozi, B., Protein cage nanostructure as drug delivery system: magnifying glass on apoferritin. *Expert Opinion on Drug Delivery* 2017, 14 (7), 825-840.
2. Jutz, G.; van Rijn, P.; Santos-Miranda, B.; Boker, A., Ferritin: a versatile building block for bionanotechnology. *Chemical Reviews* 2015, 115 (4), 1653-1701.
3. He, J.; Fan, K.; Yan, X., Ferritin drug carrier (FDC) for tumor targeting therapy. *Journal of Controlled Release* 2019, 311-312, 288-300.
4. Khoshnejad, M.; Parhiz, H.; Shuvaev, V. V.; Dmochowski, I. J.; Muzykantov, V. R., Ferritin-based drug delivery systems: hybrid nanocarriers for vascular immunotargeting. *Journal of Controlled Release* 2018, 282, 13-24.
5. Morris, G. M.; Lim-Wilby, M., Molecular docking. In *Molecular modeling of proteins*, Springer: 2008; pp 365-382.
6. Durrant, J. D.; McCammon, J. A., Molecular dynamics simulations and drug discovery. *BMC biology* 2011, 9 (1), 1-9.
7. Laufberger, V., Sur la cristallisation de la ferritine. *Bulletin de la Société de chimie biologique* 1937, 19, 1575-1582.
8. Heger, Z.; Skalickova, S.; Zitka, O.; Adam, V.; Kizek, R., Apoferritin applications in nanomedicine. *Nanomedicine* 2014, 9 (14), 2233–2245.
9. He, D.; Marles-Wright, J., Ferritin family proteins and their use in bionanotechnology. *New Biotechnology* 2015, 32 (6), 651-657.
10. Andrews, S. C.; Harrison, P. M.; Yewdall, S. J.; Arosio, P.; Levi, S.; Bottke, W.; von Darl, M.; Briat, J. F.; Laulhère, J. P.; Lobreaux, S., Structure, function, and evolution of ferritins. *Journal of Inorganic Biochemistry* 1992, 47 (1), 161-174.
11. Zhen, Z.; Tang, W.; Todd, T.; Xie, J., Ferritins as nanoplatfoms for imaging and drug delivery. *Expert Opinion on Drug Delivery* 2014, 11 (12), 1913-1922.
12. Recalcati, S.; Invernizzi, P.; Arosio, P.; Cairo, G., New functions for an iron storage protein: the role of ferritin in immunity and autoimmunity. *Journal of Autoimmunity* 2008, 30 (1-2), 84-89.
13. Cornell, T. A.; Srivastava, Y.; Jauch, R.; Fan, R.; Orner, B. P., The crystal structure of a maxi/mini-ferritin chimera reveals guiding principles for the assembly of protein cages. *Biochemistry* 2017, 56 (30), 3894-3899.
14. Wahlgren, W. Y.; Omran, H.; von Stetten, D.; Royant, A.; van der Post, S.; Katona, G., Structural characterization of bacterioferritin from *Blastochloris viridis*. *Plos One* 2012, 7 (10), e46992.
15. Ferraro, G.; Ciambellotti, S.; Messori, L.; Merlino, A., Cisplatin binding sites in human H-chain ferritin. *Inorganic Chemistry* 2017, 56 (15), 9064-9070.
16. Hempstead, P. D.; Yewdall, S. J.; Fernie, A. R.; Lawson, D. M.; Artymiuk, P. J.; Rice, D. W.; Ford, G. C.; Harrison, P. M., Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution. *Journal of Molecular Biology* 1997, 268 (2), 424-448.

17. Stillman, T. J.; Hempstead, P. D.; Artymiuk, P. J.; Andrews, S. C.; Hudson, A. J.; Treffry, A.; Guest, J. R.; Harrison, P. M., The high-resolution X-ray crystallographic structure of the ferritin (EcFtnA) of *Escherichia coli*; comparison with human H ferritin (HuHF) and the structures of the Fe³⁺ and Zn²⁺ derivatives. *Journal of Molecular Biology* 2001, 307 (2), 587-603.
18. Theil, E. C.; Liu, X. S.; Tosha, T., Gated pores in the ferritin protein nanocage. *Inorganica Chimica Acta* 2008, 361 (4), 868-874.
19. Takahashi, T.; Kuyucak, S., Functional properties of threefold and fourfold channels in ferritin deduced from electrostatic calculations. *Biophysical Journal* 2003, 84 (4), 2256-2263.
20. 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. *Journal of Biological Chemistry* 2010, 285 (3), 1950-1956.
21. Luzzago, A.; Cesareni, G., Isolation of point mutations that affect the folding of the H chain of human ferritin in *E. coli*. *The EMBO Journal* 1989, 8 (2), 569-576.
22. Liu, X.; Jin, W.; Theil, E. C., Opening protein pores with chaotropes enhances Fe reduction and chelation of Fe from the ferritin biomineral. *Proceedings of the National Academy of Sciences* 2003, 100 (7), 3653-3658.
23. Fan, K.; Jia, X.; Zhou, M.; Wang, K.; Conde, J.; He, J.; Tian, J.; Yan, X., Ferritin nanocarrier traverses the blood brain barrier and kills glioma. *ACS Nano* 2018, 12 (5), 4105-4115.
24. Stefanini, S.; Cavallo, S.; Wang, C. Q.; Tataseo, P.; Vecchini, P.; Giartosio, A.; Chiancone, E., Thermal stability of horse spleen apoferritin and human recombinant H apoferritin. *Archives of Biochemistry and Biophysics* 1996, 325 (1), 58-64.
25. Dostalova, S.; Vasickova, K.; Hynek, D.; Krizkova, S.; Richtera, L.; Vaculovicova, M.; Eckschlager, T.; Stiborova, M.; Heger, Z.; Adam, V., Apoferritin as an ubiquitous nanocarrier with excellent shelf life. *International Journal of Nanomedicine* 2017, 12, 2265-2278.
26. Martsev, S. P.; Vlasov, A. P.; Arosio, P., Distinct stability of recombinant L and H subunits of human ferritin: calorimetric and ANS binding studies. *Protein Engineering* 1998, 11 (5), 377-381.
27. Santambrogio, P.; Levi, S.; Arosio, P.; Palagi, L.; Vecchio, G.; Lawson, D. M.; Yewdall, S. J.; Artymiuk, P. J.; Harrison, P. M.; Jappelli, R., Evidence that a salt bridge in the light chain contributes to the physical stability difference between heavy and light human ferritins. *Journal of Biological Chemistry* 1992, 267 (20), 14077-14083.
28. Kim, M.; Rho, Y.; Jin, K. S.; Ahn, B.; Jung, S.; Kim, H.; Ree, M., pH-dependent structures of ferritin and apoferritin in solution: disassembly and reassembly. *Biomacromolecules* 2011, 12 (5), 1629-1640.
29. Liang, M.; Fan, K.; Zhou, M.; Duan, D.; Zheng, J.; Yang, D.; Feng, J.; Yan, X., H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection. *Proceedings of the National Academy of Sciences of the United States of America* 2014, 111 (41), 14900-14905.
30. Chen, Z.; Zhai, M.; Xie, X.; Zhang, Y.; Ma, S.; Li, Z.; Yu, F.; Zhao, B.; Zhang, M.; Yang, Y.; Mei, X., Apoferritin nanocage for brain targeted doxorubicin delivery. *Molecular Pharmaceutics* 2017, 14 (9), 3087-3097.
31. Cheng, X.; Fan, K.; Wang, L.; Ying, X.; Sanders, A. J.; Guo, T.; Xing, X.; Zhou, M.; Du, H.; Hu, Y.; Ding, H.; Li, Z.; Wen, X.; Jiang, W.; Yan, X.; Ji, J., TfR1 binding

- with H-ferritin nanocarrier achieves prognostic diagnosis and enhances the therapeutic efficacy in clinical gastric cancer. *Cell Death and Disease* 2020, 11 (2), 92.
32. Li, J. Y.; Paragas, N.; Ned, R. M.; Qiu, A.; Viltard, M.; Leete, T.; Drexler, I. R.; Chen, X.; Simone, S. C.; Mohammed, F.; Williams, D.; Lin, C. S.; Schmidt-Ott, K. M.; Andrews, N. C.; Barasch, J., Scara5 is a ferritin receptor mediating non-transferrin iron delivery. *Developmental Cell* 2009, 16 (1), 35-46.
 33. Conti, L.; Lanzardo, S.; Ruiu, R.; Cadenazzi, M.; Cavallo, F.; Aime, S.; Crich, S. G., L-Ferritin targets breast cancer stem cells and delivers therapeutic and imaging agents. *Oncotarget* 2016, 7 (41), 66713.
 34. Turino, L. N.; Ruggiero, M. R.; Stefania, R.; Cutrin, J. C.; Aime, S.; Geninatti Crich, S., Ferritin decorated PLGA/Paclitaxel loaded nanoparticles endowed with an enhanced toxicity toward MCF-7 breast tumor cells. *Bioconjugate Chemistry* 2017, 28 (4), 1283-1290.
 35. Monti, D. M.; Ferraro, G.; Merlino, A., Ferritin-based anticancer metallodrug delivery: Crystallographic, analytical and cytotoxicity studies. *Nanomedicine* 2019, 20, 101997.
 36. Meldrum, F. C.; Wade, V. J.; Nimmo, D. L.; Heywood, B. R.; Mann, S., Synthesis of inorganic nanophasse materials in supramolecular protein cages. *Nature* 1991, 349 (6311), 684.
 37. Li, L.; Zhang, L.; Knez, M., Comparison of two endogenous delivery agents in cancer therapy: Exosomes and ferritin. *Pharmacological Research* 2016, 110, 1-9.
 38. Zang, J.; Chen, H.; Zhao, G.; Wang, F.; Ren, F., Ferritin cage for encapsulation and delivery of bioactive nutrients: from structure, property to applications. *Critical Reviews in Food Science and Nutrition* 2017, 57 (17), 3673-3683.
 39. Kuruppu, A. I.; Zhang, L.; Collins, H.; Turyanska, L.; Thomas, N. R.; Bradshaw, T. D., An apoferritin-based drug aelivery system for the tyrosine kinase inhibitor Gefitinib. *Advanced Healthcare Materials* 2015, 4 (18), 2816-2821.
 40. Mosca, L.; Falvo, E.; Ceci, P.; Poser, E.; Genovese, I.; Guarguaglini, G.; Colotti, G., Use of ferritin-based metal-encapsulated nanocarriers as anticancer agents. *Applied Sciences* 2017, 7 (1), 101.
 41. Fan, K.; Cao, C.; Pan, Y.; Lu, D.; Yang, D.; Feng, J.; Song, L.; Liang, M.; Yan, X., Magnetoferritin nanoparticles for targeting and visualizing tumour tissues. *Nature Nanotechnology* 2012, 7 (7), 459-464.
 42. Xu, N.; Yang, Y. F.; Chen, L.; Xu, L.; Xu, X. J.; Lin, J., Enhancing the effect of pharmacological ascorbate in cancer therapy via acid triggered ferritin nanoparticles. *Advanced Biosystems* 2019, 3 (5), 1900006.
 43. Zhang, L.; Laug, L.; Munchgesang, W.; Pippel, E.; Gosele, U.; Brandsch, M.; Knez, M., Reducing stress on cells with apoferritin-encapsulated platinum nanoparticles. *Nano Letters* 2010, 10 (1), 219-223.
 44. Zhang, S.; Zang, J.; Chen, H.; Li, M.; Xu, C.; Zhao, G., The size flexibility of ferritin nanocage opens a new way to prepare nanomaterials. *Small* 2017, 13 (37), 1701045.
 45. Ma-Ham, A.; Wu, H.; Wang, J.; Kang, X.; Zhang, Y.; Lin, Y., Apoferritin-based nanomedicine platform for drug delivery: equilibrium binding study of daunomycin with DNA. *Journal of Materials Chemistry* 2011, 21 (24), 8700.
 46. Yang, R.; Liu, Y.; Blanchard, C.; Zhou, Z., Channel directed rutin nano-encapsulation in phytoferritin induced by guanidine hydrochloride. *Food Chemistry* 2018, 240, 935-939.
 47. Yang, R.; Liu, Y.; Meng, D.; Chen, Z.; Blanchard, C. L.; Zhou, Z., Urea-driven epigallocatechin gallate (EGCG) permeation into the ferritin cage, an innovative method

- for fabrication of protein-polyphenol co-assemblies. *Journal of Agricultural and Food Chemistry* 2017, 65 (7), 1410-1419.
48. Yang, R.; Tian, J.; Liu, Y.; Yang, Z.; Wu, D.; Zhou, Z., Thermally induced encapsulation of food nutrients into phytoferritin through the flexible channels without additives. *Journal of Agricultural and Food Chemistry* 2017, 65 (46), 9950-9955.
 49. Wang, Q.; Zhang, C.; Liu, L.; Li, Z.; Guo, F.; Li, X.; Luo, J.; Zhao, D.; Liu, Y.; Su, Z., High hydrostatic pressure encapsulation of doxorubicin in ferritin nanocages with enhanced efficiency. *Journal of Biotechnology* 2017, 254, 34-42.
 50. Bellini, M.; Mazzucchelli, S.; Galbiati, E.; Sommaruga, S.; Fiandra, L.; Truffi, M.; Rizzuto, M. A.; Colombo, M.; Tortora, P.; Corsi, F.; Prosperi, D., Protein nanocages for self-triggered nuclear delivery of DNA-targeted chemotherapeutics in Cancer Cells. *Journal of Controlled Release* 2014, 196, 184-196.
 51. Blazkova, I.; Nguyen, H. V.; Dostalova, S.; Kopel, P.; Stanisavljevic, M.; Vaculovicova, M.; Stiborova, M.; Eckschlager, T.; Kizek, R.; Adam, V., Apoferritin modified magnetic particles as doxorubicin carriers for anticancer drug delivery. *International Journal of Molecular Sciences* 2013, 14 (7), 13391-13402.
 52. Kilic, M. A.; Ozlu, E.; Calis, S., A novel protein-based anticancer drug encapsulating nanosphere: apoferritin-doxorubicin complex. *Journal of Biomedical Nanotechnology* 2012, 8 (3), 508-514.
 53. Crichton, R. R.; Bryce, C. F., Subunit interactions in horse spleen apoferritin. Dissociation by extremes of pH. *Biochemical Journal* 1973, 133 (2), 289-299.
 54. Ruozi, B.; Veratti, P.; Vandelli, M. A.; Tombesi, A.; Tonelli, M.; Forni, F.; Pederzoli, F.; Belletti, D.; Tosi, G., Apoferritin nanocage as streptomycin drug reservoir: Technological optimization of a new drug delivery system. *International Journal of Pharmaceutics* 2017, 518 (1-2), 281-288.
 55. Pontillo, N.; Pane, F.; Messori, L.; Amoresano, A.; Merlino, A., Cisplatin encapsulation within a ferritin nanocage: a high-resolution crystallographic study. *Chemical Communications* 2016, 52 (22), 4136-4139.
 56. Hofmann, T.; Harrison, P. M., The structure of apoferritin: degradation into and molecular weight of subunits. *Journal of Molecular Biology* 1963, 6 (4), 256-267.
 57. Listowsky, I.; Blauer, G.; Englard, S.; Bethel, J. J., Denaturation of horse spleen ferritin in aqueous guanidinium chloride solutions. *Biochemistry* 1972, 11 (11), 2176-2182.
 58. Santambrogio, P.; Levi, S.; Cozzi, A.; Rovida, E.; Albertini, A.; Arosio, P., Production and characterization of recombinant heteropolymers of human ferritin H and L chains. *Journal of Biological Chemistry* 1993, 268 (17), 12744-12748.
 59. Jiang, B.; Zhang, R.; Zhang, J.; Hou, Y.; Chen, X.; Zhou, M.; Tian, X.; Hao, C.; Fan, K.; Yan, X., GRP78-targeted ferritin nanocaged ultra-high dose of doxorubicin for hepatocellular carcinoma therapy. *Theranostics* 2019, 9 (8), 2167-2182.
 60. Lei, Y.; Hamada, Y.; Li, J.; Cong, L.; Wang, N.; Li, Y.; Zheng, W.; Jiang, X., Targeted tumor delivery and controlled release of neuronal drugs with ferritin nanoparticles to regulate pancreatic cancer progression. *Journal of Controlled Release* 2016, 232, 131-142.
 61. Li, Y.; Wang, X.; Yan, J.; Liu, Y.; Yang, R.; Pan, D.; Wang, L.; Xu, Y.; Li, X.; Yang, M., Nanoparticle ferritin-bound erastin and rapamycin: a nanodrug combining autophagy and ferroptosis for anticancer therapy. *Biomaterials Science* 2019, 7 (9), 3779-3787.
 62. Zhang, Q.; Chen, J.; Shen, J.; Chen, S.; Liang, K.; Wang, H.; Chen, H., Inlaying radiosensitizer onto the polypeptide shell of drug-loaded ferritin for imaging and combinational chemo-radiotherapy. *Theranostics* 2019, 9 (10), 2779-2790.

63. Kim, S.; Kim, G. S.; Seo, J.; Rangaswamy, G. G.; So, I. S.; Park, R. W.; Lee, B. H.; Kim, I. S., Double-chambered ferritin platform: dual-function payloads of cytotoxic peptides and fluorescent protein. *Biomacromolecules* 2016, 17 (1), 12-19.
64. Su, W.; Tan, H.; Janowski, R.; Zhang, W.; Wang, P.; Zhang, J.; Zhai, H.; Li, J.; Niessing, D.; Sattler, M.; Zou, P., Ferritin-displayed GLP-1 with improved pharmacological activities and pharmacokinetics. *Molecular Pharmaceutics* 2020, 1663-1673.
65. Kang, Y. J.; Park, D. C.; Shin, H. H.; Park, J.; Kang, S., Incorporation of thrombin cleavage peptide into a protein cage for constructing a protease-responsive multifunctional delivery nanoplatform. *Biomacromolecules* 2012, 13 (12), 4057-4064.
66. Shuvaev, V. V.; Khoshnejad, M.; Pulsipher, K. W.; Kiseleva, R. Y.; Arguiri, E.; Cheung-Lau, J. C.; LeFort, K. M.; Christofidou-Solomidou, M.; Stan, R. V.; Dmochowski, I. J.; Muzykantov, V. R., Spatially controlled assembly of affinity ligand and enzyme cargo enables targeting ferritin nanocarriers to caveolae. *Biomaterials* 2018, 185, 348-359.
67. Palombarini, F.; Di Fabio, E.; Boffi, A.; Macone, A.; Bonamore, A., Ferritin nanocages for protein delivery to tumor cells. *Molecules* 2020, 25 (4).
68. Koorts, A. M.; Viljoen, M., Ferritin and ferritin isoforms I: structure-function relationships, synthesis, degradation and secretion. *Archives of Physiology And Biochemistry* 2007, 113 (1), 30-54.
69. Liu, X.; Wei, W.; Huang, S.; Lin, S.-S.; Zhang, X.; Zhang, C.; Du, Y.; Ma, G.; Li, M.; Mann, S.; Ma, D., Bio-inspired protein-gold nanoconstruct with core-void-shell structure: beyond a chemo drug carrier. *Journal of Materials Chemistry B* 2013, 1 (25), 3136-3143.
70. Zhen, Z.; Tang, W.; Chen, H.; Lin, X.; Todd, T.; Wang, G.; Cowger, T.; Chen, X.; Xie, J., RGD-modified apoferritin nanoparticles for efficient drug delivery to tumors. *ACS Nano* 2013, 7 (6), 4830-4837.
71. Luo, Y.; Wang, X.; Du, D.; Lin, Y., Hyaluronic acid-conjugated apoferritin nanocages for lung cancer targeted drug delivery. *Biomaterials Science* 2015, 3 (10), 1386-1394.
72. Wang, C.; Zhang, C.; Li, Z.; Yin, S.; Wang, Q.; Guo, F.; Zhang, Y.; Yu, R.; Liu, Y.; Su, Z., Extending half life of H-ferritin nanoparticle by fusing albumin binding domain for doxorubicin encapsulation. *Biomacromolecules* 2018, 19 (3), 773-781.
73. Li, M.; Wu, D.; Chen, Y.; Shan, G.; Liu, Y., Apoferritin nanocages with Au nanoshell coating as drug carrier for multistimuli-responsive drug release. *Materials Science & Engineering C* 2019, 95, 11-18.
74. Zhen, Z.; Tang, W.; Zhang, W.; Xie, J., Folic acid conjugated ferritins as photosensitizer carriers for photodynamic therapy. *Nanoscale* 2015, 7 (23), 10330-10333.
75. Kwon, C.; Kang, Y. J.; Jeon, S.; Jung, S.; Hong, S. Y.; Kang, S., Development of protein-cage-based delivery nanoplatforms by polyvalently displaying beta-cyclodextrins on the surface of ferritins through copper(I)-catalyzed azide/alkyne cycloaddition. *Macromolecular Bioscience* 2012, 12 (11), 1452-1458.
76. Hainfeld, J. F., Uranium-loaded apoferritin with antibodies attached: molecular design for uranium neutron-capture therapy. *Proceedings of the National Academy of Sciences* 1992, 89 (22), 11064-11068.
77. Zheng, Q.; Cheng, W.; Zhang, X.; Shao, R.; Li, Z., A pH-induced reversible assembly system with resveratrol-controllable loading and release for enhanced tumor-targeting chemotherapy. *Nanoscale Research Letters* 2019, 14 (1), 305.
78. Falvo, E.; Tremante, E.; Fraioli, R.; Leonetti, C.; Zamparelli, C.; Boffi, A.; Morea, V.; Ceci, P.; Giacomini, P., Antibody-drug conjugates: targeting melanoma with

- cisplatin encapsulated in protein-cage nanoparticles based on human ferritin. *Nanoscale* 2013, 5 (24), 12278-12285.
79. Vannucci, L.; Falvo, E.; Fornara, M.; Di Micco, P.; Benada, O.; Krizan, J.; Svoboda, J.; Hulikova-Capkova, K.; Morea, V.; Boffi, A.; Ceci, P., Selective targeting of melanoma by PEG-masked protein-based multifunctional nanoparticles. *International Journal of Nanomedicine* 2012, 7, 1489-1509.
 80. Lawson, D. M.; Treffry, A.; Artymiuk, P. J.; Harrison, P. M.; Yewdall, S. J.; Luzzago, A.; Cesareni, G.; Levi, S.; Arosio, P., Identification of the ferroxidase centre in ferritin. *FEBS Letters* 1989, 254 (1-2), 207-210.
 81. Levi, S.; Luzzago, A.; Franceschinelli, F.; Santambrogio, P.; Cesareni, G.; Arosio, P., Mutational analysis of the channel and loop sequences of human ferritin H-chain. *Biochemical Journal* 1989, 264 (2), 381-388.
 82. de Turrís, V.; Cardoso Trabuco, M.; Peruzzi, G.; Boffi, A.; Testi, C.; Vallone, B.; Celeste Montemiglio, L.; Georges, A. D.; Calisti, L.; Benni, I.; Bonamore, A.; Baiocco, P., Humanized archaeal ferritin as a tool for cell targeted delivery. *Nanoscale* 2017, 9 (2), 647-655.
 83. Macone, A.; Masciarelli, S.; Palombarini, F.; Quaglio, D.; Boffi, A.; Trabuco, M. C.; Baiocco, P.; Fazi, F.; Bonamore, A., Ferritin nanovehicle for targeted delivery of cytochrome C to cancer cells. *Scientific Reports* 2019, 9 (1), 11749.
 84. Wang, Z.; Dai, Y.; Wang, Z.; Jacobson, O.; Zhang, F.; Yung, B. C.; Zhang, P.; Gao, H.; Niu, G.; Liu, G.; Chen, X., Metal ion assisted interface re-engineering of a ferritin nanocage for enhanced biofunctions and cancer therapy. *Nanoscale* 2018, 10 (3), 1135-1144.
 85. Jeon, J. O.; Kim, S.; Choi, E.; Shin, K.; Cha, K.; So, I. S.; Kim, S. J.; Jun, E.; Kim, D.; Ahn, H. J.; Lee, B. H., Designed nanocage displaying ligand-specific Peptide bunches for high affinity and biological activity. *ACS Nano* 2013, 7 (9), 7462-7471.
 86. Falvo, E.; Tremante, E.; Arcovito, A.; Papi, M.; Elad, N.; Boffi, A.; Morea, V.; Conti, G.; Toffoli, G.; Fracasso, G.; Giacomini, P.; Ceci, P., Improved doxorubicin encapsulation and pharmacokinetics of ferritin-fusion protein nanocarriers bearing proline, serine, and alanine elements. *Biomacromolecules* 2016, 17 (2), 514-522.
 87. Fracasso, G.; Falvo, E.; Colotti, G.; Fazi, F.; Ingegnere, T.; Amalfitano, A.; Doglietto, G. B.; Alfieri, S.; Boffi, A.; Morea, V.; Conti, G.; Tremante, E.; Giacomini, P.; Arcovito, A.; Ceci, P., Selective delivery of doxorubicin by novel stimuli-sensitive nano-ferritins overcomes tumor refractoriness. *Journal of Controlled Release* 2016, 239, 10-18.
 88. Falvo, E.; Malagrino, F.; Arcovito, A.; Fazi, F.; Colotti, G.; Tremante, E.; Di Micco, P.; Braca, A.; Opri, R.; Giuffrè, A.; Fracasso, G.; Ceci, P., The presence of glutamate residues on the PAS sequence of the stimuli-sensitive nano-ferritin improves in vivo biodistribution and mitoxantrone encapsulation homogeneity. *Journal of Controlled Release* 2018, 275, 177-185.
 89. Han, J. A.; Kang, Y. J.; Shin, C.; Ra, J. S.; Shin, H. H.; Hong, S. Y.; Do, Y.; Kang, S., Ferritin protein cage nanoparticles as versatile antigen delivery nanoplatfoms for dendritic cell (DC)-based vaccine development. *Nanomedicine* 2014, 10 (3), 561-569.
 90. Lee, B. R.; Ko, H. K.; Ryu, J. H.; Ahn, K. Y.; Lee, Y. H.; Oh, S. J.; Na, J. H.; Kim, T. W.; Byun, Y.; Kwon, I. C.; Kim, K.; Lee, J., Engineered human ferritin nanoparticles for direct delivery of tumor antigens to lymph node and cancer immunotherapy. *Scientific Reports* 2016, 6, 35182.

91. Kim, S. E.; Ahn, K. Y.; Park, J. S.; Kim, K. R.; Lee, K. E.; Han, S. S.; Lee, J., Fluorescent ferritin nanoparticles and application to the aptamer sensor. *Analytical Chemistry* 2011, 83 (15), 5834-5843.
92. Cioloboc, D.; Kurtz, D. M., Jr., Targeted cancer cell delivery of arsenate as a reductively activated prodrug. *Journal of Biological Inorganic Chemistry* 2020, 1-9.
93. Choi, S. H.; Choi, K.; Kwon, I. C.; Ahn, H. J., The incorporation of GALA peptide into a protein cage for an acid-inducible molecular switch. *Biomaterials* 2010, 31 (19), 5191-5198.
94. Fan, K.; Jiang, B.; Guan, Z.; He, J.; Yang, D.; Xie, N.; Nie, G.; Xie, C.; Yan, X., Fenobody: a ferritin-displayed nanobody with high apparent affinity and half-life extension. *Analytical Chemistry* 2018, 90 (9), 5671-5677.
95. Kim, S.; Jeon, J. O.; Jun, E.; Jee, J.; Jung, H. K.; Lee, B. H.; Kim, I. S.; Kim, S., Designing peptide bunches on nanocage for bispecific or superaffinity targeting. *Biomacromolecules* 2016, 17 (3), 1150-1159.
96. Lee, N. K.; Lee, E. J.; Kim, S.; Nam, G. H.; Kih, M.; Hong, Y.; Jeong, C.; Yang, Y.; Byun, Y.; Kim, I. S., Ferritin nanocage with intrinsically disordered proteins and affibody: A platform for tumor targeting with extended pharmacokinetics. *Journal of Controlled Release* 2017, 267, 172-180.
97. Li, X.; Qiu, L.; Zhu, P.; Tao, X.; Imanaka, T.; Zhao, J.; Huang, Y.; Tu, Y.; Cao, X., Epidermal growth factor-ferritin H-chain protein nanoparticles for tumor active targeting. *Small* 2012, 8 (16), 2505-2514.
98. Uchida, M.; Flenniken, M. L.; Allen, M.; Willits, D. A.; Crowley, B. E.; Brumfield, S.; Willis, A. F.; Jackiw, L.; Jutila, M.; Young, M. J., Targeting of cancer cells with ferrimagnetic ferritin cage nanoparticles. *Journal of the American Chemical Society* 2006, 128 (51), 16626-16633.
99. Guo, J.; Xu, N.; Yao, Y.; Lin, J.; Li, R.; Li, J. W., Efficient expression of recombinant human heavy chain ferritin (FTH1) with modified peptides. *Protein Expression and Purification* 2017, 131, 101-108.
100. Jaaskelainen, A.; Soukka, T.; Lamminmaki, U.; Korpimaki, T.; Virta, M., Development of a denaturation/renaturation-based production process for ferritin nanoparticles. *Biotechnology and Bioengineering* 2009, 102 (4), 1012-1024.
101. Subhadarshane, B.; Mohanty, A.; Jagdev, M. K.; Vasudevan, D.; Behera, R. K., Surface charge dependent separation of modified and hybrid ferritin in native PAGE: Impact of lysine 104. *BBA-Proteins and Proteomics* 2017, 1865 (10), 1267-1273.
102. Santambrogio, P.; Pinto, P.; Levi, S.; Cozzi, A.; Rovida, E.; Albertini, A.; Artymiuk, P.; Harrison, P. M.; Arosio, P., Effects of modifications near the 2-, 3-and 4-fold symmetry axes on human ferritin renaturation. *Biochemical Journal* 1997, 322 (2), 461-468.
103. Lin, X.; Xie, J.; Zhu, L.; Lee, S.; Niu, G.; Ma, Y.; Kim, K.; Chen, X., Hybrid ferritin nanoparticles as activatable probes for tumor imaging. *Angewandte Chemie International Edition* 2011, 50 (7), 1569-1572.
104. Ji, T.; Zhao, Y.; Wang, J.; Zheng, X.; Tian, Y.; Zhao, Y.; Nie, G., Tumor fibroblast specific activation of a hybrid ferritin nanocage-based optical probe for tumor microenvironment imaging. *Small* 2013, 9 (14), 2427-2431.
105. Lin, X.; Xie, J.; Niu, G.; Zhang, F.; Gao, H.; Yang, M.; Quan, Q.; Aronova, M. A.; Zhang, G.; Lee, S.; Leapman, R.; Chen, X., Chimeric ferritin nanocages for multiple function loading and multimodal imaging. *Nano Letters* 2011, 11 (2), 814-819.

106. Ahn, B.; Lee, S. G.; Yoon, H. R.; Lee, J. M.; Oh, H. J.; Kim, H. M.; Jung, Y., Four-fold channel-nicked human ferritin nanocages for active drug loading and pH-responsive drug release. *Angewandte Chemie International Edition* 2018, 57 (11), 2909-2913.
107. Boumaiza, M.; Carmona, F.; Poli, M.; Asperti, M.; Gianoncelli, A.; Bertuzzi, M.; Ruzzenenti, P.; Arosio, P.; Marzouki, M. N., Production and characterization of functional recombinant hybrid heteropolymers of camel hepcidin and human ferritin H and L chains. *Protein Engineering, Design & Selection* 2017, 30 (2), 77-84.
108. Lee, J. H.; Seo, H. S.; Song, J. A.; Kwon, K. C.; Lee, E. J.; Kim, H. J.; Lee, E. B.; Cha, Y. J.; Lee, J., Proteinticle engineering for accurate 3D diagnosis. *ACS Nano* 2013, 7 (12), 10879-10886.
109. Jin, P.; Sha, R.; Zhang, Y.; Liu, L.; Bian, Y.; Qian, J.; Qian, J.; Lin, J.; Ishimwe, N.; Hu, Y.; Zhang, W.; Liu, Y.; Yin, S.; Ren, L.; Wen, L. P., Blood circulation-prolonging peptides for engineered nanoparticles identified via phage display. *Nano Letters* 2019, 19 (3), 1467-1478.
110. Liang, M.; Tan, H.; Zhou, J.; Wang, T.; Duan, D.; Fan, K.; He, J.; Cheng, D.; Shi, H.; Choi, H. S.; Yan, X., Bioengineered H-ferritin nanocages for quantitative imaging of vulnerable plaques in atherosclerosis. *ACS Nano* 2018, 12 (9), 9300-9308.
111. Gumulec, J.; Fojtu, M.; Raudenska, M.; Sztalmachova, M.; Skotakova, A.; Vlachova, J.; Skalickova, S.; Nejd, L.; Kopel, P.; Knopfova, L.; Adam, V.; Kizek, R.; Stiborova, M.; Babula, P.; Masarik, M., Modulation of induced cytotoxicity of doxorubicin by using apoferritin and liposomal cages. *International Journal of Molecular Sciences* 2014, 15 (12), 22960-22977.
112. Pandolfi, L.; Bellini, M.; Vanna, R.; Morasso, C.; Zago, A.; Carcano, S.; Avvakumova, S.; Bertolini, J. A.; Rizzuto, M. A.; Colombo, M.; Prosperi, D., H-ferritin enriches the curcumin uptake and improves the therapeutic efficacy in triple negative breast cancer cells. *Biomacromolecules* 2017, 18 (10), 3318-3330.
113. Dostalova, S.; Polanska, H.; Svobodova, M.; Balvan, J.; Krystofova, O.; Haddad, Y.; Krizkova, S.; Masarik, M.; Eckschlager, T.; Stiborova, M.; Heger, Z.; Adam, V., Prostate-specific membrane antigen-targeted site-directed antibody-conjugated apoferritin nanovehicle favorably influences *in vivo* side effects of doxorubicin. *Scientific Reports* 2018, 8 (1), 8867.
114. Zhang, Y.; Dong, Y.; Li, X.; Wang, F., Proanthocyanidin encapsulated in ferritin enhances its cellular absorption and antioxidant activity. *Journal of Agricultural and Food Chemistry* 2019, 67 (41), 11498-11507.
115. Bonizzi, A.; Truffi, M.; Sevieri, M.; Allevi, R.; Sitia, L.; Ottria, R.; Sorrentino, L.; Sottani, C.; Negri, S.; Grignani, E.; Mazzucchelli, S.; Corsi, F., Everolimus nanoformulation in biological nanoparticles increases drug responsiveness in resistant and low-responsive breast cancer cell lines. *Pharmaceutics* 2019, 11 (8), 384.
116. Tilocca, B.; Soggiu, A.; Sanguinetti, M.; Musella, V.; Britti, D.; Bonizzi, L.; Urbani, A.; Roncada, P., Comparative computational analysis of SARS-CoV-2 nucleocapsid protein epitopes in taxonomically related coronaviruses. *Microbes and Infection* 2020, 22 (4-5), 188-194.
117. Demirci, M. D. S.; Adan, A., Computational analysis of microRNA-mediated interactions in SARS-CoV-2 infection. *PeerJ* 2020, 8, e9369.
118. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 2015, 1, 19-25.
119. Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y., CHARMM-GUI input generator for NAMD,

- GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Journal of Chemical Theory and Computation* 2016, 12 (1), 405-413.
120. Holcomb, J.; Spellmon, N.; Zhang, Y.; Doughan, M.; Li, C.; Yang, Z., Protein crystallization: eluding the bottleneck of X-ray crystallography. *AIMS Biophysics* 2017, 4 (4), 557.
 121. Kay, L. E., NMR studies of protein structure and dynamics. *Journal of Magnetic Resonance* 2011, 213 (2), 477-491.
 122. Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpí, J. L., Molecular dynamics simulations: advances and applications. *Advances and Applications in Bioinformatics and Chemistry: AABC* 2015, 8, 37.
 123. Dorn, M.; e Silva, M. B.; Buriol, L. S.; Lamb, L. C., Three-dimensional protein structure prediction: Methods and computational strategies. *Computational Biology and Chemistry* 2014, 53, 251-276.
 124. Maupetit, J.; Derreumaux, P.; Tufféry, P., A fast method for large scale De Novo peptide and miniprotein structure prediction. *Journal of Computational Chemistry* 2010, 31 (4), 726-738.
 125. Thevenet, P.; Shen, Y.; Maupetit, J.; Guyon, F.; Derreumaux, P.; Tuffery, P., PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Research* 2012, 40 (W1), W288-W293.
 126. Maupetit, J.; Derreumaux, P.; Tuffery, P., PEP-FOLD: an online resource for de novo peptide structure prediction. *Nucleic Acids Research* 2009, 37 (suppl_2), W498-W503.
 127. Kukic, P.; Kannan, A.; Dijkstra, M. J.; Abeln, S.; Camilloni, C.; Vendruscolo, M., Mapping the protein fold universe using the camtube force field in molecular dynamics simulations. *PLoS Computational Biology* 2015, 11 (10), e1004435.
 128. Rose, P. W.; Prlić, A.; Altunkaya, A.; Bi, C.; Bradley, A. R.; Christie, C. H.; Costanzo, L. D.; Duarte, J. M.; Dutta, S.; Feng, Z., The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Research* 2016, gkw1000.
 129. Panchal, N. K.; Bhale, A.; Verma, V. K.; Beevi, S. S., Computational and molecular dynamics simulation approach to analyze the impact of XPD gene mutation on protein stability and function. *Molecular Simulation* 2020, 46 (15), 1200-1219.
 130. Bjelkmar, P.; Larsson, P.; Cuendet, M. A.; Hess, B.; Lindahl, E., Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *Journal of Chemical Theory and Computation* 2010, 6 (2), 459-466.
 131. Singh, A.; Munshi, S.; Raghavan, V., Effect of external electric field stress on gliadin protein conformation. *Proteomes* 2013, 1 (2), 25-39.
 132. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 2004, 25 (13), 1605-1612.
 133. Gao, Y.-D.; Huang, J. F., An extension strategy of Discovery Studio 2.0 for non-bonded interaction energy automatic calculation at the residue level, 2011, 32(3): 262-266.
 134. Zhang, L.; Lua, L. H.; Middelberg, A. P.; Sun, Y.; Connors, N. K., Biomolecular engineering of virus-like particles aided by computational chemistry methods. *Chemical Society Reviews* 2015, 44 (23), 8608-8618.
 135. Maity, B.; Li, Z.; Niwase, K.; Ganser, C.; Furuta, T.; Uchihashi, T.; Lu, D.; Ueno, T., Single-molecule level dynamic observation of disassembly of the apo-ferritin cage in solution. *Physical Chemistry Chemical Physics* 2020, 22 (33), 18562-18572.

136. Zhang, L.; Tang, R.; Bai, S.; Connors, N. K.; Lua, L. H.; Chuan, Y. P.; Middelberg, A. P.; Sun, Y., Molecular energetics in the capsomere of virus-like particle revealed by molecular dynamics simulations. *The Journal of Physical Chemistry B* 2013, 117 (18), 5411-5421.
137. Jung, C., Insight into protein structure and protein–ligand recognition by Fourier transform infrared spectroscopy. *Journal of Molecular Recognition* 2000, 13 (6), 325-351.
138. Pacholarz, K. J.; Garlish, R. A.; Taylor, R. J.; Barran, P. E., Mass spectrometry based tools to investigate protein–ligand interactions for drug discovery. *Chemical Society Reviews* 2012, 41 (11), 4335-4355.
139. Patching, S. G., Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 2014, 1838 (1), 43-55.
140. Siligardi, G.; Hussain, R.; Patching, S. G.; Phillips-Jones, M. K., Ligand-and drug-binding studies of membrane proteins revealed through circular dichroism spectroscopy. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 2014, 1838 (1), 34-42.
141. Nitsche, C.; Otting, G., NMR studies of ligand binding. *Current Opinion in Structural Biology* 2018, 48, 16-22.
142. Chatterjee, T.; Pal, A.; Dey, S.; Chatterjee, B. K.; Chakrabarti, P., Interaction of virstatin with human serum albumin: spectroscopic analysis and molecular modeling. *PloS One* 2012, 7 (5), e37468.
143. Prieto-Martínez, F. D.; Arciniega, M.; Medina-Franco, J. L., Molecular docking: current advances and challenges. *TIP. Revista Especializada En Ciencias Químico-biológicas* 2018, 21.
144. Pagadala, N. S.; Syed, K.; Tuszynski, J., Software for molecular docking: a review. *Biophysical Reviews* 2017, 9 (2), 91-102.
145. Subramanian, V.; Evans, D. G., A molecular dynamics and computational study of ligand docking and electron transfer in ferritins. *The Journal of Physical Chemistry B* 2012, 116 (31), 9287-9302.
146. Shahwan, M.; Khan, M. S.; Husain, F. M.; Shamsi, A., Understanding binding between donepezil and human ferritin: molecular docking and molecular dynamics simulation approach. *Journal of Biomolecular Structure and Dynamics* 2020, 1-9.
147. Agrawal, P.; Singh, H.; Srivastava, H. K.; Singh, S.; Kishore, G.; Raghava, G. P., Benchmarking of different molecular docking methods for protein-peptide docking. *BMC Bioinformatics* 2019, 19 (13), 426.
148. Ciemny, M.; Kurcinski, M.; Kamel, K.; Kolinski, A.; Alam, N.; Schueler-Furman, O.; Kmiecik, S., Protein–peptide docking: opportunities and challenges. *Drug Discovery Today* 2018, 23 (8), 1530-1537.
149. De Vries, S. J.; Van Dijk, M.; Bonvin, A. M., The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols* 2010, 5 (5), 883.
150. Deplazes, E.; Davies, J.; Bonvin, A. M.; King, G. F.; Mark, A. E., Combination of ambiguous and unambiguous data in the restraint-driven docking of flexible peptides with HADDOCK: The binding of the spider toxin PcTx1 to the acid sensing ion channel (ASIC) 1a. *Journal of Chemical Information and Modeling* 2016, 56 (1), 127-138.
151. Koukos, P. I.; Faro, I.; van Noort, C. W.; Bonvin, A. M., A membrane protein complex docking benchmark. *Journal of Molecular Biology* 2018, 430 (24), 5246-5256.
152. Volkov, A. N.; Ferrari, D.; Worrall, J. A.; Bonvin, A. M.; Ubbink, M., The orientations of cytochrome c in the highly dynamic complex with cytochrome b5 visualized by NMR and docking using HADDOCK. *Protein Science* 2005, 14 (3), 799-811.

Chapter 3

MOLECULE DESIGN, EXPRESSION OF FUNCTIONALISED HUMAN HEAVY-CHAIN FERRITINS AND COMPUTATIONAL INVESTIGATION OF SELF-ASSEMBLING

3.1 Introduction

As was reviewed in **Chapter 2**, human heavy-chain ferritin (HF_n), because of its significant thermal and chemical stability, has become of interest for protein nanocages for anti-tumour drug delivery.¹ Researchers have developed various approaches to load anti-tumour drugs into HF_n and explored their anti-tumour efficacy.²⁻⁴ Chemical conjugation and gene technology have been applied to broaden HF_n application and infuse extra functions.^{5, 6}

Because of the greater site-specificity and ordered presence manner of intended alterations in gene technology, fusion is the most widely adopted functionalisation approach. Common functional peptide insertion sites in fusion include subunit N-terminal, C terminal and DE turn of HF_n subunit.⁷⁻⁹

However, insoluble expression and expression level reduction of functionalised HF_ns have been detected. This is because the selection of functional peptide is critical. Long and structured or cysteine containing peptide is more likely to end up with greater insoluble expression level. For example, 2 functionalised HF_ns with either a cysteine containing peptide, or a single-chain antibody fragment inserted at N terminal were both expressed as inclusion bodies (IBs).^{10, 11} Denaturation and reassembly of them were required and led to low protein yields.

In addition, inserted functional peptide may pose an impact on HF_n conformation and biochemical property, and these changes depend on the location and the biochemical property of the inserted peptide. A C-terminal inserted GALA peptide, the structure of which turns from random coil to helix in neutral pH, has equipped ferritin with a neutral pH disassembly feature.¹² The study of the functionalised HF_ns structure and inserted peptide impact on HF_n can shed light on the findings in functionalised HF_ns characterisations, deepen the understanding of HF_n and the inserted peptide, and provide valuable information for the design of new functionalised HF_n. It has been reviewed in **Chapter 2**, molecular dynamic (MD) simulation and protein-peptide molecular docking are powerful techniques to predict functionalised protein conformation and study the impact of foreign peptide on the original protein. These are potentially feasible to study functionalised HF_ns.

In this thesis, HF_n was explored in 2 different functionalisation strategies to improve its anti-tumour performance as a drug delivery platform through fusion. The functional peptides selected contain no cysteine residues and are either short or non-structured for the reduction of the impact on HF_n conformation after insertion.

The first strategy is to equip HF_n with tumour targeting ability and extended half-life and then load small molecule chemotherapeutic. To achieve this, 2 functional peptides were inserted into HF_n subunit C-terminus, and 3 constructs were designed. One functional peptide is a tetrapeptide named RGDK. It belongs to tumour penetration Peptide (TPP) and possesses dual functions as it enhances drug tumour delivery and drug distribution inside whole tumour tissue instead of only tumour cells alongside tumour vessels.¹³ Its –COOH needs to be exposed at C-terminal of the new protein for its function.¹⁴ This is the reason for the selection of HF_n subunit C-terminus as fusion site in this strategy. Another inserted peptide called PAS consists of repetitive P, A and S residues. It was designed by Schlapschy and aimed to extend the retention time of drug in blood circulation.¹⁵ In a previous study, Kim has fused 40 aa and 75 aa PAS peptide to human ferritin subunit at N-terminal and increased half-life in circulation by 4.6 and 5.8 times.¹⁶ Both PAS and RGDK themselves are random coils and connected to HF_n subunit through a flexible 15 aa linker to reduce the impact of peptide insertion on HF_n conformation. Three functionalised HF_ns all contain PAS peptide and were named as HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK, respectively (refer to **Figure 3.1**). GFLG and PLGLAG are 2 enzyme responsive sites which can be selectively cleaved by cathepsin B and matrix metalloproteinase-2 (MMP-2), respectively. Cathepsin B and MMP-2 expression is upregulated in tumour cells.^{17, 18} With these enzyme cleavable sites, functional peptides will be dissociated from HF_n. HF_n was constructed as a control for 3 PAS functionalised HF_ns in this strategy.

The second strategy is to fuse a peptide drug called pro-apoptotic peptide (P) to HF_n. P is a small helix consisting of 14 aa (KLAKLAKKLAKLAK). Its design draws upon cationic antimicrobial peptides which could disrupt cell membranes with positive charges. Because tumour cell overexpresses negatively charged anionic phospholipids, P has demonstrated a selective cytotoxicity against tumour cells.¹⁹⁻²¹ In the design of P functionalised HF_ns, C-terminal and N-terminal fusion were both employed. In C-terminal fusion, short HF_n (sHF_n), which is E-helix truncated HF_n, was used instead of HF_n, because compared with HF_n, it

allows for larger foreign peptides in C-terminus fusion. A 15 aa flexible linker was used in C-terminal fusion. In N-terminal fusion, a 30 aa-long flexible linkers were employed to sperate P and HF_n. Three P functionalised HF_ns were P-HF_n, sHF_n-P and sHF_n-P-RGDK (refer to **Figure 3.1**). RGDK in sHF_n-P-RGDK was for an enhanced tumour targeting ability. In P functionalisation strategy, sHF_n was constructed as a control for sHF_n-P and sHF_n-P-RGDK, and HF_n was the control for P-HF_n.

In total, 8 constructs, 2 controls (HF_n and sHF_n), 3 PAS functionalised HF_ns and 3 P functionalised HF_ns were designed. In this chapter, plasmids of sHF_n, sHF_n-P and sHF_n-P-RGDK were synthesized in laboratory and the other 5 genes were purchased. All 8 corresponding plasmids were transformed into *E. coli* BL21 (DE3) through heat-shock transformation. Expression of proteins were compared in shake-flask fermentation. A 5L fermenter fermentation was used to express control HF_n. Transmission electron microscopy (TEM), size exclusion chromatography (SEC) and reducing 12 % sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) were used to detect the expression forms and levels. Computational analysis including molecular dynamic simulation and protein-peptide molecular docking was applied to understand the mechanism of self-assembly in functionalised HF_ns.

3.2 Materials and methods

3.2.1 Materials

Primers for PCR were synthesised by BGI Company (China). Nde I, Xho I restriction endonucleases, T4 ligation kit were purchased from NEB (UK). Agarose gel DNA extract kit, plasmid extract kit and *E. coli* BL21 (DE3) competent cells were bought from Tiangen Biological Company (China). Plasmid sequencing service was offered by BGI Company (China). Tryptone, yeast extract, isopropyl β-D-thiogalactoside (IPTG) and kanamycin were purchased from Thermo Fisher Scientific (USA). All other reagents of analytical grade were purchased from Chem-supply (Australia) and Sigma-Aldrich (USA). Milli Q water was produced from a Millipore purification system (Merck, USA) and used throughout the whole process.

3.2.2 HFn-based protein sequence design

There are 8 HFn-based proteins in this thesis. Plasmid construction schematic is shown in **Figure 3.1**. Primary sequences are shown in **Table A1 (Appendix A)**. pET30a plasmid was applied as target protein gene vector. Genes of HFn, 3 PAS functionalised HFns (HFn-PAS, HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK) and P-HFn, were synthesised and inserted into pET30a plasmid between Nde I and Xho I restriction enzyme sites by BGI Company (China). P-HFn was constructed by adding pro-apoptotic peptide (KLAKLAKKLAKLAK) to the N-terminal HFn subunit through a 30 aa flexible linker (GGGSGGGTGGGSGGG GGGSGGGTGGGSGGG). HFn-PAS was constructed by fusing GFLG ASPAAPAPASPAAPAPSAPAASPAAPAPASPAAPAPSAPA GGSGG to HFn subunit C-terminus through a flexible 15 residues linker (GGGSGGGTGGGSGGG). 40 underlined residues make up PAS peptide and GFLG is the Cathepsin B cleavable site. HFn-GFLG-PAS-RGDK has 4 extra C-terminus residues, RGDK, compared with HFn-PAS. In HFn-PLGLAG-PAS-RGDK, enzyme-cleavable site in HFn-GFLG-PAS-RGDK is replaced with a 6 residue MMP-2 cleavable site PLGLAG.

Three types of plasmids bearing gene of sHFn, sHFn-P and sHFn-P-RGDK, respectively, were constructed in the lab. In sHFn-P, a 15 aa flexible linker (GGSGG GGSGG GGSGG) was adopted as the spacer for sHFn subunit C-terminal and pro-apoptotic peptide. Extra 7 residues of GGG RGDK were added to sHFn-P C-terminal to construct sHFn-P-RGDK.

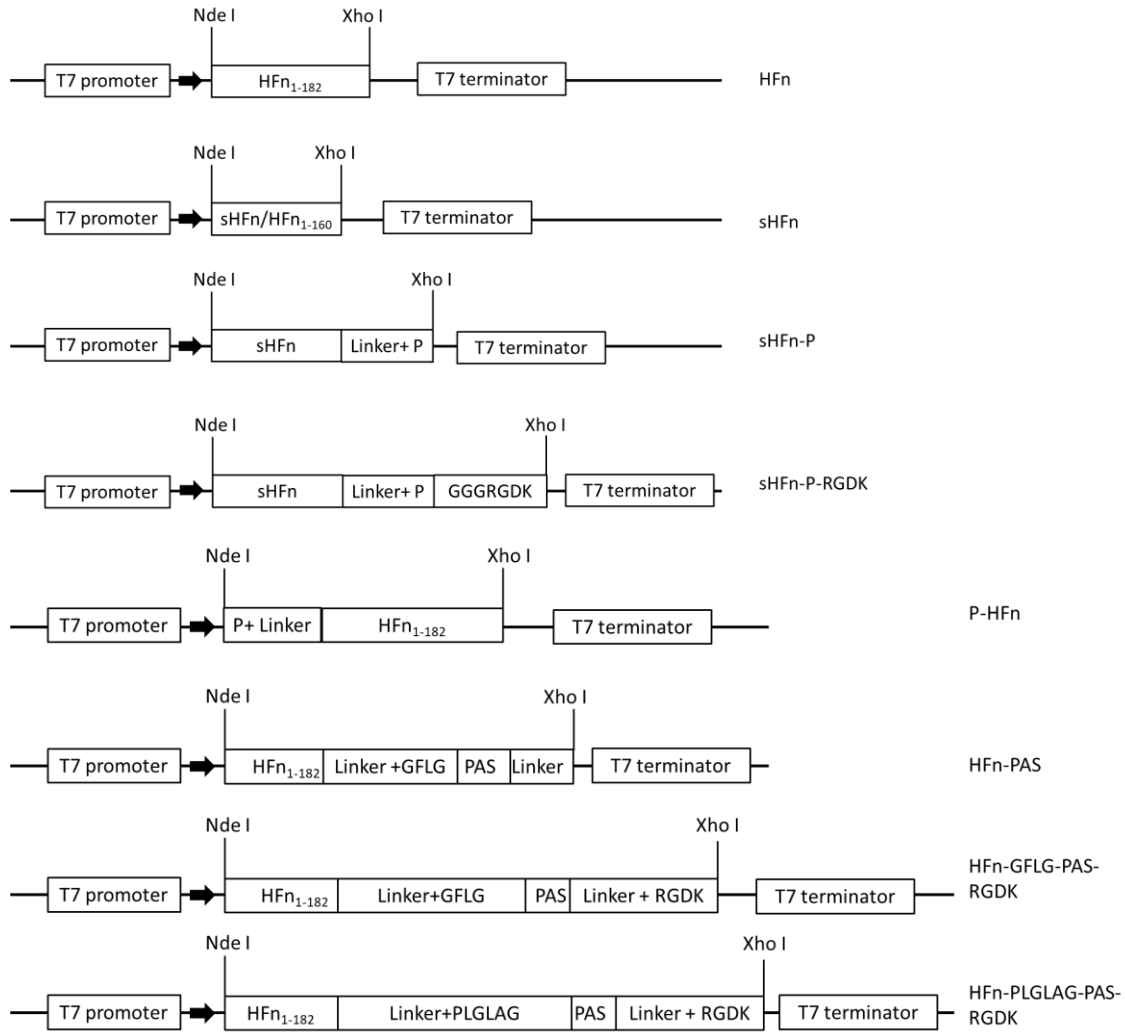


Figure 3.1 Schematic of expression cassettes of the pET30a plasmids harbouring genes for all HFn-based proteins.

3.2.3 Plasmid construction for sHF_n, sHF_n-P and sHF_n-P-RGDK

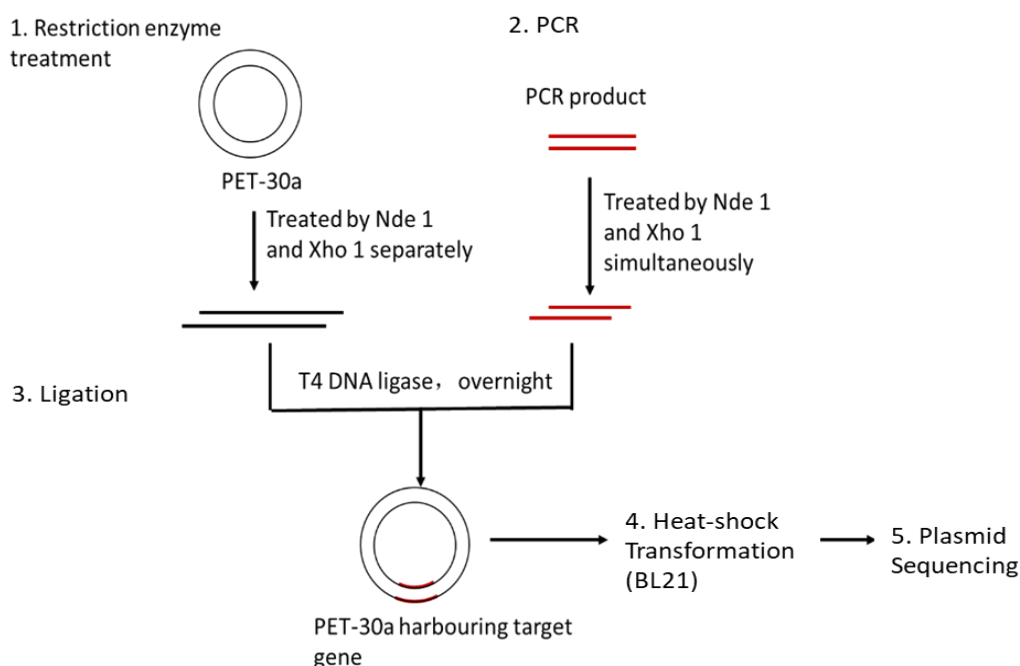


Figure 3.2 Schematic of *E. coli* construction.

Figure 3.2 illustrates the process adopted for the construction of bacteria in the laboratory. It consists of 5 steps, 1) the enzyme treatment of plasmid vector pET30a to expose sticky ends. Nde I and Xho I endonuclease were applied. 2) Polymerase chain reaction (PCR) for the amplification and synthesis of genes encoding target proteins (sHF_n, sHF_n-P, sHF_n-P-RGDK). In PCR, template genes and primers were synthesised by BGI Company (China). Forward and reverse primers are listed in **Table 3.1**. PCR products were then treated with the same endonucleases used for plasmid vector digestion to generate the same sticky ends. 3) Ligation of endonuclease treated vector and PCR products to obtain new plasmids containing target genes. 4) Transformation of new plasmids into competent cells (BL21 DE3). 5) Sequencing of the plasmids transformed into *E. coli* to confirm the gene sequences and insertion location.

Table 3.1 Primer sequences for PCR.

Target gene	Forward primer sequence	Reverse primer sequence
sHF _n	5'GGAATTCATATGACTACTGCTTCT ACC 3'	5'CCGCTCGAGTTAGGCGCCCATTTTA CG 3'
sHF _n -P	5'GGAATTCATATGACTACTGCTTCT ACC 3'	5'CCGCTCGAGTCATTTTGCCAGCTTTG CCAGTTTTTTGGC 3'
sHF _n -P- RGDK	5'GGAATTCATATGACTACTGCTTCT ACC 3'	5'CCGCTCGAGTTATTTGTCACCGCGA CC 3'

3.2.3.1 *Nde I* and *Xho I* digestion of plasmid vector pET30a

30 μL of pET30a plasmids ($5 \mu\text{g mL}^{-1}$ in Mili Q water) was mixed with $1 \mu\text{L}$ *Nde I* endonuclease and $3 \mu\text{L}$ $10\times$ H buffer (NBE, UK). The mixture was incubated at 37°C for 12 h. 1.0 % agarose electrophoresis was conducted to analyse and separate the *Nde I* digested plasmid. Briefly, 0.3g agarose was added to 30 mL Tris-acetate-EDTA (TAE), pH 8.3 buffer and melted using microwave. When the temperature of agarose solution cooled to around 50°C , $1 \mu\text{L}$ of DNA staining dye (Sigma-Aldrich, USA) was added and the mixed thoroughly. Agarose solution was poured into a gel mould to form agarose gel. Undigested pET30a and *Nde I* digested sample were mixed with $10\times$ DNA loading buffer with a volume ratio of 9:1. Two samples and DNA marker D 15,000 were loaded to the gel and the electrophoresis was conducted in a horizontal chamber. Voltage was 90 V and the electrophoresis lasted 1 h. Gel image was taken in Gel Documentation System G:BOX (Syngene , Australia). The target *Nde I* digested plasmid was extracted using a DNA extract kit (Tiangen Biological Company, China) from the agarose gel according to the manual included in the kit.

Extracted *Nde I* digested plasmid was digested with *Xho I* at 37°C for 12 h. 1 % agarose gel electrophoresis and DNA extraction were conducted to obtain *Nde I* and *Xho I* digested pET30a for the following ligation.

3.2.3.2 PCR

PCR was conducted for synthesis and amplification of 4 genes encoding sHFn, HFn, sHFn-P and sHFn-P-RGDK on an automated thermal cycler (Thermo Fisher Scientific, USA). $25 \mu\text{L}$ PCR mixtures were prepared for each gene synthesis. **Table 3.2** lists the reagents for PCR. Reagents were added to PCR tubes in the order from the top to the bottom of the table.

Table 3.2 Reagents for PCR.

Reagent	Volume
Mili Q water	$14.5 \mu\text{L}$
$5\times$ Taq buffer	$5 \mu\text{L}$
dNTP (2.5 mM)	$2 \mu\text{L}$
Forward primer (10 μM)	$1 \mu\text{L}$
Reverse primer (10 μM)	$1 \mu\text{L}$
Template (10 μM)	$1 \mu\text{L}$
Taq DNA polymerase ($0.5 \text{ U } \mu\text{L}^{-1}$)	$0.5 \mu\text{L}$

PCR consisted of 3 steps and the settings were as follows: 1) denaturation of DNA template, 94 °C 5 min; 2) 35 cycles of denaturation, annealing and elongation. In each cycle, denaturation was set at 94 °C for 40 s. Annealing was set at a proper annealing temperature for 40 s and elongation was set at 72 °C, 50 s. Annealing temperatures were selected within the range from T_m value -5 °C of forward to T_m value -5 °C of reverse primers. Primer T_m value was estimated using Annhyb software based on nucleotide sequence. Selected annealing temperatures for 3 PCR were 58 °C (sHF_n), 65 °C (sHF_n-P) and 62 °C (sHF_n-P-RGDK); 3) Elongation, set at 72 °C for 10 min to finish the possibly incomplete elongation in step 2.

On completion of PCR, products underwent 1.2 % agarose gel electrophoresis. DNA marker D2000 was used to help analyse if the molecular weights (M_ws) of amplified genes were correct. DNA bands with molecular weights the same as in theory were extracted using agarose gel DNA extract kit (Tiangen Biological Company, China) following the manual in the kit.

3.2.3.3 PCR product restriction endonuclease digestion

15 µL gel extracted PCR products were incubated with 10 × H buffer 2 µL, Nde I 2 µL and Xho I 1 µL at 37 °C for 12 h. Digested products were analysed using 1.5 % or 1.2 % agarose gel electrophoresis and recovered by extraction from the gel.

3.2.3.4 Ligation

1.5 µL of endonuclease digested pET30a and 12 µL of digested PCR products were mixed with 7.5 µL T4 ligation mix (NBE, UK). The mixtures were incubated at 16 °C for 12 h. Products after ligation were kept in -20 °C until heat-shock transformation.

3.2.4 Heat-shock transformation and plasmid sequencing

1 µL BGI synthesised plasmids (HF_n, P-HF_n, HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK) (50 ng mL⁻¹) or 10 µL ligation products (plasmids for sHF_n, sHF_n-P and sHF_n-P-RGDK) were incubated with 25 µL competent *E. coli* BL21 (DE3) (Tiangen Biological Company, China) in ice bath for 20 min to starve the cells in laminar flow cabinet (Thermo Fisher Scientific, USA). Following that, mixtures were shocked at 42 °C for 90 s and then put in ice bath (0 °C) for 2 min. 800 µL antibiotic free Luria-Bertani (LB) medium was added to each mixture and cells were cultured at 37 °C for 1 h at a rotation speed of 150 rpm

in New Brunswick™ Innova® 40/40R Shakers (Thermo Fisher Scientific, USA). After that, cultured cells were centrifuged at 1,000 rpm, 4 °C for 3 min using a centrifuge (Eppendorf, USA) and 700 µL of supernatant was discarded. Cells were resuspended using the remaining medium and seeded onto solid LB medium plates containing 100 µg mL⁻¹ kanamycin and 1 % agar. Plates were incubated upside down in oven overnight at 37 °C. Colonies grown on plates the next day were sent to BGI Company (China) for plasmid sequencing to check if target genes were correctly synthesised and inserted between Nde I and Xho I restriction sites in pET30a plasmids.

3.2.5 Shake flask and fermenter fermentation of different *E. coli* strains

3.2.5.1 Shake flask fermentation

37 °C induction shake flask fermentation was performed for 8 bacterial strains expressing HF_n, sHF_n and functionalised HF_ns. Firstly, 100 µL *E. coli* strain was seeded in a conical flask containing 50 mL Luria-Bertani (LB) medium and 100 µg mL⁻¹ kanamycin and cultured at 37 °C for 14 h under shaking at 200 rpm for pre-culture. 40 mL of the pre-cultured cells in the exponential growth phase were transferred evenly to 4 flasks each containing 500 mL fresh LB medium supplemented with 100 µg mL⁻¹ kanamycin and continued to grow in the same conditions for 4 h. The rest pre-cultured cells were mixed with autoclaved 50 % (v/v) glycerol with a ratio of 1:1 (v/v) and divided into 1 mL per tube for cryopreservation of bacterial strains for future use. IPTG was injected into the 4 flasks of 500 mL medium with a final concentration of 0.5 mM for protein expression induction. After 4 h, cells were harvested by centrifugation at 4,000 rpm, 4 °C for 20 min and wet biomass were measured.

A low temperature induction shake flask fermentation was conducted for P-HF_n expression. The first 2 steps were the same as in 37 °C induction shake flask fermentation. In the induction step, incubator temperature was set at 30 °C and the bacteria were induced for 12 h by 0.5 mM IPTG.

3.2.5.2 5 L fermenter fermentation

To scale up fermentation, a 5 L fermenter fed-batch fermentation was performed for bacteria expressing HF_n using BioFlo® 320 Vessels fermenter (Eppendorf, USA). The first 2 steps were exactly the same as in 37 °C induction shake flask fermentation. After 4 h culture in 500

mL medium, cells were transferred to fermenter in which there was 5 L medium supplemented with 100 $\mu\text{g mL}^{-1}$ kanamycin. Fermenter vessel temperature was controlled at 37 °C. Medium pH was set at 6.9, adjusted using phosphoric acid and ammonium hydroxide automatically using the cascade function in control tower. Compressed air was used as oxygen source. Dissolved oxygen range was set at 10 - 35 %. Agitation speed and gas flow rate were linked to dissolved oxygen level using 'cascade' function in control tower and adjusted automatically by the system. Sample of fermentation broth was taken out hourly to double check the medium pH and measure the optical density at 600 nm (OD_{600}). When OD_{600} reached above 5, 0.5 mM IPTG was injected into the fermenter and 1 L of supplement 10 times LB medium with glycerol was fed to the vessel at a constant rate. OD_{600} was still recorded hourly after induction. After 4.5 h, fermentation stopped, and cells were harvested. Wet biomass mass was recorded.

3.2.5.3 Cell disruption and Bradford assay

Harvested cell pellets were re-suspended using 20 mM phosphate buffer (PB), 2 mM EDTA, pH 7 buffer with a ratio of 1: 20 (mass of bacteria (g): volume of lysis buffer (mL)) and subjected to ultra-sonication or high-pressure homogenisation for cell disruption. Ultra-sonication was conducted with the settings as follows: 360 W power, 4-s on, 6-s off, 20 min in total using ultra-sonicator (Ningbo Keyi Technology, China). Resuspended bacteria were put in ice-bath and the temperature during ultra-sonication was kept under 30 °C.

In high pressure homogenisation, APV 2000 (SPX Flow, USA) was used. 500 mL of resuspended bacteria was fed into the homogeniser for 15 min homogenisation and the pressure was set at 400 bar. Watering cooling system was connected to homogeniser to keep the whole system below 30 °C.

Afterwards, bacterial lysates were separated into supernatants and sediments by centrifugation at 12,000 rpm, 30 min, 4 °C. The supernatants of cell lysate were kept for the following analysis. Supernatant protein concentrations were measured using Bradford assay. In Bradford assay, bovine serum albumin (BSA) (Sigma-Aldrich, USA) with different concentrations were used as references to build up standard correlation curve of 595 nm absorbance (A_{595}) to C_{protein} (mg mL^{-1}). Standard sample concentrations were 1, 0.75, 0.5, 0.25, 0.1 and 0 mg mL^{-1} . 3 mL of Bradford dye (Bio-Rad, USA) was mixed with 60 μL of standard protein samples in a 1 cm path length glass cuvette and the mixture was settled for 2 min to reach the protein-dye

interaction equilibrium. A_{595} was measured using a UV/vis spectrometer (VWR, Australia). Linear fitting of the A_{595} against protein concentration was conducted to obtain the standard curve. Bacterial lysate supernatants protein concentrations were diluted 10 times and then measured using the same procedure.

3.2.5.4 12 % reducing SDS-PAGE

Protein expression results were analysed using a 12 % reducing SDS-PAGE. Samples included 4 types of bacteria before IPTG induction, 4 types of bacteria after 4 h IPTG induction, 4 types of bacterial lysate supernatants and sediments. For bacterial and sediments samples, they were resuspended using Mili Q water. All samples were mixed with 5 × loading buffer (0.25 M Tris-HCl, 0.25 % bromophenol blue, 0.5 M DTT, 50 % glycerol, 10 % SDS, pH 6.8) and 10 % SDS with a volume ratio of 3: 1: 1, following by being heated at 100 °C for 20 min. Electrophoresis voltage was set at 90 V when samples were in 5 % stacking gel and 150 V when in separating gel. Electrophoresis was done using a vertical chamber (Bio-Rad, USA), and it was stopped when the bromophenol blue reached the bottom of the gel.

3.2.6 Transmission electron microscopy (TEM) and size exclusion chromatography (SEC) analysis of bacterial lysate supernatants

To check if HF_n-based proteins expressed in soluble form were self-assemblies, TEM analysis was adopted. 5 µL of bacterial lysate supernatant (0.5 mg mL⁻¹) was applied to carbon-coated, glow-discharged EM grids and negatively stained with 2 % uranyl acetate. Micrographs were recorded on a TVIPS F224HD 2k × 2k CCD camera (Gauting, Germany) using a FEI Tecnai G2 Spirit TEM (Eindhoven, The Netherlands). Voltage was set at 100 kV.

SEC determines target protein retention volume. By comparing functionalised HF_ns retention volume with that of HF_n, SEC results can preliminarily confirm if functionalised HF_ns are expressed in soluble monomers, assemblies or aggregates. Bacterial lysate supernatants for HF_n, sHF_n-PAS and sHF_n-P-RGDK were characterised using SEC. In SEC, Superose 6 10/300 gl column (GE Healthcare, USA) was connected to AKTA pure system (GE Healthcare, USA) and equilibrated with 20 mM PB, 0.15 M Na₂SO₄, pH 7. 0.5 mL bacterial lysate supernatant was loaded to the column. Absorbance at 260 and 280 nm was recorded. Fractions

were collected to undergo 12 % reducing SDS-PAGE analysis to confirm the target protein retention volume.

3.2.7 Computational analysis of the impact of pro-apoptotic peptide on HF_n and sHF_n subunit property and structure

Interactions between interface residues from different subunits are the basis for HF_n assembly. Residues located in helix A-D bundle form hydrogen bonds, salt bridges and hydrophobic interactions to form the assembly. All these non-covalent interactions are closely related to hydrophobicity and charge properties of relevant residues. In functionalised HF_ns, the insertion of foreign peptide can affect assembling through roughly 2 distinct ways: 1) the inserted peptide may directly interact with residues related to assembling and disrupt the original assembling related bond formation, or, 2) the peptide affects assembling in an indirect way through altering the assembling related area biochemical properties and conformation.

Computational analysis is a reliable and fast means to shed light on protein structure, properties and protein-ligand interactions. Because all 3 pro-apoptotic peptide functionalised HF_ns (sHF_n-P, sHF_n-P-RGDK and P-HF_n) were not able to form correct assemblies while sHF_n and HF_n controls were self-assemblies according to experiment results, it was very likely that the pro-apoptotic peptide was the main cause of assembly failure. Therefore, comparison of sHF_n-P and P-HF_n with pro-apoptotic peptide, HF_n and sHF_n were performed using computational methods to help understand the impact of P on HF_n assembly. In the case of sHF_n-P and P-HF_n, a 15 aa long linker and a 30 aa long linker were used to separate P with sHF_n C-terminal and HF_n N-terminal. The long length of linker leads to a reasonable assumption that HF_n or sHF_n subunit and P exist in their theoretically correct conformation. **Figure 3.3** illustrates the theoretical conformations of P-HF_n and sHF_n-P when the linkers fully stretch. As in **Figure 3.3**, both the 15 aa linker and 30 aa linkers are sufficiently long for P to reach the opposite end of subunit, which means P is accessible to all residues of HF_n subunit outer surface.

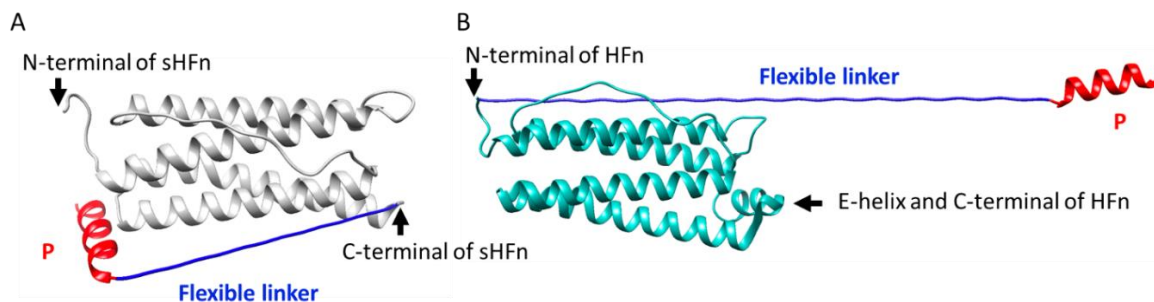


Figure 3.3 Schematics of theoretical P-HFn and sHFn-P subunits with fully stretched linkers. **A**, sHFn-P. **B**, P-HFn. Created using Discovery Studio Visualizer and UCSF Chimera. Grey part is HFn or sHFn subunit, blue part is the stretched 15aa or 30aa linker and red part is pro-apoptotic peptide.

Therefore, 2 assumptions have been made to analyse the possible ways of P affecting HFn or sHFn subunit assembling. Assumption one (interaction assumption) is P interacts with HFn or sHFn subunit residues through non-covalent bonds. As shown in **Figure 3.3**, because of the length of the linker, P could reach and interact with all available residues on HFn or sHFn subunit outer surface. This will affect subunit assembly when P pre-occupies the assembling related residues on HFn or sHFn subunit. To simplify the analysis, P and HFn or sHFn subunit were analysed separately as individual molecules and HADDOCK 2.4, a protein-peptide docking webserver was used to identify the possible interactions poses.²²

Another assumption (no interaction assumption) is P does not interact with the residues on HFn or sHFn subunit. When the linker bends, P primarily affects the property and conformation of HFn N-terminal in P-HFn and sHFn C-terminal in sHFn-P to disrupt the assembly of subunits. Because it is difficult and time-consuming to form long peptide fused protein structure, and our aim is to study the local impact of inserted P, a 3-residue linker GSG was usable to represent the 15-residue linker in this case. Structures of P-HFn and sHFn-P with 3 aa linker (P3HFn and sHFn3P) were generated and their hydrophobicity and charge properties would be analysed in comparison with those of HFn and sHFn. When the linker does not bend, P floats around Hfn or sHFn subunit and probably disrupts assembling by repelling another P. Charge and hydrophobic property of P surface may elucidate the reason.

The computation analysis followed different steps for each assumption, as presented in **Figure 3.4**. 1) 3D structure PDB files generation. Raw HFn subunit PDB file (ID: 2FHA) was obtained from RSCB PDB.²³ Raw sHFn subunit PDB was generated using Discovery Studio Visualizer based on HFn PDB by directly deleting the E-helix. Raw PDB files of P and P with 3 aa linker

(GSG) at N- or C terminal were generated using the online PEP-FOLD server.^{24, 25} Raw PDB files of sHFn3P subunit and P3HFn subunit were obtained using USCF Chimera and Discovery Studio Visualizer by putting together the PDB file of sHFn and PDB file of linker GSG with P, and HFn subunit PDB file and P with linker GSG PDB file, respectively. Because the 3D structure of sHFn, P, sHFn3P and P3HFn were artificial PDB files, in order to assess the stability and obtain stable PDB files of these 4 proteins, 5 ns 300 K molecular dynamic simulations were applied using Gromacs 2020.4 software. OPLS-AA/L all-atom force field, Lincs constraint algorithm, Verlet cut-off scheme, Particle Mesh Ewald coulomb type were used in this MD simulation. HFn subunit 5 ns MD simulation was also conducted as a control. RMSD and radius of gyrate in simulation was recorded to assess structure stabilities. 2) Protein-peptide docking analysis for interaction assumption. 5 ns MD simulated PDB files of HFn subunit, sHFn subunit and P were the input structures in docking. Active residues were surface residues of protein and peptide. Interactions within final complexes PDB files with the highest probability were visualised and analysed using Discovery Studio Visualiser. Non-interaction assumption did not include this step. 3) The hydrophobicity and charge property comparison. In interaction assumption, the top complex surface hydrophobicity and charge property were compared with those of HFn and sHFn subunit. With no interaction assumption, the MD simulated sHFn3P and P3HFn subunit hydrophobicity and charge property were analysed. P hydrophobicity and charge property were also visualised.

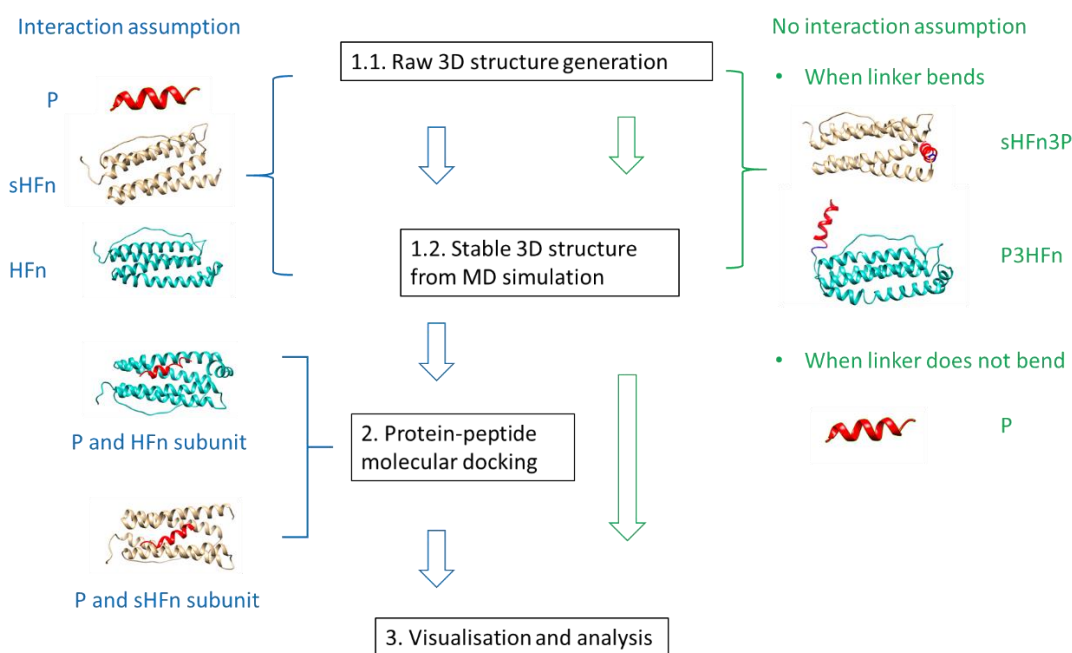


Figure 3.4 Procedure for computational analysis of P impact on HFn assembling.

3.3 Results and discussion

3.3.1 Characterisations of plasmid of sHF_n, sHF_n-P and sHF_n-P-RGDK

3.3.1.1 Characterisation and recovery of endonuclease treated pET30a vector

In the restriction endonuclease treatment of plasmid vector, one enzyme was used at a time to ensure digestion efficiency. As is shown in **Figure 3.5A**, after Nde I treatment, the apparent molecular weight of the vector increased, implying the change of vector conformation. Theoretically, after restriction endonuclease treatment, the ring-shaped plasmid would turn into a chain-like structure and its migration speed would lower down because of the enlargement of its size. A single band in each lane suggests a high enzyme treatment efficiency. After Xho I treatment, the nucleotides between Nde I and Xho I were removed from the vector and its migration speed slightly increased (**Figure 3.5B**). After the recovery of digested plasmid from agarose gel, 30 μ L Nde I and Xho I digested vector was obtained.

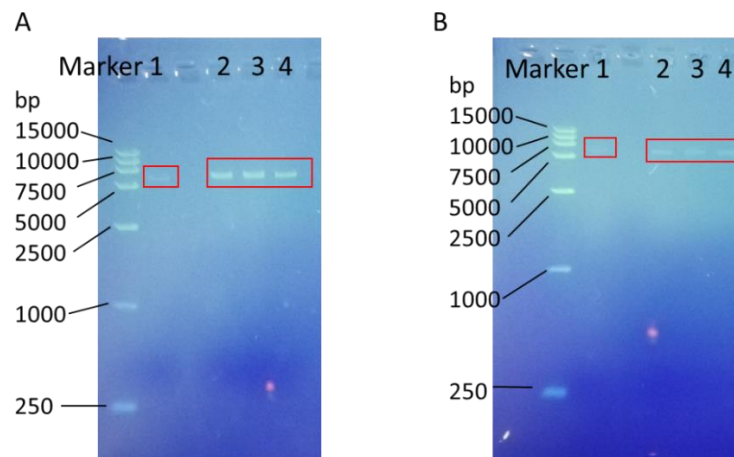


Figure 3.5 1 % agarose gel electrophoresis results of pET30a vector restriction enzyme treatment. **A**, Nde I treatment result, lane 1, pET30a plasmid; lane 2-4, Nde I digested pET30a. **B**, Nde I and Xho I treatment result, lane 1, Nde I digested pET30a plasmid, lane 2-4, Xho I and Nde I digested pET30a.

3.3.1.2 Characterisation and recovery of target genes for sHF_n, sHF_n-P and sHF_n-P-RGDK

Genes obtained from PCR consist of additional base pairs for enhancement of restriction endonuclease performance, start and stop codon and the codons encoding target protein residues. The theoretical PCR products for sHF_n, sHF_n-P and sHF_n-P-RGDK are 505, 592 and 613 bp, respectively. In **Figure 3.6A** and **C**, the Mw of the strongest bands in lanes are close to the theoretical Mw and these bands were recovered. Other impurities were derived from incorrect primer binding and incomplete elongation. Because of the additional base pairs at

both ends, PCR products are relatively easy to be digested compared with plasmid, so the two restriction endonuclease digestion was combined in one step. In the digestion step, only the additional base pairs and part of the restriction nuclease recognition sites were cut off. Therefore, the Mw after digestion were similar to that before digestion (**Figure 3.6B and D**).

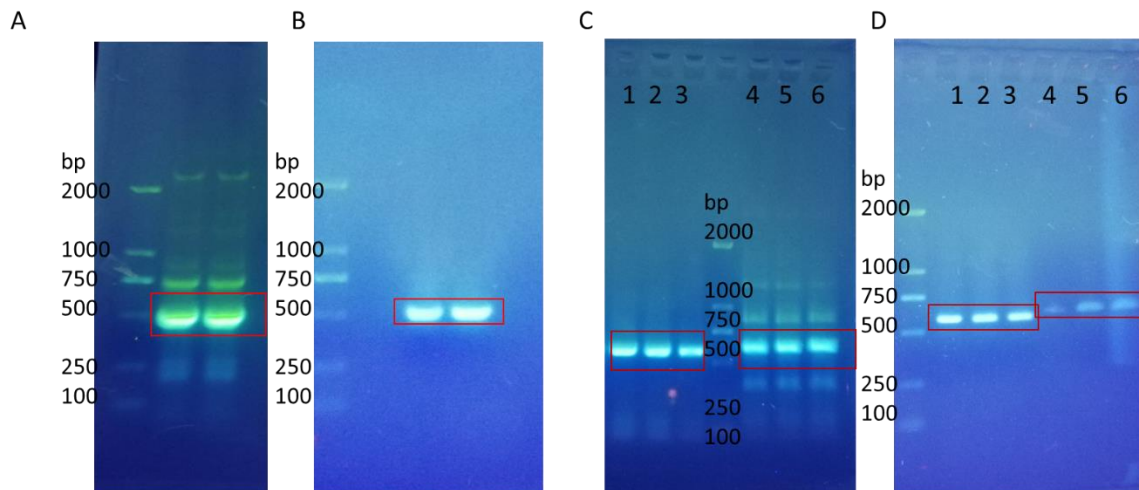


Figure 3.6 1.5 % agarose gel electrophoresis results of PCR for sHFn, sHFn-P and sHFn-P-RGDK. **A**, sHFn PCR product result. **B**, restriction enzyme treated sHFn PCR product result. **C**, sHFn-P and sHFn-P-RGDK PCR products results, lane 1-3, PCR products for sHFn-P; lane 4-6, PCR products for sHFn-P-RGDK. **D**, restriction enzyme treated sHFn-P and sHFn-P-RGDK PCR product results, lane 1-3, enzyme treated PCR products for sHFn-P; lane 4-6, enzyme treated PCR products for sHFn-P-RGDK.

3.3.2 Confirmation of plasmid transformation into *E. coli*

Plasmids constructed by BGI Company (China) for HFn, P-HFn, HFn-PAS, HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK and ligation products for sHFn, sHFn-P and sHFn-P-RGDK were transformed into competent *E. coli* BL21 (DE3) cells. 1 % agar plate with kanamycin was used for selectively growth of cells with plasmids. As is shown in **Figure 3.7**, there were colonies growing on all 8 plates, suggesting the success of plasmid transformation. Three colonies were picked from plates and sent to BGI Company (China) for plasmid sequencing analysis. The sequencing results demonstrate that all correct target genes were inserted between Nde I and Xho I recognition sites in pET30a (**Appendix A, Table A2**).

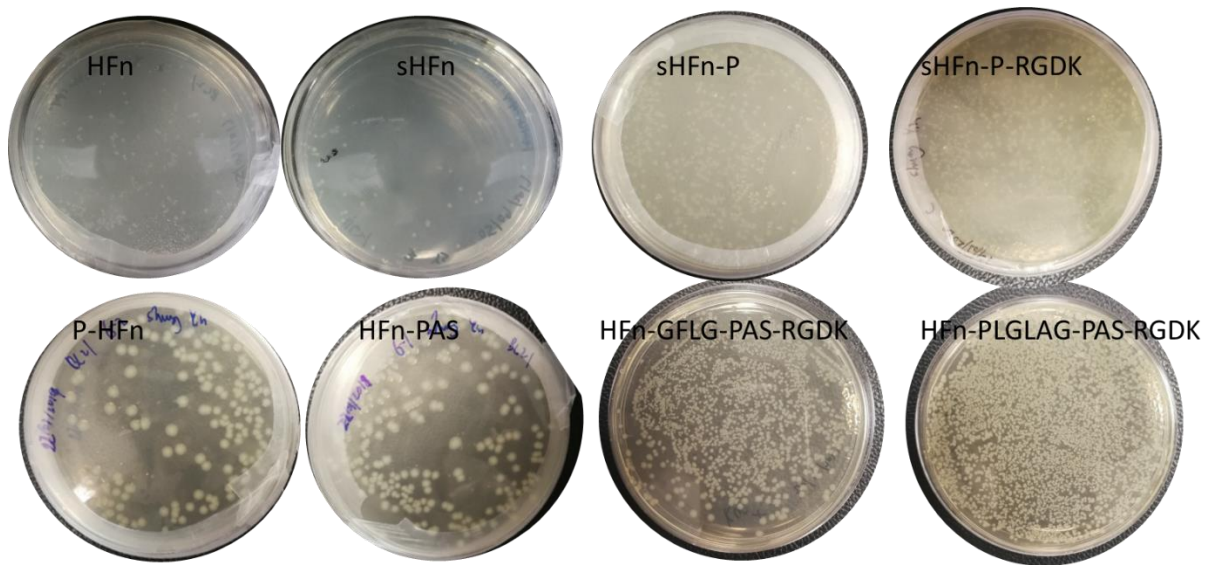


Figure 3.7 Pictures of *E. coli* BL21 (DE3) colonies grown on the agar LB plates.

3.3.3 Expression comparison of 8 HFn-based proteins

In 37 °C induction shake-flask fermentation, after centrifuge, wet biomass of bacteria from 2 L fermentation broth of 8 types of bacteria were around 8 g. In low temperature induction shake flask fermentation of P-HFn, 6.6 g wet biomass was harvested form 2 L fermentation broth. In 5 L fermenter fermentation of HFn, a greater cell density was achieved and around 90 g wet biomass was collected after centrifuge.

Figure 3.8 illustrates the bacterial growth curve in batch fermenter fermentation over time. For the first 3 h in fermenter, bacteria grew slowly. When the OD600 reached around 5, cells begun to grow exponentially and IPTG induction was started. The final OD when harvesting cells was 22.

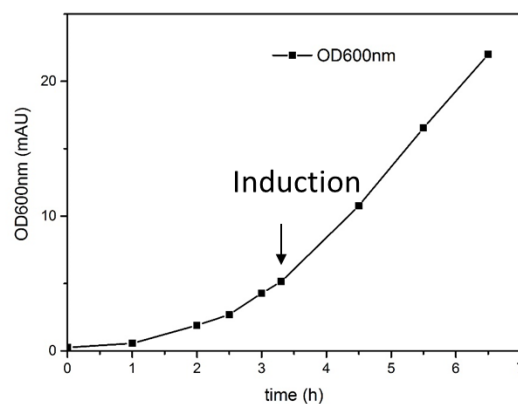


Figure 3.8 Bacterial growth curve in batch fermenter fermentation.

In **Figure 3.9**, reducing SDS-PAGE result demonstrates that all 7 HF_n-based proteins (P-HF_n not included) were successfully expressed in both soluble and insoluble forms in *E. coli* under 37 °C induction shake-flask fermentation. All protein expression levels ranged from 25 % to 32 % in soluble forms, calculated by density scan using software Image J.²⁶ HF_n and sHF_n have relatively high soluble expression levels (32 % and 30 %) than other functionalised HF_ns. Three PAS functionalised HF_ns have relatively high soluble expression levels (28 %) than those of P functionalised HF_ns, sHF_n-P and sHF_n-P-RGDK (25 %). This is in accordance with some previously reported cases where fusion has led to a reduced expression level.^{27, 28} Fermenter cultured bacteria also expressed HF_n in both soluble and insoluble forms with a similar level to HF_n shake-flask fermentation (SDS-PAGE results not shown).

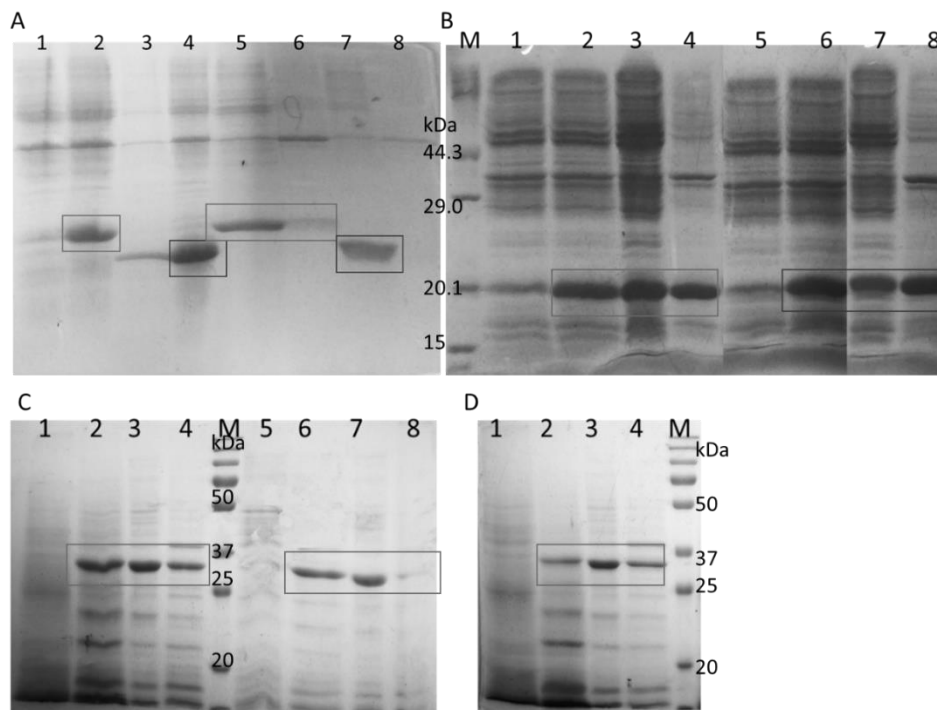


Figure 3.9 Expression results of HF_n-based proteins except for P-HF_n.

A, expression results of HF_n and sHF_n, lane 1: bacterial for HF_n before IPTG induction; lane 2: bacterial for HF_n after IPTG induction; lane 3: bacterial for sHF_n before IPTG induction, lane 4: bacterial for sHF_n after IPTG induction, lane 5: supernatant of bacterial lysate for HF_n; lane 6: sediment of bacterial lysate for HF_n, lane 7: supernatant of bacterial lysate for sHF_n; lane 8: sediment of bacterial lysate for sHF_n. **B**, expression results for sHF_n-P and sHF_n-P-RGDK, M: marker, lane 1: bacterial for sHF_n-P before IPTG induction; lane 2: bacterial for sHF_n-P after IPTG induction; lane 3: supernatant of bacterial lysate; lane 4: sediment of bacterial lysate, lane 5: bacterial for sHF_n-P-RGDK before IPTG induction; lane 6: bacterial for sHF_n-P-RGDK after IPTG induction; lane 7: supernatant of bacterial lysate; lane 8: sediment of bacterial lysate. **C**, expression results of HF_n-PAS and HF_n-GFLG-PAS-RGDK, lane 1: bacterial for HF_n-PLGLAG-PAS-RGDK before IPTG induction; lane 2: bacterial for HF_n-PLGLAG-PAS-RGDK after IPTG induction; lane 3: supernatant of bacterial lysate; lane 4: sediment of bacterial lysate; M:protein marker, lane 5: bacterial for HF_n-PAS before IPTG induction; lane 6: bacterial for HF_n-PAS after IPTG induction; lane 7: supernatant of bacterial lysate; lane 8: sediment of bacterial lysate. M: protein marker. **D**, expression result of HF_n-PLGLAG-PAS-RGDK,

lane 1: bacterial before IPTG induction; lane 2: bacterial after IPTG induction; lane 3: supernatant of bacterial lysate; lane 4: sediment of bacterial lysate. M: protein marker.

In **Figure 3.9B**, apparent Mw of sHFn-P subunit in gel was close to 20.1 kDa marker band, which is near its theoretical Mw, 21.14 kDa. However, in 3 PAS containing HFns, target bands in gel were in the middle of 25 kDa and 37 kDa, which are greater than their theoretical Mw (HFn-PAS 26 kDa, HFn-GFLG-PAS-RGDK 26.5 kDa, HFn-PLGLAG-PAS-RGDK, 26.6 kDa) (**Figure 3.9C and D**). The reason for the increase of apparent Mw is the water binding ability of PAS peptide and this phenomenon has also been discovered in previous PAS fused protein.¹⁶

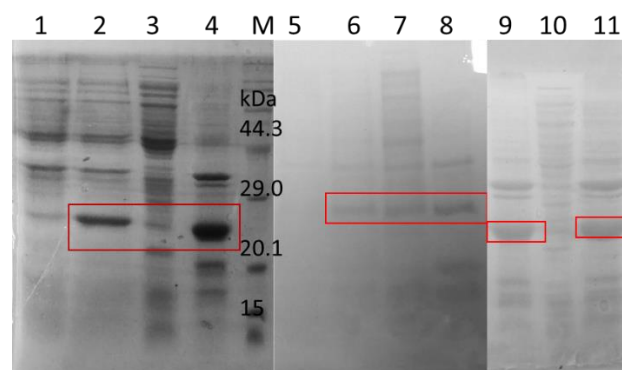


Figure 3.10 Expression results for standard and low temperature induction of P-HFn.

Lane 1: bacterial before 37 °C IPTG induction; lane 2: bacterial after 37 °C IPTG induction; lane 3: supernatant of bacterial lysate after 37 °C induction; lane 4: sediment of bacterial lysate after 37 °C induction, M: protein marker, lane 5: bacterial before 30 °C IPTG induction, lane 6: bacterial after 30 °C IPTG induction, lane 7: supernatant of bacterial lysate after 30 °C induction and cell disruption; lane 8: sediment of bacterial lysate after 30 °C induction and cell disruption; lane 9: supernatant of 30 °C IPTG induction bacterial lysate after cell disruption; lane 10: supernatant of re-centrifuged bacterial lysate supernatant after 4 °C 0.5 h storage; lane 11: precipitate of re-centrifuged bacterial lysate supernatant after 4 °C 0.5 h storage.

Regarding P-HFn, in **Figure 3.10**, the results show in 37 °C induction, almost no target protein was expressed in soluble form (lane 3). In 30 °C induction, it seems that target protein was in both supernatant and sediment of bacterial lysate. However, P-HFn in the supernatant precipitated 0.5 h after cell disruption at 4 °C, as shown in lane 10 and 11. The precipitates after re-centrifuge had similar protein composition as in IBs. This means the expressed P-HFn in supernatant was not stable and was very likely to precipitate.

In summary, P-HFn was expressed as IBs. Other 7 HFn-based proteins were expressed in both stable soluble forms and IBs. However, because of the influence of inserted peptides on HFn monomer and assembly structure, the proteins expressed in soluble forms were assemblies or not needs to be further confirmed.

3.3.4 Expression conformation comparison of soluble expression of HF_n-based proteins

TEM analysis aimed to confirm if the 7 HF_n-based proteins expressed in soluble forms were self-assemblies. In **Figure 3.11**, hollow spherical structures were observed in 5 types of bacterial lysate supernatants, including HF_n, sHF_n, HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK. Their sizes are close to each other and also close to theoretical HF_n size, 12 nm in diameter. This indicates these 5 HF_n-based proteins self-assembled when being expressed in *E. coli*. sHF_n assembly demonstrates that the truncation of E helix did not affect the assembly of subunits. This has also been proven by a previous research.²⁹ In the cases of 3 PAS fused HF_ns, results show that the insertion of PAS peptide and RGDK peptide did not affect HF_n assembly. This is because although PAS peptide is long, its residues are hydrophilic and bear no net charge at neutral pH. It is a flexible random coil and very likely to pose a minor impact on HF_n inter-subunit interactions. RGDK peptide only consists of 4 residues and is away from the HF_n subunit residues because of the flexible linkers and PAS peptide, it, therefore, has a negligible influence on HF_n inter-subunit interactions. In contrast, in 2 soluble expressed P fused HF_ns, sHF_n-P and sHF_n-P-RGDK, clusters of fragment-like structures were observed, and no spherical structure existed, as is shown in **Figure 3.11**. Therefore, it must be the functional peptide P that has negatively affected self-assembling, and it is worth further investigation of the mechanism.

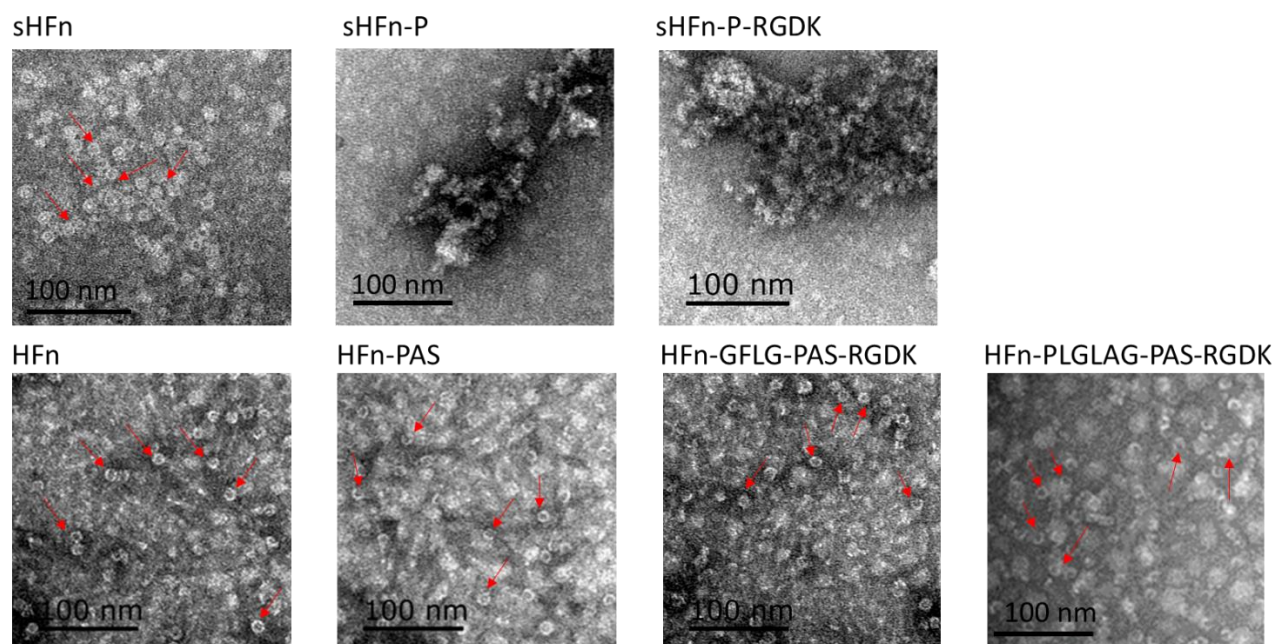


Figure 3.11 TEM images of bacterial lysate supernatants. Red arrows indicate some of the self-assembled spheres, no spherical structures were found in sHF_n-P and sHF_n-P-RGDK.

In order to further analyse if the expressed forms of sHFn-P and sHFn-P-RGDK in bacterial supernatants were soluble monomers or soluble aggregates, SEC was adopted to separate proteins in supernatants according to their hydrodynamic volumes. The SEC fractions were collected according to retention volumes and fraction protein compositions were analysed using reducing SDS-PAGE. In Superose 6 10/100 gl column (column volume 25 mL, GE Healthcare, USA), chromatogram of supernatants containing HFn has a peak at 13.9 mL, which was confirmed to be HFn assembly (SDS-PAGE result not shown) (**Figure 3.12A**). In theory, if sHFn-P and sHFn-P-RGDK are soluble aggregates, their retention volumes would be lower than 13.9 mL, and if they were soluble monomers or oligomers, due to the small sizes, their retention volume would be greater than 13.9 mL. As is shown in **Figure 3.12D** and **E**, both proteins were eluted between 16-21 mL and peaked at 17-19 mL, indicating that they were probably soluble monomers. Combined with the IBs expression result of P-HFn, it proves that the pro-apoptotic peptide has made a significant negative impact on HFn inter-subunit interactions and HFn assembly, even though the pro-apoptotic peptide is just 13 aa long.

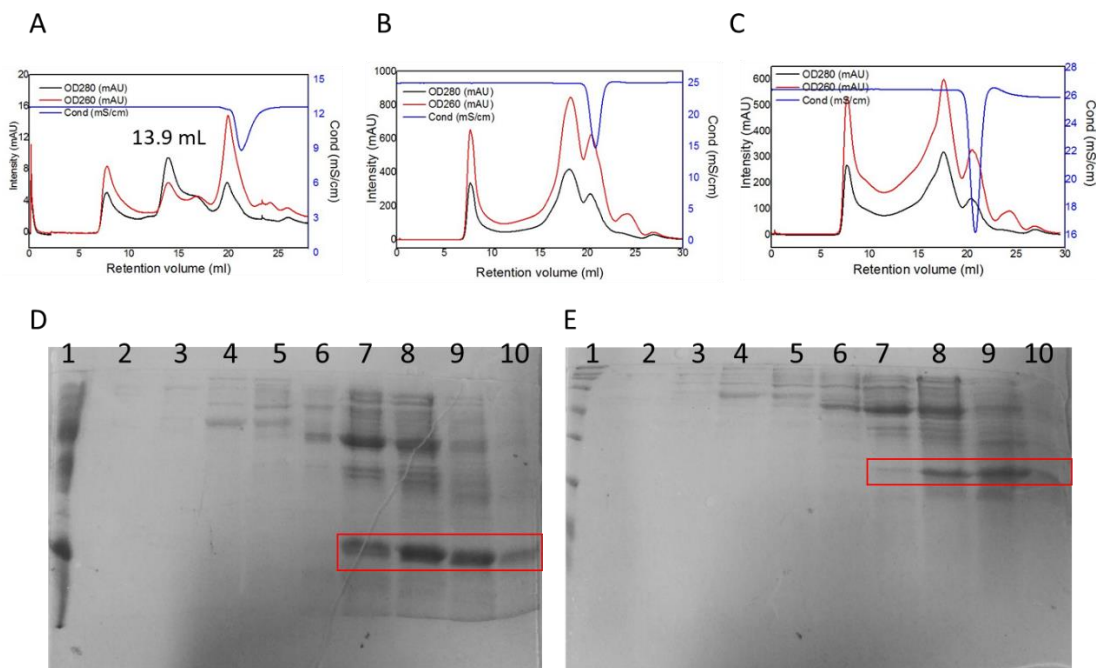


Figure 3.12 SEC of bacterial lysate supernatants of sHFn-P and sHFn-P-RGDK and reducing SDS-PAGE results of SEC fractions.

A, chromatogram of bacterial lysate supernatant of HFn. **B**, chromatogram of bacterial lysate supernatant of sHFn-P. **C**, chromatogram of bacterial lysate supernatant of sHFn-P-RGDK. **D**, SDS-PAGE results of sHFn-P bacteria lysate supernatant SEC fractions. **E**, SDS-PAGE results of sHFn-P-RGDK bacteria lysate supernatant SEC fractions. Lane 1: bacterial lysate supernatant (**D**) and protein marker (**E**); lane 2: 7-11 mL SEC fraction; lane 3: 11-13 mL fraction; lane 4: 13-14 mL fraction; lane 5: 14-15 mL fraction; lane 6: 15-16 mL fraction; lane 7: 16-17 mL fraction; lane 8: 17-18 mL fraction; lane 9: 18-19 mL fraction; lane 10: 19-21 mL fraction.

3.3.5 Impact of pro-apoptotic peptide on HF_n and sHF_n biochemical property and structure by computational analysis

As there were 2 assumptions where the mechanism behind the impact of P on HF_n structure is different, computational analysis included both cases. In the first step of computational analysis, stable 3D structures of P, HF_n subunit, sHF_n subunit, P3HF_n and sHF_n3P were obtained from MD simulation. During the 5 ns MD simulation, the root mean square diameter (RMSD) and radius of gyrate (Rg) over time imply the stability of the structure in water at neutral pH. The greater the RMSD value and the more fluctuated Rg is, the less stable the structure is. **Figures 3.13** and **3.14** show the RMSD and Rg during the 5 ns MD simulation of all raw structures. Comparing HF_n with P3HF_n, and sHF_n with sHF_n3P, the addition of P has decreased the stability of HF_n/sHF_n subunit. P3HF_n subunit has the highest RMSD and most fluctuated Rg. sHF_n3P subunit was the second unstable structure. HF_n and sHF_n subunit are quite similar. The α -helix structure of P obtained from PEP-FOLD server was also quite stable during the simulation.

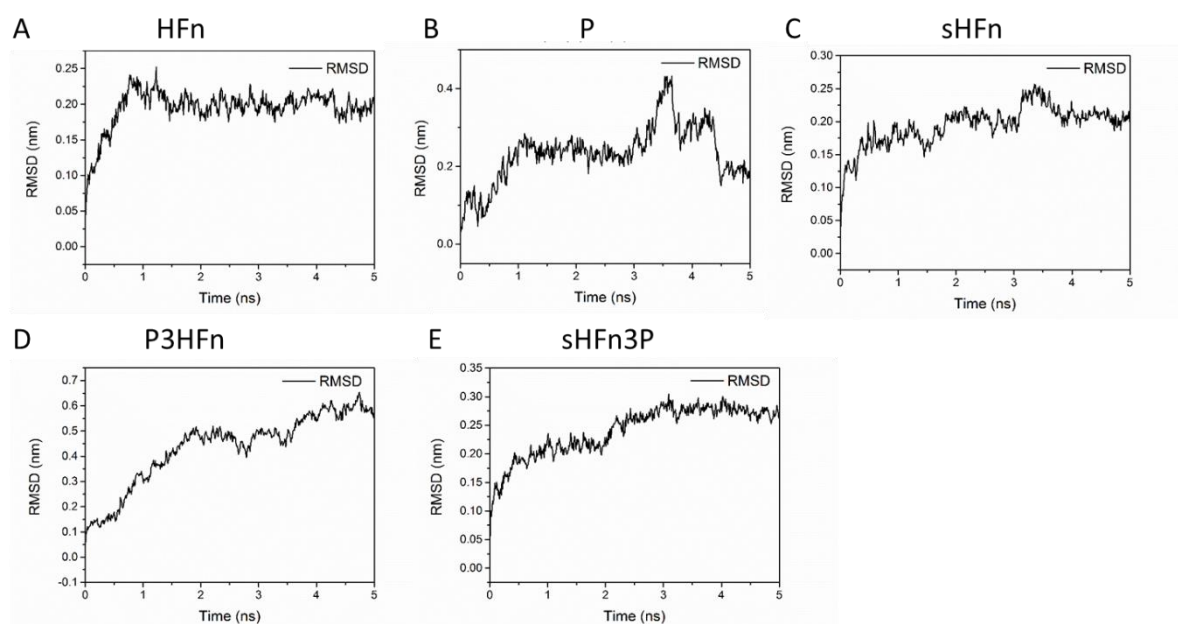


Figure 3.13 RMSD of structures in 5 ns Gromacs MD simulation.

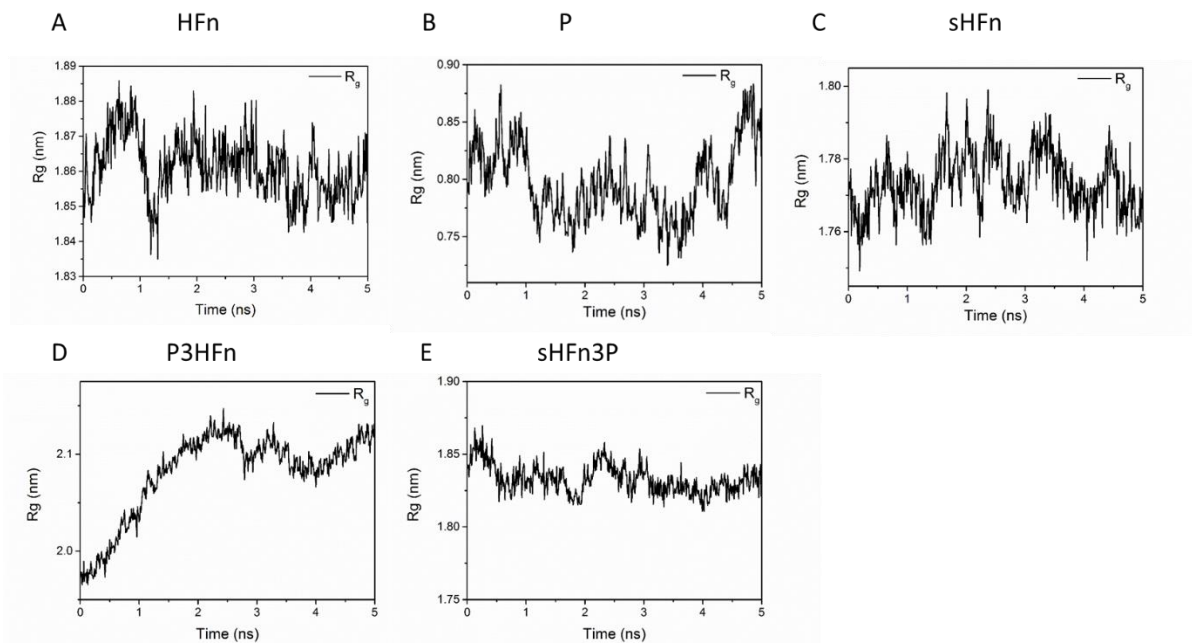


Figure 3.14 Radius of gyration of structures in 5 ns Gromacs MD simulation.

3.3.5.1 Assumption of *P* and *HFn/sHFn* interaction

In the interaction assumption, because of the long flexible linker, *P* is able to freely interact with any surface residues of *HFn* or *sHFn* subunits in *P-HFn* and *sHFn-P*. Docking was adopted to analyse the possible complexes. The top 4 docking complexes of *P* with *HFn* subunit are illustrated in **Figure 3.15**. The hydrogen bond, salt bridge and pi (π) effects interactions of *HFn* subunit with *P* in the complex with the highest probability are listed in **Table 3.3**. Many of the interactions within other complexes are the same as in top 1 complex (data not shown), and top 1 complex has the most interactions, therefore its interactions are presented as an exemplar.

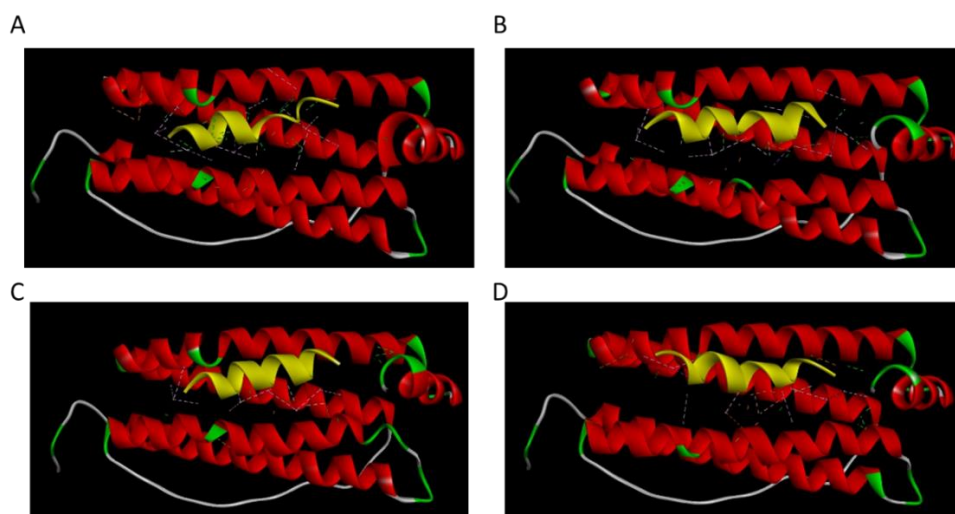


Figure 3.15 3D structure of top 4 complexes of *HFn* subunit and *P* obtained from HADDOCK 2.4 docking analysis.

The yellow molecule is P, dotted lines are the interactions between HF_n subunit and P. Structures were visualised using discovery studio.

Table 3.3 Interacted residues of HF_n and interaction types in top 1 complex.

HF _n residue	Residue location	Interaction type	HF _n residue	Residue location	Interaction type
Glu107	Helix C	H-bond, salt bridge	Gln141	Helix D	H-bond
Glu62	Helix B	H-bond, salt bridge	His60	Helix B	H-bond
Glu61	Helix B	H-bond, salt bridge	His65	Helix B	H-bond, pi-alkyl
Glu64	Helix B	H-bond, salt bridge	His128	Helix D	Pi-cation
Glu147	Helix D	H-bond, salt bridge	Lys68	Helix B	Alkyl
Asp131	Helix D	Salt bridge	Ala144	Helix D	Alkyl
Gln58	Helix B	H-bond, alkyl, pi-alkyl	lys143	Helix D	Alkyl
His136	Helix D	H-bond	leu148	Helix D	Alkyl
Leu175	C terminal	H-bond	Tyr54	Helix B	Pi-alkyl
Gly176	C terminal	H-bond	His57	Helix B	Pi-alkyl
Thr174	Helix E	H-bond	Phe132	Helix D	Pi-alkyl
Tyr34	Helix B	H-bond			

As is shown in **Figure 3.15**, in the top 4 complexes, P is bound to the four-helix bundle of HF_n with slightly different orientations. In **Table 3.3**, the majority of interacted residues of HF_n subunit in top 1 complex is from helix B and D, many of which have been proven to be involved in HF_n assembling. Interactions in the top 1 complex include 14 hydrogen bonds and 6 salt bridges, showing the binding between HF_n subunit and P is quite strong. The cationic Lys residues (Lys1, Lys4, Lys7, Lys8, Lys11, Lys14) in P are primarily responsible for the hydrogen bonds and salt bridges with HF_n residues.

Figure 3.16 is the top 4 complexes of P with sHF_n subunit and **Table 3.4** lists the interaction types and sHF_n residues in the top 1 complex. Binding location and orientation of P in top 4 complexes are quite similar to those of P to HF_n subunit. This can be explained by the similar structure of HF_n and sHF_n subunit. Even the interacted residues resemble those in top 1 complex of HF_n with P. Most of the sHF_n subunit residues involved are from helix B and D.

Interacted residues in P are also mainly the multiple Lys residues. 15 hydrogen bonds and 7 salt bridges are formed in the top 1 complex.

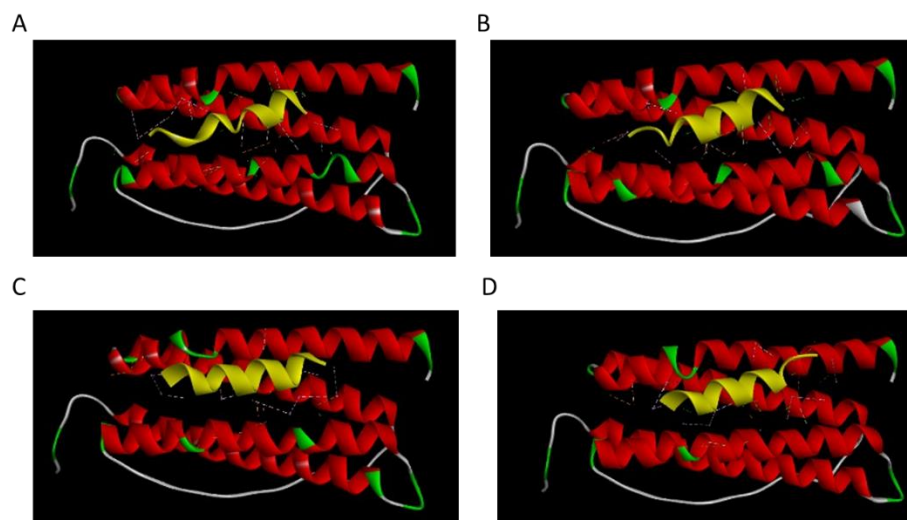


Figure 3.16 3D structure of top 4 complexes of sHFn subunit and P obtained from HADDOCK 2.4 docking analysis.

The yellow molecule is P, dotted lines are the interactions between sHFn subunit and P. Structures were visualised using discovery studio.

Table 3.4 Interacted residues of sHFn and interaction types in top 1 complex.

sHFn residue	Residue location	Interaction type	sHFn residue	Residue location	Interaction type
Glu147	Helix D	H-bond, salt bridge	Tyr137	Helix D	H-bond
Glu62	Helix B	H-bond, salt bridge	His60	Helix B	H-bond
Glu107	Helix C	H-bond, salt bridge	Gln75	Helix B	H-bond
Glu61	Helix B	H-bond, salt bridge	His136	Helix D	H-bond, pi-alkyl
Glu64	Helix B	H-bond, salt bridge	Tyr54	Helix B	H-bond, pi-alkyl
Glu67	Helix B	H-bond, salt bridge	Ala144	Helix D	Alkyl
Glu27	Helix A	H-bond, salt bridge	His57	Helix B	Pi-alkyl
Lys68	Helix B	H-bond	His65	Helix B	Pi-alkyl
Lys143	Helix D	H-bond, pi-alkyl	His128	Helix D	Pi-alkyl
Lys53	Helix B	H-bond	Phe132	Helix D	Pi-alkyl

Figure 3.17 illustrates the surface hydrophobicity and charge property of the top 1 complex of P with HFn and sHFn subunit. In contrast to HFn or sHFn, due to the addition of P, hydrophilic surface turns into hydrophobic (yellow box area), and negatively charged surface becomes

neutral in neutral pH. The reduction of charge and increase of hydrophobicity possibly decrease the repulsion force of this local area between monomers and increase the likelihood of by-product subunit aggregate such as IBs rather than correct assembly.

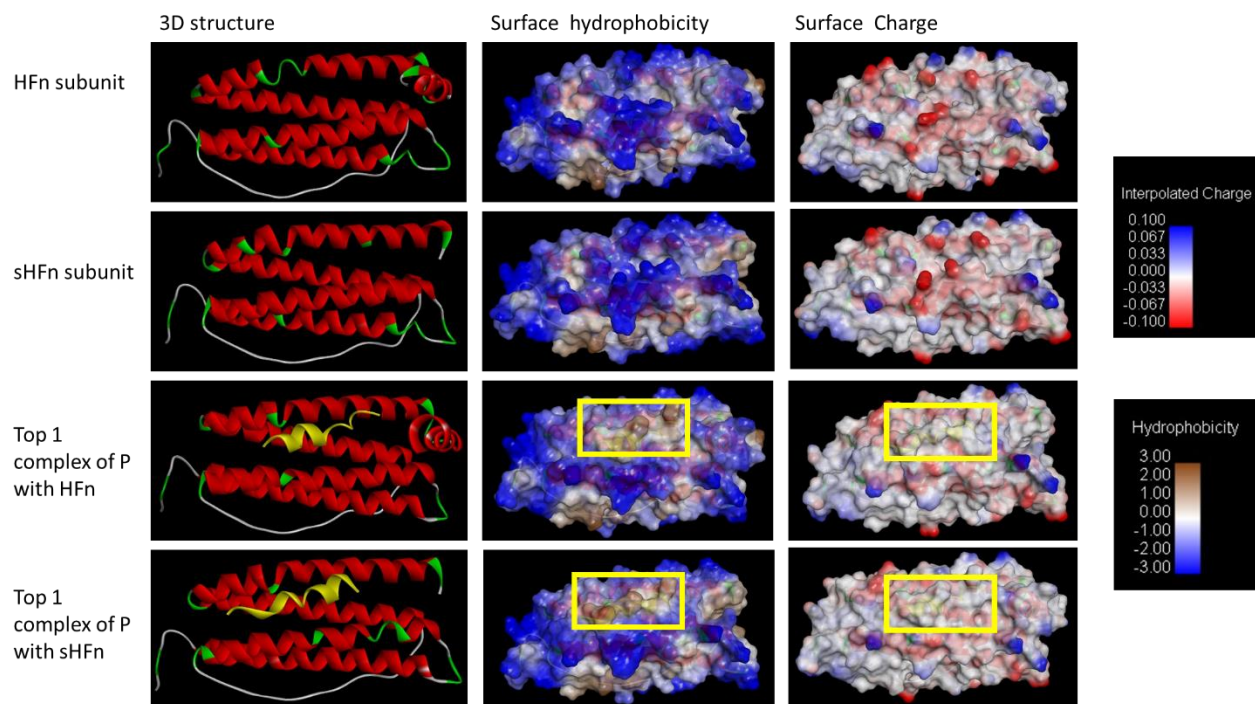


Figure 3.17 Surface hydrophobicity and charge comparison of top 1 complexes with HFfn and sHFfn subunit.

Yellow helix in 3D structure is P, the area in yellow box is surface of P in complex. Structures were visualised using discovery studio.

3.3.5.2 Assumption of no interaction between P and HFfn/sHFfn

In the no interaction assumption, when linker bends, sHFfn3P and P3HFfn were used for analysis of property changes in sHFfn-P and P-HFfn. **Figure 3.18** illustrates the view from C terminal to compare the C-terminal structure, hydrophobicity and charge at neutral pH of sHFfn3P with sHFfn and HFfn subunit.

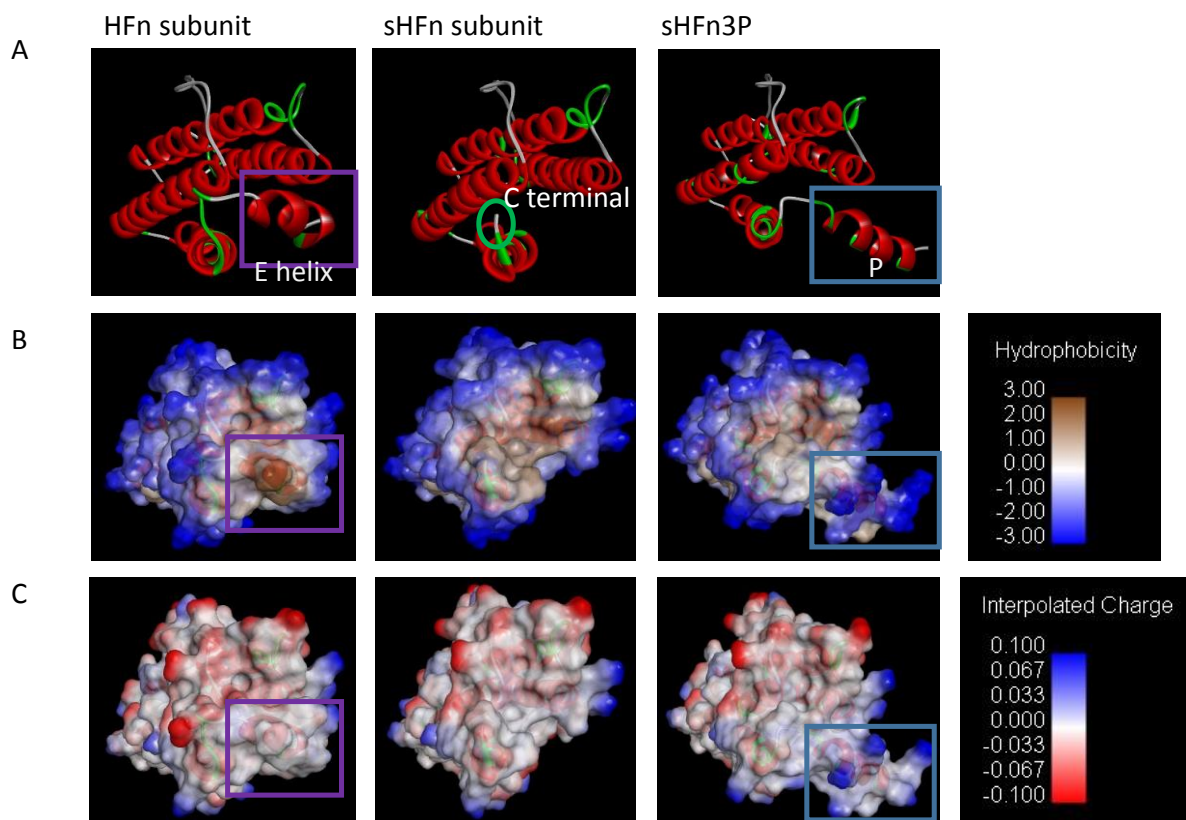


Figure 3.18 Comparison of sHFfn3P with HFfn and sHFfn subunit after Gromacs MD simulation. **A**, subunit structure. **B**, C-terminal hydrophobicity. **C**, C-terminal charge. Structures were visualised using discovery studio.

In HFfn subunit, the E-helix is hydrophobic and bears almost no charge. In HFfn assembly, 4 E-helices gather together around the hydrophobic channel, the hydrophobic attractions of which contribute to the stability of the whole assembly.²⁹ In sHFfn subunit, the E-helix and the hydrophobic attraction are missing but due to other inter-subunit non-covalent bonds, sHFfn can still assemble but with a relatively low structural stability. In contrast, in sHFfn3P subunit, P is overall quite hydrophilic, suggesting no possible hydrophobic attractions. Additionally, P bears positive charges, leading to a repulsion between C-terminals when in an assembly. Plus, the helix structure of P adds steric hindrance to the C-terminal area. All the above 3 factors are responsible for the failure of sHFfn-P subunit to self-assemble. As a result, sHFfn-P was expressed as monomers.

Figure 3.19 illustrates the structure comparison of P3HFfn, HFfn and sHFfn subunit N-terminal. In HFfn and sHFfn subunits, the N-terminal extension is a short random coil and floating outside the HFfn assembly, with no interactions of each other. As is shown in **Figure 3.19B** and **C**, it is hydrophilic and has no charge at neutral pH. In P3HFfn subunit, P is half hydrophobic and

half hydrophilic (**Figure 3.19B**), with hydrophobic part of P facing outwards, which could cause unwanted aggregation with other P3HF_n subunits. In terms of charge, positive charges are distributed in the half of P that faces the other subunit N terminal in an assembly in theory. Repulsion caused by these charges may impose a negative effect on self-assembly. In addition, the steric hindrance of the rigid helix structured P at the N-terminal can negatively affect assembly. These could explain the fact that P-HF_n is expressed in aggregated IBs instead of soluble assembly.

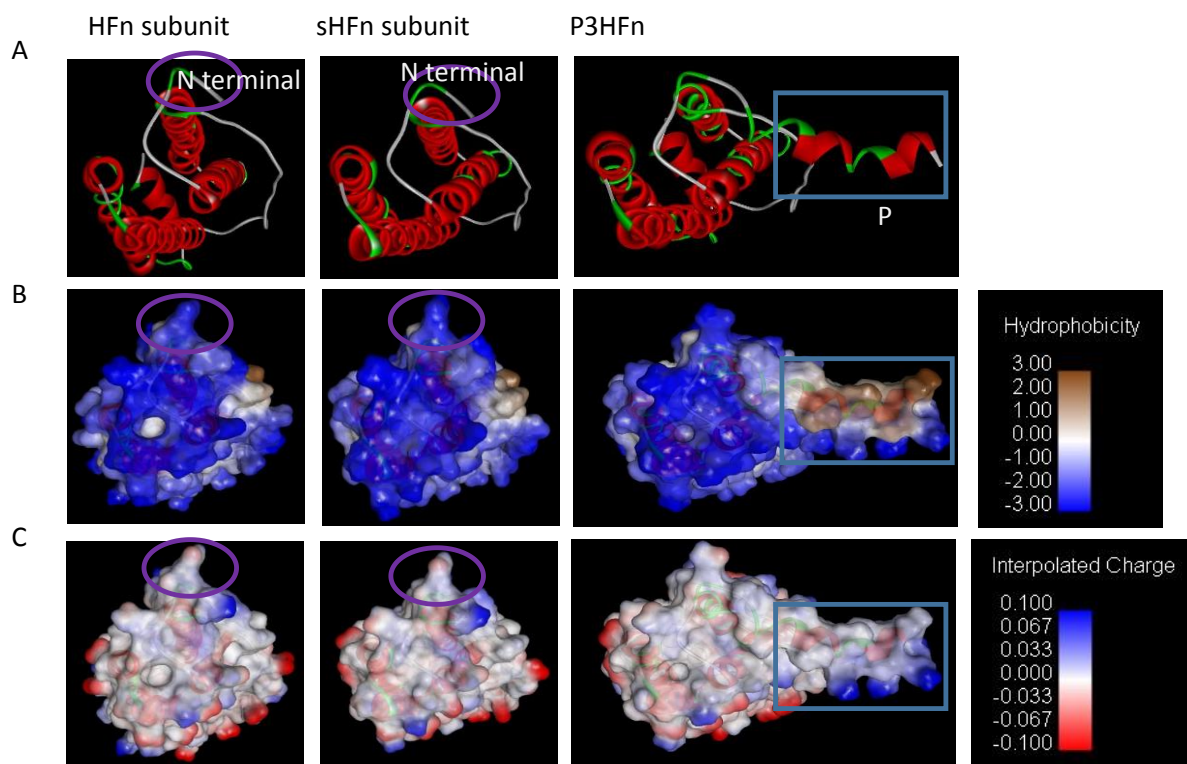


Figure 3.19 Comparison of P3HF_n with HF_n and sHF_n subunit N-terminal after Gromacs MD simulation.

A, subunit structure. **B**, N-terminal hydrophobicity. **C**, N-terminal charge. Structures were visualised using discovery studio.

Except for the hydrophobicity, charge and structural changes at fusion sites, there are other structural changes found in comparison of sHF_n3P and P3HF_n with HF_n and sHF_n subunits. As is shown in **Figure 3.20**, the transition of helix to random coil of α -helix exists in sHF_n3P and P3HF_n subunit. sHF_n3P has helix-to-random coil transition in helix B and C whilst P3HF_n has in helix B, C and D. The change of helix in these areas can also contribute to a disruption of inter-subunit non-covalent bonds and the stability decrease when forming an assembly, because helix bundle residues are essential for HF_n assembling.



Figure 3.20 Structures of HFfn, sHFfn, sHFfn3P and P3HFfn subunit after MD simulation, front view. Red ribbons are α -helix and green and white lines are random coil. Arrows show the structural differences.

In no interaction assumption, when linker does not bend and P floats around, P is very likely to repel each other to negatively affect HFfn or sHFfn assembling. **Figure 3.21** shows the surface charge property of P. The positive charges can result in an electrostatic repulsion. However, this assumption is actually the least possible interruption mechanism compared with the interaction assumption and the other possible no interaction assumption, as P has posed the least damage to HFfn or sHFfn assembling.

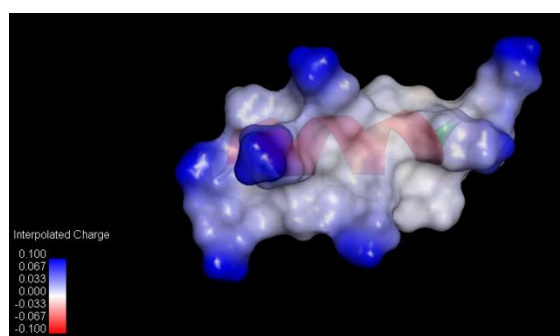


Figure 3.21 Surface charge property of P.

To conclude, in the interaction assumption, P probably interacts with multiple four-helix bundle residues of HFfn and sHFfn subunit by forming hydrogen bonds and salt bridges to block the assembling related interactions. In no interaction assumption, when linker bends, P mainly affects HFfn and sHFfn terminal hydrophobicity and charge property and impose steric hindrance to impair the stability of assembly. With the least possibility, P repels each other

through charge repulsion if P floats around HF_n or sHF_n subunit and does not interact with subunit. The multiple Lys residues which bear positive charges and tend to form hydrogen bonds are the main cause of assembling failure, and the half outer surface hydrophobicity and rigid structure of P also make a difference.

3.4 Conclusions

Six functionalised HF_ns in two functionalisation strategies were expressed by *E. coli* and the expression level and form were compared.

Based on the results in this chapter, there are 6 key findings:

1. Three PAS functionalised HF_ns, designed in the first functionalisation strategy, HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK, together with 2 controls, sHF_n and HF_n, were successfully expressed in self-assembled nanoparticles in *E. coli*. The random coil structured long PAS peptide and the short four residue peptide RGDK did not significantly affect HF_n self-assembly and the expression level.

2. In the second strategy, P functionalisation strategy, the soluble expression level of sHF_n-P, sHF_n-P-RGDK and P-HF_n were lower than controls (HF_n and sHF_n) and 3 PAS functionalised HF_ns in the first functionalisation strategy, showing a stronger impact of P on HF_n in contrast with PAS.

3. Three P functionalised HF_ns failed in self-assembly in *E. coli*. sHF_n-P and sHF_n-P-RGDK were found to be in soluble monomers whilst P-HF_n was expressed as inclusion bodies in both 37 and 30 °C induction fermentation.

4. In the design of the P functionalised strategy, there are 2 possible mechanisms of P impacting HF_n assembling, therefore 2 assumptions, the interaction and no interaction assumption were proposed. Computational analysis demonstrates that in interaction assumption, the multiple Lys residues are the main cause for the assembling failure. They form hydrogen bonds, salt bridges with HF_n or sHF_n subunit residues to pre-occupy the HF_n assembling related interactions. The hydrophobicity and the charge property change in complexes of P with HF_n or sHF_n also increase the possibility of aggregation, possibly leading to IBs formation.

5. In the no interaction assumption, when the flexible linker bends, the local charge and hydrophobicity change caused by P together with the steric hindrance result in the failure of assembling. The insertion of P at terminal of HF_n or sHF_n subunit also cause minor structural change in 4-helix bundle. With the least possibility, in no interaction assumption, when linker does not bend, P floats around HF_n or sHF_n subunit and repel each other due to the high positive charge density.

6. Based on the results of the computational analysis and the comparison of P and PAS functionalised HF_ns, the self-assembly impairment is mainly derived from the high positive charge density and the rigid structure of the inserted P. In the design of functionalised HF_n by fusion, a non-charged, non-structured peptide will be preferred.

As was reviewed in **Chapter 2**, the assembly structure of HF_n lays the foundation of HF_n as an anti-tumour drug delivery platform. Without the assembly structure, HF_n would not have the passive tumour targeting ability, the shelter-like protective ability of drugs and the relatively long half-life in circulation compared with small molecule drugs and peptide drugs. Its potentials as an anti-tumour drug delivery platform will be significantly compromised. In this chapter, it was found that 3 P functionalised HF_n assembling failed whilst 3 PAS functionalised HF_ns succeeded. P functionalised HF_ns are very likely to be unable to form correct assembly due to the significant negative impact of P, as is analysed in computation analysis. Therefore, in the following chapters, the study of P functionalised HF_ns is not continued, and PAS functionalisation strategy is the research focus.

In **Chapter 4**, purification procedure is established for 3 PAS functionalised HF_ns and control HF_n, and structures of these are characterised and compared using multiple techniques.

3.5 References

1. He, D.; Marles-Wright, J., Ferritin family proteins and their use in bionanotechnology. *Nature Biotechnology* 2015, 32 (6), 651-657.
2. Liu, X.; Wei, W.; Huang, S.; Lin, S. S.; Zhang, X.; Zhang, C.; Du, Y.; Ma, G.; Li, M.; Mann, S.; Ma, D., Bio-inspired protein-gold nanoconstruct with core-void-shell structure: beyond a chemo drug carrier. *Journal of Material Chemistry B* 2013, 1 (25), 3136-3143.
3. Liang, M.; Fan, K.; Zhou, M.; Duan, D.; Zheng, J.; Yang, D.; Feng, J.; Yan, X., H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a

- single-dose injection. *Proceedings of National Academy of Sciences of the United States of America* 2014, 111 (41), 14900-14905.
4. Lei, Y.; Hamada, Y.; Li, J.; Cong, L.; Wang, N.; Li, Y.; Zheng, W.; Jiang, X., Targeted tumor delivery and controlled release of neuronal drugs with ferritin nanoparticles to regulate pancreatic cancer progression. *Journal of Controlled Release* 2016, 232, 131-142.
 5. Zheng, Q.; Cheng, W.; Zhang, X.; Shao, R.; Li, Z., A pH-Induced reversible assembly system with resveratrol-controllable loading and release for enhanced tumor-targeting chemotherapy. *Nanoscale Research Letters* 2019, 14 (1), 305.
 6. Su, W.; Tan, H.; Janowski, R.; Zhang, W.; Wang, P.; Zhang, J.; Zhai, H.; Li, J.; Niessing, D.; Sattler, M.; Zou, P., Ferritin-displayed GLP-1 with improved pharmacological activities and pharmacokinetics. *Molecular Pharmaceutics* 2020, 17 (5), 1663-1673.
 7. Fan, K.; Jiang, B.; Guan, Z.; He, J.; Yang, D.; Xie, N.; Nie, G.; Xie, C.; Yan, X., Fenobody: A ferritin-displayed nanobody with high apparent affinity and half-life extension. *Analytical Chemistry* 2018, 90 (9), 5671-5677.
 8. Jiang, B.; Zhang, R.; Zhang, J.; Hou, Y.; Chen, X.; Zhou, M.; Tian, X.; Hao, C.; Fan, K.; Yan, X., GRP78-targeted ferritin nanocaged ultra-high dose of doxorubicin for hepatocellular carcinoma therapy. *Theranostics* 2019, 9 (8), 2167-2182.
 9. Lee, N. K.; Lee, E. J.; Kim, S.; Nam, G. H.; Kih, M.; Hong, Y.; Jeong, C.; Yang, Y.; Byun, Y.; Kim, I. S., Ferritin nanocage with intrinsically disordered proteins and affibody: A platform for tumor targeting with extended pharmacokinetics. *Journal of Controlled Release* 2017, 267, 172-180.
 10. Jääskeläinen, A., Harinen, R. R., Soukka, T., Lamminmäki, U., Korpimäki, T., Virta, M., Biologically produced bifunctional recombinant protein nanoparticles for immunoassays. *Analytical Chemistry* 2008, 80 (3), 583-587.
 11. Boumaiza, M.; Carmona, F.; Poli, M.; Asperti, M.; Gianoncelli, A.; Bertuzzi, M.; Ruzzenenti, P.; Arosio, P.; Marzouki, M. N., Production and characterization of functional recombinant hybrid heteropolymers of camel hepcidin and human ferritin H and L chains. *Protein Engineering Design & Selection* 2017, 30 (2), 77-84.
 12. Choi, S.-H.; Choi, K.; Kwon, I. C.; Ahn, H. J., The incorporation of GALA peptide into a protein cage for an acid-inducible molecular switch. *Biomaterials* 2010, 31 (19), 5191-5198.
 13. Wang, K.; Zhang, X.; Liu, Y.; Liu, C.; Jiang, B.; Jiang, Y., Tumor penetrability and anti-angiogenesis using iRGD-mediated delivery of doxorubicin-polymer conjugates. *Biomaterials* 2014, 35 (30), 8735-8747.
 14. Wang, J.; Lei, Y.; Xie, C.; Lu, W.; Yan, Z.; Gao, J.; Xie, Z.; Zhang, X.; Liu, M., Targeted gene delivery to glioblastoma using a C-end rule RGERPPR peptide-functionalised polyethylenimine complex. *International Journal of Pharmaceutics* 2013, 458 (1), 48-56.
 15. Harari, D.; Kuhn, N.; Abramovich, R.; Sasson, K.; Zozulya, A. L.; Smith, P.; Schlapschy, M.; Aharoni, R.; Koster, M.; Eilam, R.; Skerra, A.; Schreiber, G., Enhanced in vivo efficacy of a type I interferon superagonist with extended plasma half-life in a mouse model of multiple sclerosis. *Journal of Biological Chemistry* 2014, 289 (42), 29014-29029.
 16. Falvo, E.; Tremante, E.; Arcovito, A.; Papi, M.; Elad, N.; Boffi, A.; Morea, V.; Conti, G.; Toffoli, G.; Fracasso, G.; Giacomini, P.; Ceci, P., Improved doxorubicin encapsulation and pharmacokinetics of ferritin-fusion protein nanocarriers bearing proline, serine, and alanine elements. *Biomacromolecules* 2016, 17 (2), 514-522.

17. Shim, M. K.; Moon, Y.; Yang, S.; Kim, J.; Cho, H.; Lim, S.; Yoon, H. Y.; Seong, J.-K.; Kim, K., Cancer-specific drug-drug nanoparticles of pro-apoptotic and cathepsin B-cleavable peptide-conjugated doxorubicin for drug-resistant cancer therapy. *Biomaterials* 2020, 261, 120347.
18. Yao, Z.; Yuan, T.; Wang, H.; Yao, S.; Zhao, Y.; Liu, Y.; Jin, S.; Chu, J.; Xu, Y.; Zhou, W., MMP-2 together with MMP-9 overexpression correlated with lymph node metastasis and poor prognosis in early gastric carcinoma. *Tumor Biology* 2017, 39 (6), 1010428317700411.
19. Chen, W. H.; Xu, X. D.; Luo, G. F.; Jia, H. Z.; Lei, Q.; Cheng, S. X.; Zhuo, R. X.; Zhang, X. Z., Dual-targeting pro-apoptotic peptide for programmed cancer cell death via specific mitochondria damage. *Scientific Reports* 2013, 3, 3468.
20. Chu, D. S.; Bocek, M. J.; Shi, J.; Ta, A.; Ngambenjwong, C.; Rostomily, R. C.; Pun, S. H., Multivalent display of pendant pro-apoptotic peptides increases cytotoxic activity. *Journal of Controlled Release* 2015, 205, 155-161.
21. Ellerby, H. M.; Arap, W.; Ellerby, L. M.; Kain, R.; Andrusiak, R.; Del Rio, G.; Krajewski S.; Lombardo, C. R.; Rao, R.; Ruoslahti, E.; Bredesen, D. E., Anti-cancer activity of targeted pro-apoptotic peptides. *Nature Medicine* 1999, 5 (9), 1032-1038.
22. Van Zundert, G.; Rodrigues, J.; Trellet, M.; Schmitz, C.; Kastiris, P.; Karaca, E.; Melquiond, A.; van Dijk, M.; De Vries, S.; Bonvin, A., The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *Journal of Molecular Biology* 2016, 428 (4), 720-725.
23. Hempstead, P. D.; Yewdall, S. J.; Fernie, A. R.; Lawson, D. M.; Artymiuk, P. J.; Rice, D. W.; Ford, G. C.; Harrison, P. M., Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution. *Journal of Molecular Biology* 1997, 268 (2), 424-448.
24. Shen, Y.; Maupetit, J.; Derreumaux, P.; Tufféry, P., Improved PEP-FOLD approach for peptide and miniprotein structure prediction. *Journal of Chemical Theory and Computation* 2014, 10 (10), 4745-4758.
25. Thévenet, P.; Shen, Y.; Maupetit, J.; Guyon, F.; Derreumaux, P.; Tufféry, P., PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Research* 2012, 40 (W1), W288-W293.
26. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W., NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 2012, 9 (7), 671-675.
27. Uchida, M., Flenniken, M.L., Allen, M., Willits, D.A., Crowley, B.E., Brumfield, S., Willis, A. F., Jackiw, L., Jutila, M., Young, M. J., Douglas, T., Targeting of cancer cells with ferrimagnetic ferritin cage nanoparticles. *Journal of the American Chemical Society* 2006, 128, 16626-16633.
28. Li, X.; Qiu, L.; Zhu, P.; Tao, X.; Imanaka, T.; Zhao, J.; Huang, Y.; Tu, Y.; Cao, X., Epidermal growth factor–Ferritin H-chain protein nanoparticles for tumor active targeting. *Small* 2012, 8 (16), 2505-2514.
29. A.Luzzago, G. C., Isolation of point mutations that affect the folding of the H chain of human ferritin in E.coli. *The EMBO Journal* 1989, 8 (2), 569-576.

Chapter 4

PURIFICATION PROCESS DEVELOPMENT AND STRUCTURE COMPARISON OF HUMAN HEAVY-CHAIN FERRITIN AND FUNCTIONALISED HUMAN HEAVY-CHAIN FERRITINS

4.1 Introduction

The purity and conformation of human heavy-chain ferritin (HF_n) and functionalised HF_ns are crucial to their bioactivity in application as an anti-tumour drug delivery platform. *Escherichia coli* (*E. coli*) expressed HF_n and functionalised HF_ns are in co-existence with host cell proteins (HCPs) and nucleic acid contaminants.¹ Residual HCPs and nucleic acid contaminants in HF_n products are likely to cause safety risks in real-life application and problems in characterisation. Purification of HF_n and functionalised HF_ns to remove HCPs and nucleic acid contaminants is essential. In addition, the maintaining of the native assembly conformation in purification procedure matters, because the assembly is the basis of HF_n drug delivery.

In recombinant HF_n purification process, its high thermal stability has inspired researchers to heat *E. coli* lysate to precipitate host cell proteins (HCPs). Also its nanometre-scale size allows for a usage of size-exclusion chromatography (SEC) and ultra-centrifugation to separate HCPs. Ion exchange chromatography (IEC), and DNase and RNase treatment are commonly used to remove host cell nucleic acid. Li et al. used a 70 °C 15 min heating of *E. coli* lysate followed by sucrose density ultra-centrifugation or Ammonium sulphate (AS) precipitation with SEC.² Masuda et al. first heated lysate at 60 °C for 10 min and then did AS precipitate, prior to Q Sepharose IEC and Superdex 200 SEC.⁴ A 4 step purification pathway comprising 70 °C heating, Q Sepharose IEC, ultra-centrifugation and AS precipitation was adopted in two previous studies.^{5,6} In other research, lysate heating followed by AS precipitation, plus DNase and RNase treatment and SEC were performed to remove HCPs and nucleic acid.⁷⁻⁹ Currently, most HF_n purification process is long and consists of 3 to 4 steps. Time-consuming steps such as ultracentrifugation and SEC, and cost-inefficient step as DNase and RNase treatment are also frequently included.

The purification process of recombinant HF_n was generally directly adopted in the purification of functionalised ferritin.^{10,11} However, in some cases, adjustments in purification pathway of functionalised HF_ns are reported but the mechanisms were not clearly investigated. For example, IEC was used instead of heating at 60 °C for 15 min for purification of GE11 peptide fused ferritin before SEC.¹² This is probably due to a thermal stability decrease after fusion. The concentration of AS in AS precipitation decreased to 30%-50 % from 40 %-60 % saturation in mutated HF_n purification.¹³ It is likely to result from the hydrophobicity

difference caused by the mutation. Overall, the adjustments in purification pathway are closely related to the genetic modifications. Although PAS and RGDK peptide do not affect HF_n assembly in *E. coli*, as has been shown in **Chapter 3**, it is possible that these would influence HF_n biochemical properties and affect purification performance.

In terms of molecular structures of functionalised HF_ns, PAS peptide can bind to the surrounding water molecules and lead to the increase of HF_n hydrodynamic volume.¹⁴ Structure comparison are needed to confirm this. Additionally, the confirmation of assembly conformance maintenance of all HF_n-based proteins after purification, and the investigation of RGDK functionalisation impact are necessary.

In this chapter, different purification processes of HF_n and functionalised HF_ns were developed and structures of purified HF_n and 3 PAS functionalised HF_ns were compared. In purification, a heat-acid precipitation followed by chromatography process was explored to remove HCPs and nucleic acid from HF_n. HF_n purification pathway was also applied in 3 functionalised HF_ns (HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK) to compare purification performance of these with HF_n. In the first step, heat-acid precipitation, HF_n, HF_n-PAS and HF_n-GFLG-PAS-RGDK were studied as examples to screen temperature and pH factors. In the second step, Q fast flow (FF) IEC and hydrophobic interaction chromatography (HIC) were compared in nucleic acid removal efficiency of HF_n and functionalised HF_ns. However, the final purity of PAS functionalised HF_ns using HF_n purification pathway were undesirable. Therefore, mono Q IEC was also explored after heat-acid precipitation for functionalised HF_n purification.

Structures of purified HF_n and 3 PAS functionalised HF_ns were compared on multiple levels using analytical techniques inclusive of Polyacrylamide Gel Electrophoresis (PAGE), Circular Dichroism (CD), Intrinsic Fluorescence spectroscopy (IF), Dynamic Light Scattering (DLS), Transmission Electron Microscopy (TEM) and High-Performance Size Exclusion Chromatography coupled with Multiple Angle Laser Light Scattering (HPSEC-MALLS).

4.2 Materials and methods

4.2.1 Materials

E. coli lysate supernatants containing HF_n and 3 PAS functionalised HF_ns were prepared as was described in **Chapter 3**, and used as the initial purification samples in this chapter. All materials involved in preparation of the *E. coli* lysate supernatants are listed in **Chapter 3, Section 3.2.1**. All chromatography columns used in this chapter were bought from GE Healthcare (USA). All reagents of analytical grade were purchased from Chem-Supply (Australia). Milli Q water was obtained from a Millipore purification system (Merck, USA) and used throughout the whole process.

4.2.2 Comparison of host cell proteins (HCPs) removal by heat-acid precipitation

A 2-step purification, heat-acid precipitation followed by chromatography, was explored for HF_n and 3 PAS functionalised HF_ns.

To obtain a high protein purity and protein recovery, buffer pH and heating temperature in heat-acid precipitation of HF_n, HF_n-PAS and HF_n-GFLG-PAS-RGDK were screened to find the optimal conditions. This optimal condition of HF_n-GFLG-PAS-RGDK was also applied to HF_n-PLGLAG-PAS-RGDK. Briefly, 0.5 mL *E. coli* lysate supernatant (8 mg mL⁻¹) was mixed at a volume ratio of 1:1 (v/v) with acidic buffer, 200 mM acetic acid-sodium acetate, 2 M NaCl, with pH of 4.0, 4.5 and 5.0. Then the mixtures were heated at 50 or 60 °C for 5 min. Samples were mixed with a stir bar at 100 rpm throughout the heating process. Following the heating, the mixture was centrifuged at 12,000 rpm, 4 °C for 15 min to remove precipitated proteins. Supernatant protein concentrations after centrifugation were measured through Bradford assay. The protein compositions in supernatants after centrifugation was analysed by 12 % reducing SDS-PAGE (Bio-Rad, USA). Densitometry scan was used to determine the purity of target protein after purification. Software Image J was adopted.¹⁵ Recovery yield (step) was calculated using **Equation 4.1**. In heat-acid precipitation, ‘the amount of target protein at purification stage n’ means the amount of target protein in heat-acid precipitation supernatant. ‘The amount of target protein at purification stage n-1’ means the target protein amount in sample before heat-acid precipitation.

$$\text{Recovery yield (step) (\%)} = 100 \% \times \frac{\text{the amount of target protein at purification stage } n}{\text{the amount of target protein at purification stage } n-1}$$

4.2.3 Nucleic acid removal by Q FF IEC

Q FF IEC and HIC were compared in the second step to remove nucleic acid contaminants. AKTA pure (GE Healthcare, USA) was employed for all following chromatography.

In regard to IEC, Hitrap Q fast flow (FF) column (1 mL, GE Healthcare, USA), was explored for HF_n, HF_n-PAS and HF_n-GFLG-PAS-RGDK. Three pH conditions were examined, pH 7, 8 and 9. Equilibration buffers and elution buffers at pH 7 and 8 were 20 mM phosphate buffer (PB) without and with 1 M NaCl. 20 mM Tris-HCl without and with 1 M NaCl buffers were pH 9 equilibration and elution buffer. Supernatants obtained after heat-acid precipitation underwent Hitrap G25 desalting chromatography (5 mL, GE Healthcare, USA) into corresponding equilibration buffer before being loaded onto Q FF column. Protein loading amount in each chromatography run was approximate 10 mg. After sample loading and flow through (FT) peak, the column was eluted from 0 - 1 M NaCl linearly within 5 column volume (CV). Absorbance at 280 and 260 nm were recorded. Flow rate was 1 mL min⁻¹. Collected peaks underwent Bradford protein concentration determination (Bio-Rad, USA) and Quan-iT 1 × dsDNA HS assay kit nucleic acid concentration determination (Invitrogen, Thermo Fisher Scientific, USA) and 12% reducing SDS-PAGE (Bio-Rad, USA) analysis. Purity of target peaks were obtained from densitometry scan and protein recovery yields (step) were calculated as in **Equation 4.1**. In chromatography step, ‘the amount of target protein at purification stage n’ means the amount of target protein in peak. ‘The amount of target protein at purification stage n-1’ means the target protein amount in loading sample. Nucleic acid removal ratio of the optimal peaks was calculated as in **Equation 4.2**:

Nucleic acid removal ratio (%) =

$$100 (\%) \times \frac{\text{loading sample nucleic acid amount (mg)} - \text{final product nucleic acid amount (mg)}}{\text{loading sample nucleic acid amount (mg)}}$$

(4.2)

4.2.4 Nucleic acid removal by HIC

HIC screening was done for HF_n and HF_n-GFLG-PAS-RGDK. The optimal condition was applied to the other 2 proteins HF_n-PAS and HF_n-PLGLAG-PAS-RGDK. Prior to HIC, the supernatants after heat-acid precipitation were diluted 5 times with 100 mM phosphate buffer

(PB), 1.2 M AS, incubated at 4 °C for 0.5 h and then filtered using 0.45 µm membrane. During HIC, absorbance at 260 and 280 nm was recorded. Hitrap Octyl FF or Hitrap Butyl FF column (1 mL, GE Healthcare, USA) was equilibrated with 100 mM PB, 1.0 M AS, pH 6.5 before loading the diluted supernatant. After FT peak finished, column was washed with 0 - 100 % elution buffer (20 mM PB, pH 6.5), 1 CV gradient. Flow rate was 1 mL min⁻¹. Peaks collected from HIC were collected and protein concentrations were determined using Bradford assay. Protein purities in peaks were analysed using densitometry scan of SDS-PAGE gel. Recovery yield (overall), the recovery yield after heat-acid precipitation and chromatography, was calculated as in **Equation 4.3**:

$$\text{Recovery yield (overall) (\%)} = 100 \% \times \frac{\text{the amount of target protein after chromatography}}{\text{the amount of target protein in bacterial lysate (supernatant)}} \quad (4.3)$$

Purity, protein recovery yield (step), and nucleic acid removal ratio were also calculated.

4.2.5 Scale up purification of HF_n

In a larger scale of HF_n purification, sample volume for heat-acid precipitation increased to 100 mL. Prior to HIC, supernatant after heat-acid precipitation was diluted 5 times using 100 mM PB, 1.2 M AS, placed at 4 °C for 30 min and then centrifuged at 12,000 rpm, 4 °C for 30 min. In HIC, a 20 mL XK16 self-packed Butyl FF column (GE Healthcare, USA) and AKTA pure (GE Healthcare, USA) were used. Buffers and loading sample treatment were exactly the same as in small scale Butyl FF purification. Elution pattern was conducted with one step using 100 % elution buffer. Flow rate was 3 mL min⁻¹. Eluted peaks were collected and analysed by reducing 12 % SDS-PAGE and Bradford assay. Purity, protein recovery yield (overall) for heat-acid precipitation and HIC, and the nucleic acid removal ratio in HIC were calculated.

4.2.6 Mono Q purification of functionalised HF_ns after heat-acid precipitation

The supernatants after heat-acid precipitation of 3 functionalised HF_ns underwent buffer exchange to 20 mM PB, pH 7 using Hitrap G25 desalting column (GE Healthcare, USA). Mono Q 5/50 GL column (GE Healthcare, USA) was equilibrated with 20 mM PB, pH 7 and loaded with approximate 10 mg desalted sample. When flow through peak finished, the column was washed with a linear gradient from 0- 0.5 M NaCl, 8CV, followed by another 2 steps of washing

using 1 M NaCl and 2 M NaCl, respectively. Absorbance at 280 and 260 nm were recorded. Purity and protein recovery yield of target peaks were calculated.

4.2.7 Secondary and tertiary structural comparison

Purified HF_n-based proteins (HF_n, HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK) were buffer exchanged into 20 mM PB, pH 7 through Hitrap G25 desalting chromatography (GE Healthcare, USA). Protein concentrations were adjusted to 0.2 mg mL⁻¹. CD spectroscopy was measured on a J-810 spectrometer (Jasco, Japan) at 25 °C using a 1.0 mm path length quartz cuvette. The wavelength scanning was conducted from 260 to 190 nm at a rate of 500 nm min⁻¹ with a bandwidth of 1 nm. The average of 3 scans of each sample was presented. IF spectroscopy was performed on F-4500 fluorescence spectrophotometer (Hitachi, Japan). The excitation wavelength was 280 nm, and the emission was recorded from 300 to 400 nm with a scanning rate of 1200 nm min⁻¹. 1.0 cm path length cuvette was used. Each sample was also subjected to scan for 3 times.

4.2.8 Nanoparticle structural comparison

Purified HF_n, HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK were buffer exchanged into 20 mM PB, pH 7 through Hitrap G25 desalting column (GE Healthcare, USA) for nanoparticle structural characterisations by DLS, 5 % native-PAGE, HPSEC-MALLS and TEM.

DLS aimed to characterise protein sizes and zeta potentials. These were measured on a Zetasizer Nano ZS90 (Malvern, UK). Before measurement, purified protein samples were adjusted to 0.5 mg mL⁻¹ and centrifuged at 10, 000 rpm, 4 °C for 20 min. Instrument was equilibrated at 25 °C and every sample was measured 3 times.

6 µg of each protein was loaded to 5 % native-PAGE gel and run at 90 V for 5 h in Bio-Rad electrophoresis set (USA). Loading buffer and gel pH were 6.8.

HPSEC-MALLS was utilised for protein hydrodynamic volume comparison and molecular weight determination. In HPSEC-MALLS analysis, TSK G4000 SWxl column (Tosoh bioscience, Japan) was connected to high performance liquid chromatography (HPLC)

(Shimadzu, Japan) coupled with DAWN MALLS (Wyatt, USA) and Optilab refractive index (RI) detector (Wyatt, USA). Equilibration buffer was 20 mM PB, 0.1 M Na₂SO₄, pH 7.0. Flow rate was 0.8 mL min⁻¹. Absorbance of fractions at 280 nm was monitored. Equilibration buffer was 20 mM PB, pH 7.0. For TEM analysis, 5 µL of each purified HF_n-based protein were applied and the method was the same as in **Chapter 3, Section 3.2.6**.

4.3 Results and discussion

4.3.1 Optimisation of HCPs removal by heat-acid precipitation

Bacterial lysate supernatants containing HF_n and functionalised HF_ns underwent heat-acid precipitation to remove HCPs. **Figure 4.1** shows the SDS-PAGE images, and recovery yield and purity. **Table 4.1** lists the detailed purity and protein recovery yield (step) data. Because HF_n has a relatively high thermal and pH stability than most of HCPs, it remained in the supernatant after thermal and acidic treatment while HCPs precipitated (**Figure 4.1A**). As the decrease of pH, the amount of precipitated HCPs increased and thus the purity of HF_n increased. HF_n achieved the best purities under 50 °C pH 4.0 (64.77%) and 60 °C pH 4.5 (65.50 %) (**Figure 4.1A** and **D**). Because of the greater recovery yield (step) in 60 °C pH 4.5 (99.69 %) than in 50 °C pH 4.0 (88.70 %) (**Figure 4.1A** and **F**), 60 °C pH 4.5 was chosen as the first step purification of HF_n. Protein concentrations were calculated using the Bradford standard curve in **Figure B1 (Appendix B)**.

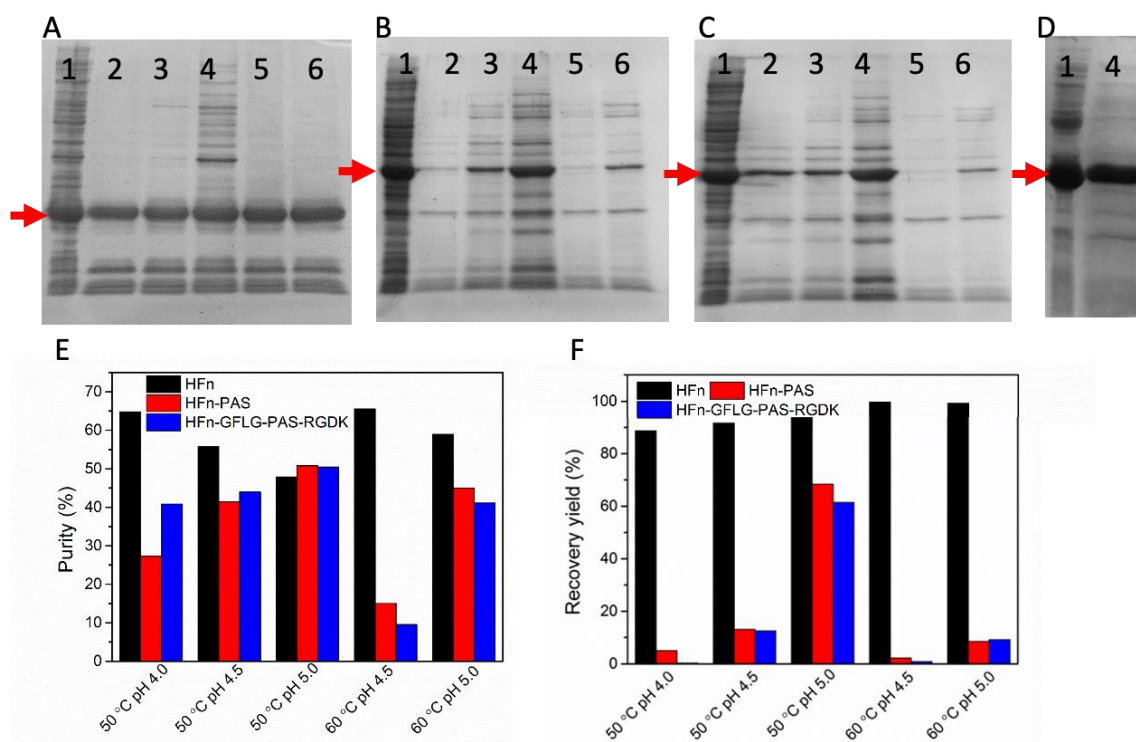


Figure 4.1 Heat-acid precipitation of HFn and functionalised HFns.

A, HFn SDS-PAGE image. **B**, HFn-PAS SDS-PAGE image. **C**, HFn-GFLG-PAS-RGDK SDS-PAGE image. **D**, HFn-PLGLAG-PAS-RGDK SDS-PAGE image. Lane 1, bacterial lysate supernatant; 2, 50 °C pH 4.0 heat-acid precipitation supernatant; 3, 50 °C pH 4.5 heat-acid precipitation supernatant; 4, 50 °C pH 5.0 heat-acid precipitation supernatant; 5, 60 °C pH 4.5 heat-acid precipitation supernatant; 6, 60 °C pH 5.0 heat-acid precipitation supernatant. **E**, Purity of HFn, HFn-PAS and HFn-GFLG-PAS-RGDK in heat-acid precipitation. **F**, Protein recovery yield (step) of HFn, HFn-PAS and HFn-GFLG-PAS-RGDK in heat-acid precipitation.

Table 4.1 Purity and protein recovery yield (step) in heat-acid precipitation.

Protein	Conditions	50 °C pH	50 °C pH	50 °C pH	60 °C pH	60 °C pH
		4.0	4.5	5.0	4.5	5.0
HFn	Purity (%)	64.77	55.77	47.87	65.50	58.92
	Recovery yield (%)	88.70	91.58	93.82	99.69	99.22
HFn-PAS	Purity (%)	27.31	41.43	50.77	15.02	45.01
	Recovery yield (%)	4.98	13.14	68.41	2.26	8.56
HFn-GFLG-PAS-RGDK	Purity (%)	40.85	43.98	50.41	9.50	41.10
	Recovery yield (%)	0.32	12.55	61.45	0.88	9.19
HFn-PLGLAG-PAS-RGDK	Purity (%)	/	/	59.63	/	/
	Recovery yield (%)	/	/	66.78	/	/

Interestingly, 3 functionalised HFns show the increasing purity with the increase in pH (**Figure 4.1E**). The reason is that compared with HFn, 3 functionalised HFns have a significantly

weaker resistance against the combination of acidic pH and heat. At the conditions of pH 4.0, 50 °C and pH 4.5, 60 °C, functionalised HFns cannot resist to the acid condition and precipitate together with HCPs. No obvious target protein can be observed in supernatant indicated by **Figure 4.1B** and **C** lane 2 and 5. The purities, therefore, were very low. As pH increased to the level which modified HFns could resist, the modified HFns remained in supernatant, whilst the majority of HCPs precipitated and the purity of modified HFns increased. Recovery yield (step) of functionalised HFns was in the range of 1% to 15 % under all tested conditions, except 50 °C pH 5.0 (above 60 %), in contrast with that of HFn, above 85 % at all tested conditions (**Figure 4.1F**). Particularly, at 60 °C pH 4.5, recovery yield (step) of HFn was around 100 %.

This significant stability drop results from the C-terminal insertion of foreign peptides. Stability decrease after insertion of foreign peptide has been reported several times in different proteins.¹⁶⁻¹⁸ Comparing HFn-PAS with HFn-GFLG-PAS-RGDK, recovery yield differences under all conditions were within 5 %, showing that the short RGDK peptide did not significantly affect stability, and the property change mainly stemmed from PAS peptide insertion.

4.3.2 Nucleic acid removal by Q FF IEC

In the heat-acid precipitation, many HCPs were removed, but most of the host cell nucleic acid remained in supernatants and co-existed with target proteins. Therefore, the second step focused on the nucleic acid removal.

Ion exchanger has been commonly applied to separate protein and nucleic acids based on the differences of their binding strengths.¹⁹ In this research, an anion exchanger, Q FF, was selected, because of the stability and the isoelectric point (PI) of HFn and functionalised HFns. HFn nanocage remains stable in pH range of 3.4-10.²⁰ Functionalised HFns have a significant lower tolerance against acidic pH, as is observed in heat-acid precipitation. HFn PI is around 4.5-5.0, and the inserted PAS peptide theoretical PI is around 5.2, calculated by ExPASy PI calculation tool (Swiss Institute of Bioinformatics).²¹ Therefore, it is very likely that the functionalised HFns and HFn nanocage will be damaged to some degree, if using cation exchanger, because pH in use would be at least below 4.0. As a result, Q FF and pH 7, 8 and 9 were selected to separate nucleic acid and HFn based on binding strength difference.

Figure 4.2 shows SDS-PAGE images of HF_n, HF_n-PAS and HF_n-GFLG-PAS-RGDK at pH 7, 8 and 9. **Figure 4.3** shows the Q FF chromatograms and the purity, recovery yield and nucleic acid removal. **Table 4.2** lists all the purity, protein recovery yield and nucleic acid removal ratio of peaks. Nucleic acid concentrations of peaks were calculated using the standard curve in **Figure B2 (Appendix B)**. Theoretically, HF_n and functionalised HF_ns can bind to Q FF because of their acidic PI. As is shown in **Figure 4.3J**, surprisingly, in all tested pH values, the majority (65 %-92 %) of 3 HF_n-based proteins were in FT peaks, especially HF_n. This shows a very weak binding of HF_n and functionalised HF_ns to Q FF column. All FT and elution peaks had greater OD₂₆₀ than OD₂₈₀ values, which indicates the presence of nucleic acid in both FT and elution peaks (**Figure 4.3A-I**). Therefore, the nucleic acid removal was not satisfactory. In pH 9 FT peaks, nucleic acid removal was 46.00 % of HF_n, 65.17 % of HF_n-PAS and 46.81 % of HF_n-PAS-RGDK.

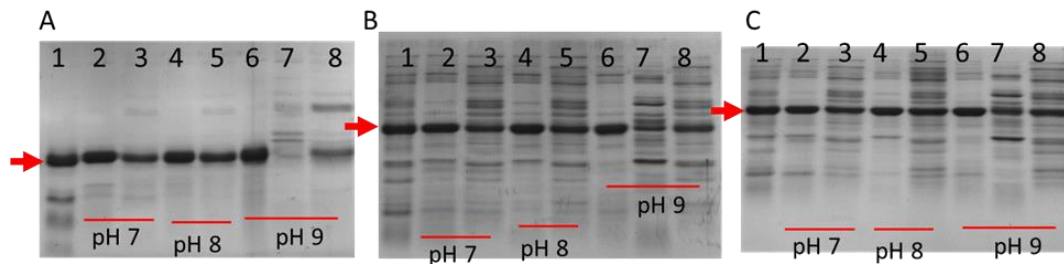


Figure 4.2 12 % reducing SDS-PAGE images of peaks collected from Q FF column. **A**, 12 % reducing SDS-PAGE results of HF_n Q FF chromatography. **B**, 12 % reducing SDS-PAGE results of HF_n-PAS Q FF chromatography. **C**, 12 % reducing SDS-PAGE results of HF_n-GFLG-PAS-RGDK Q FF chromatography. Lane 1, heat-acid precipitation supernatant; 2, pH 7 flow through peak; 3, pH 7 eluted peak 1; 3, pH 8 flow through peak; 4, pH 8 eluted peak 1; 5, pH 9 flow through peak; 6, eluted peak 1; 7: pH 9 eluted peak 2.

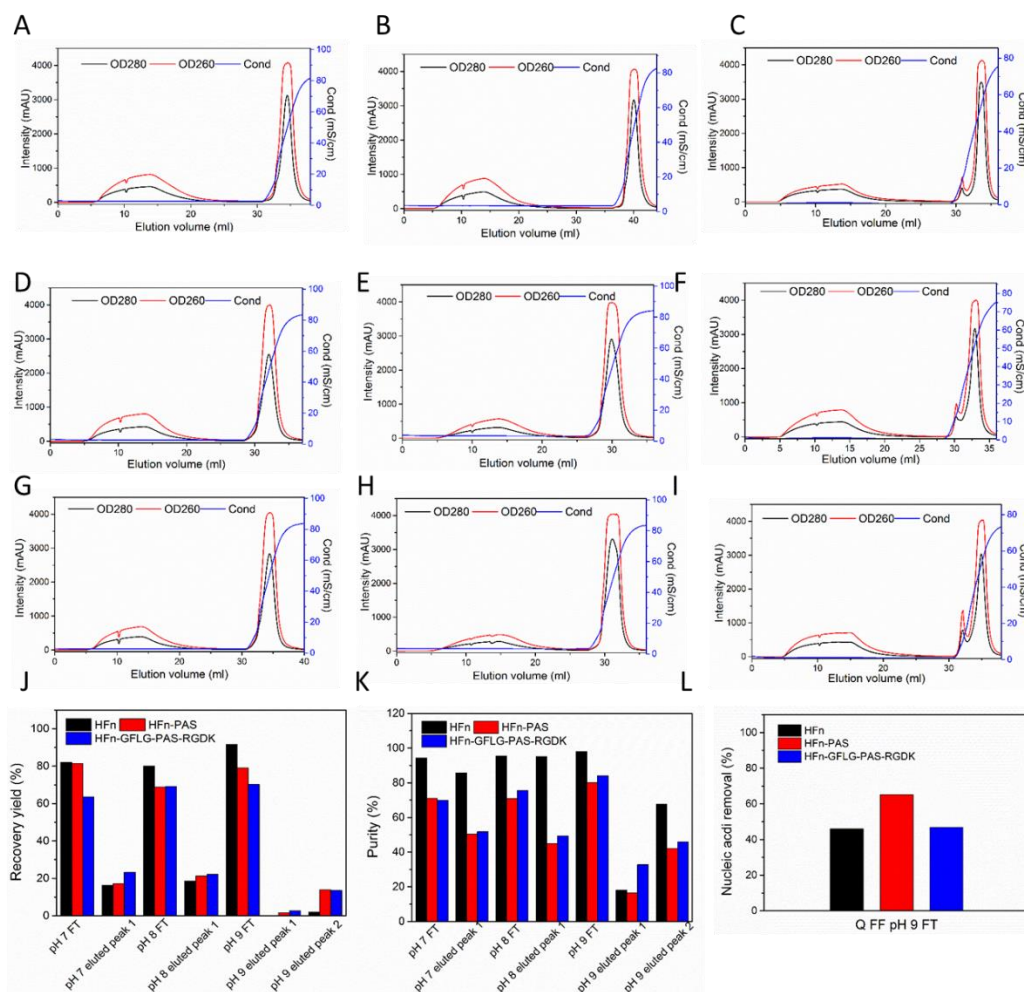


Figure 4.3 Q FF chromatography purification of HFn, HFn-PAS and HFn-GFLG-PAS-RGDK. **A**, pH 7 HFn chromatogram. **B**, pH 8 HFn chromatogram. **C**, pH 9 HFn chromatogram. **D**, pH 7 HFn-PAS chromatogram. **E**, pH 8 HFn-PAS chromatogram. **F**, pH 9 HFn-PAS chromatogram. **G**, pH 7 HFn-GFLG-PAS-RGDK chromatogram. **H**, pH 8 HFn-GFLG-PAS-RGDK chromatogram. **I**, pH 9 HFn-GFLG-PAS-RGDK chromatogram. **J**, recovery yield of all peaks from Q FF chromatography. **K**, purity of all peaks from Q FF chromatography. **L**, nucleic acid removal of pH 9 FT peaks.

Table 4.2 Purity, recovery yield (step) and nucleic acid removal of all peaks in Q FF chromatography.

Protein	Q FF Peak	pH 7	pH 7 P1	pH 8	pH 8 P1	pH 9	pH 9 P1	pH 9 P2
		FT	FT	FT	FT	FT	FT	FT
HFn	Purity (%)	94.27	85.69	95.56	95.08	97.98	18.00	67.73
	Recovery yield (%)	82.12	16.34	80.07	18.47	91.68	0.14	2.01
	Nucleic acid removal (%)	/	/	/	/	46.00	/	/
HFn-PAS	Purity (%)	70.96	50.45	70.95	44.95	80.13	16.41	42.06
	Recovery yield (%)	81.40	17.17	68.93	21.33	79.06	1.59	13.83
	Nucleic acid removal (%)	/	/	/	/	65.17	/	/
	Purity (%)	69.92	51.72	75.60	49.40	83.96	32.75	45.96

HFn-GFLG-	Recovery yield (%)	63.47	23.17	69.17	22.21	70.25	2.67	13.55
PAS-RGDK	Nucleic acid removal (%)	/	/	/	/	46.81	/	/

10 mg protein loaded in each run. FT: flow through peak. '/': no measurement.

The phenomenon of nucleic acid appearing in FT and elution fractions indicates the non-uniform charge distribution, which may be caused by the nature of digested nucleic acid fragments resulting from heat-acid condition. The observation that all 3 HFn-based proteins appeared in FT peaks in tested pH shows that nanometre sized proteins are hard to diffuse inside resin spheres to interact with the functional groups on resin and the binding is generally weaker compared with small sized proteins, because of the lower value of surface net charge/weight.^{22, 23} What is more, HFn is more negatively charged in inner surface than outer surface because more acidic residues were buried inside HFn assembly.²⁴

pH 9 FT peak had the highest purity and recovery yield (**Table 4.2** and **Figure 4.3J** and **K**), compared with pH 7 and 8. This is probably because the increase of binding affinity from pH 7 to 9 of the contaminants was greater than that of target proteins.

In comparison of HFn with functionalised HFns, the Q FF column binding strength of functionalised HFns was greater than that of HFn, according to the protein recovery yield (step) of eluted peaks in **Table 4.2** and **Figure 4.3J**. The difference of recovery yield (step) in comparison of HFn-PAS and HFn-GFLG-PAS-RGDK, was smaller than the difference between HFn and HFn-PAS. This is perhaps because PAS peptide is much longer than RGDK peptide and makes a stronger impact on HFn. PAS peptide theoretically bears negative charges at pH 8 and 9 and is primarily responsible for charge difference between HFn and functionalised HFns.

4.3.3 Nucleic acid removal by HIC

In HIC, the strategy is the binding of target proteins and the flow through of nucleic acids. Nucleic acids are highly hydrophilic and hypothetically cannot bind to hydrophobic interaction columns. HFn, by contrast, has been proven to be able to bind to hydrophobic interaction column in a previous study.²⁵ **Figure 4.4** is the SDS-PAGE images of peaks form HIC. **Figure 4.5** lists the chromatograms and nucleic acid removal, purity and recovery yield (step) in HIC. **Table 4.3** list the detailed results of HIC. As shown in **Figure 4.5A-D**, OD260 nm of the FT

peaks from butyl and octyl columns were greater than twice of OD280, suggesting they were mainly composed of nucleic acids. The eluted peaks containing target proteins, on the contrary, had greater OD280 than OD260. Both butyl and octyl FF achieved an above 98 % nucleic acid removal for HFn and HFn-GFLG-PAS-RGDK (**Figure 4.5E** and **Table 4.3**), which was significantly greater than in Q FF IEC. This indicates HIC is very efficient in nucleic acid removal.

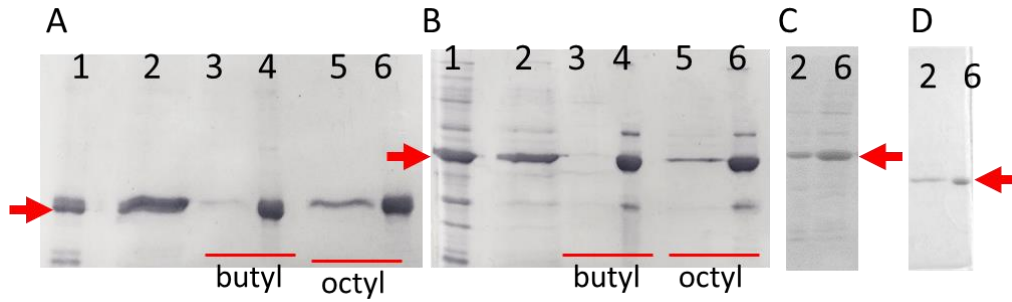


Figure 4.4 SDS-PAGE images of HFn and 3 PAS functionalised HFns after HIC. **A**, 12 % reducing SDS-PAGE results of peaks from HFn-GFLG-PAS-RGDK HIC chromatography. **B**, 12 % reducing SDS-PAGE results of peaks from HFn-PLGLAG-PAS-RGDK butyl FF chromatography. **C**, 12 % reducing SDS-PAGE results of peaks from HFn-PAS butyl FF chromatography. **D**, 12 % reducing SDS-PAGE results of peaks from HFn-PLGLAG-PAS-RGDK butyl FF chromatography. Lane 1, heat-acid precipitation supernatant; 2, Loading sample for HIC; 3, flow through peak of butyl FF; 4, eluted peak of butyl FF; 5, flow through peak of octyl FF; 6, eluted peak of octyl FF.

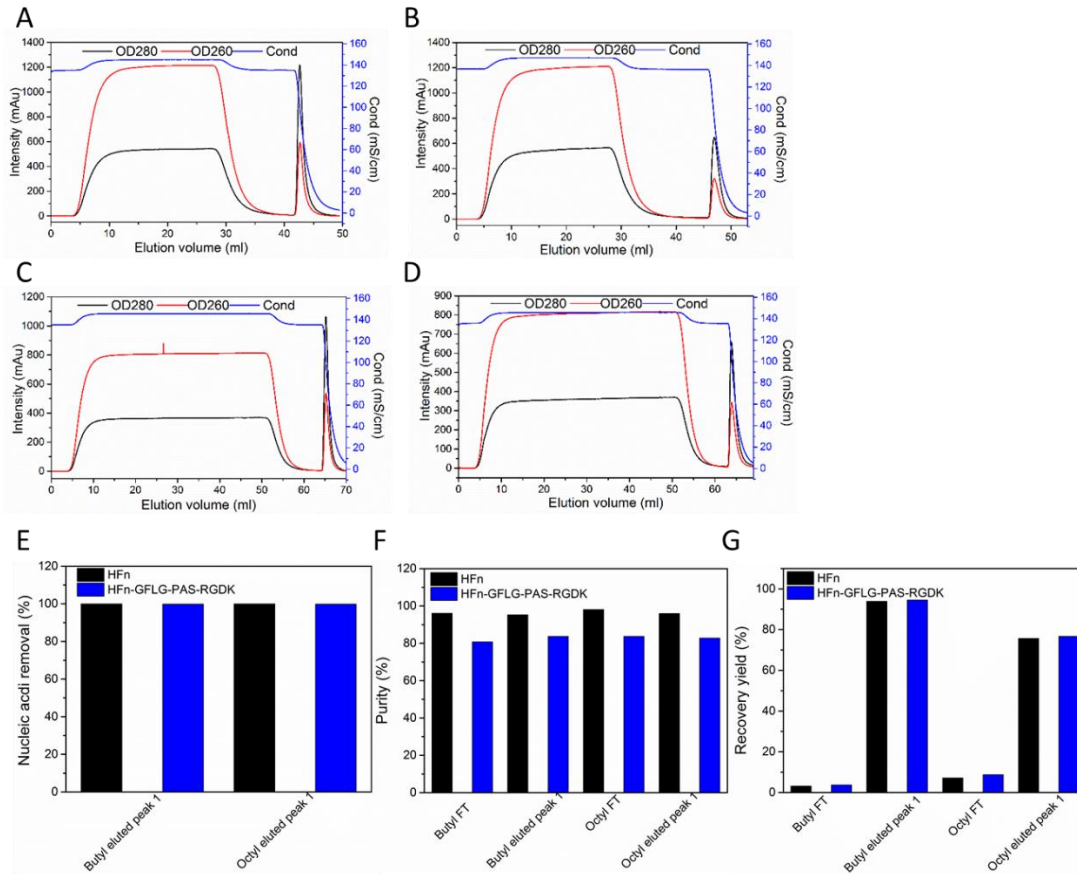


Figure 4.5 HIC chromatography purification for HFn and functionalised HFns. **A**, HFn butyl chromatogram. **B**, HFn octyl chromatogram. **C**, HFn-GFLG-PAS-RGDK butyl chromatogram. **D**, HFn-GFLG-PAS-RGDK octyl chromatogram. **E**, nucleic acid removal in HIC peaks. **F**, purity in HIC peaks. **G**, recovery yield (step) in HIC peaks.

Table 4.3 Purity, protein recovery yield (step) and nucleic acid removal in HIC.

Protein	Chromatography peak	Butyl FF FT	Butyl FF P1	Octyl FF FT	Octyl FF P1
HFn	Purity	96.09	95.24	98.02	95.94
	Recovery yield (%)	3.18	93.87	7.13	75.65
	Nucleic acid removal (%)	/	99.93	/	99.95
HFn-PAS	Purity	83.48	82.32	/	/
	Recovery yield (%)	5.05	93.46	/	/
	Nucleic acid removal (%)	/	99.81	/	/
HFn-GFLG-PAS-RGDK	Purity	80.78	83.74	83.75	82.83
	Recovery yield (%)	3.68	94.54	8.79	76.66
	Nucleic acid removal (%)	/	99.81	/	99.78
HFn-PLGLAG-PAS-RGDK	Purity	85.88	85.03	/	/
	Recovery yield (%)	4.26	92.64	/	/

Purity of FT peak is not as highly accurate as eluted peak purity because of the much lighter bands in gel. ‘/’: no measurement.

In terms of purity, the 100 mM PB 1.2 M AS 5 × time dilution treatment of supernatant after heat-acid precipitation removed part of HCPs in both HF_n and HF_n-GFLG-PAS-RGDK (**Figure 4.4A** and **B**). Purity after HIC was above 95 % of HF_n. However, purity of HF_n-GFLG-PAS-RGDK was below 90 % (**Table 4.3**). This is primarily because 1) HIC was not effective in HCPs removal but focused on nucleic acid removal and 2) purity of HF_n-GFLG-PAS-RGDK after heat-acid precipitation was lower than that of HF_n due to the lower purity in heat-acid precipitation, which results from the decreased stability.

Comparing recovery yields (step) of functionalised HF_ns and HF_n in HIC (**Figure 4.5G** and **Table 4.3**), there is no significant differences. This shows that the inserted PAS and RGDK peptide did not affect HF_n HIC behaviour. This is probably because 1) both the inserted functional peptides were at C-terminal of the E-helix, and residues constituting the nanocage outer surface undergo no changes, and 2) both functional peptides are random coils and highly hydrophilic, which do not interact with resins.

In comparison of butyl and octyl resins, the total recovery yield of eluted peak plus flow through peak from butyl was greater than from octyl (**Table 4.3**). Some tightly bound target proteins were eluted in column washing step (chromatogram not shown). However, flow through peaks of octyl column contained more target protein than butyl column. This shows that HF_n and HF_n-GFLG-PAS-RGDK behaved more diversely in octyl column compared with in butyl column. Because of the greater protein recovery yield in eluted peak and more uniform binding behaviour, butyl HIC was considered better than octyl. It was tried in HF_n-PAS and HF_n-PLGLAG-PAS-RGDK and SDS-PAGE images are shown in **Figure 4.4C** and **D**. Chromatograms of HF_n-PAS and HF_n-PLGLAG-PAS-RGDK were the same as that of HF_n-GFLG-PAS-RGDK (not shown).

In summary, purity of HF_n after heat-acid precipitation followed by butyl FF HIC was above 95 % but purities of functionalised HF_ns were below 90 %. Recovery yields (overall) were: 92.98 % (HF_n), 63.93 % (HF_n-PAS), 58.09 % (HF_n-GFLG-PAS-RGDK) and 61.86 % (HF_n-PLGLAG-PAS-RGDK), respectively. Therefore, heat-acid precipitation followed by butyl FF

HIC was effective in HF_n but failed to achieve desirable purification performance in functionalised HF_ns. Other purification methods should be explored to improve functionalised HF_n purity. The final HF_n amount per gram of wet *E. coli* cell was 32.25 mg. Nucleic acid removal by butyl FF HIC decreased nucleic acid amounts from 2000-3000 µg in loading sample to 2-4 µg in the final products.

4.3.4 Scalable HCPs and nucleic acid removal pathway of HF_n

Figure 4.6 is the butyl FF chromatogram and 12 % reducing SDS-PAGE result. In this scale up of purification of HF_n, heat-acid precipitation and dilution step has increased HF_n purity to 97.25 %. After the butyl chromatography, the final HF_n purity reached 98.52 %, greater than that in small scale. HF_n recovery yield (overall) was 98.18 %. Nucleic acid removal ratio was 99.97 %. Remaining nucleic acid concentration was 0.58 µg mL⁻¹. These results show that the two-step purification pathway is scalable for HF_n purification.

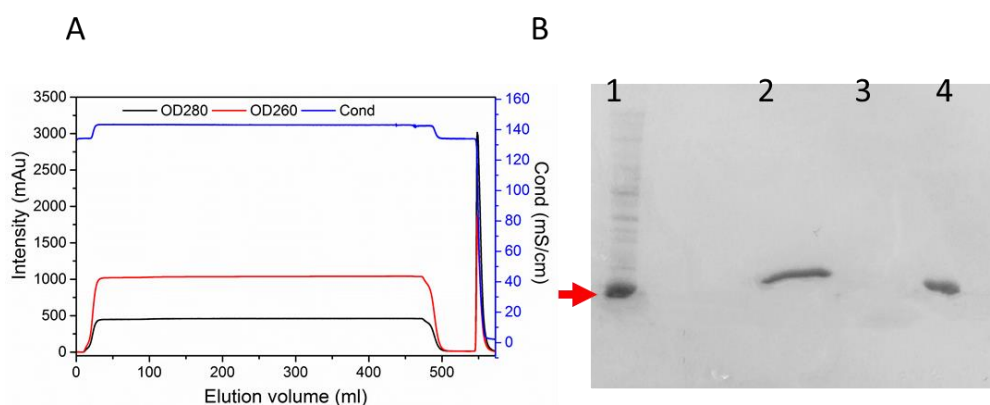


Figure 4.6 Large-scale butyl chromatogram and SDS-PAGE image.

A, butyl chromatogram. **B**, 12 % reducing SDS-PAGE image. Lane 1, cell lysate supernatant; 2, loading sample for butyl FF chromatography; 3, flow through peak of butyl FF; 4, eluted peak of butyl FF.

4.3.5 Functionalised HF_ns purification by Mono Q IEC after heat-acid precipitation

In order to obtain functionalised HF_ns with desirable purity, mono Q IEC was explored. **Figure 4.7** shows the chromatograms and SDS-PAGE images of purified functionalised HF_ns. Three functionalised HF_ns had similar mono Q chromatograms and were all eluted when the conductivity was around 7- 20 mS cm⁻¹, showing the same binding strength. As is shown in **Figure 4.7D**, purity of functionalised HF_ns reached above 95 % after mono Q chromatography. Consequently, it was selected as the second step of functionalised HF_n purification. The recovery yield (overall) after heat-acid precipitation followed by mono Q IEC were 59.27 % (HF_n-PAS), 55.45 % (HF_n-GFLG-PAS-RGDK) and 57.61 % (HF_n-PLGLAG-PAS-RGDK).

Final protein amounts per gram of biomass were 19.21 mg (HF_n-PAS), 19.30 mg (HF_n-GFLG-PAS-RGDK) and 18.87 mg (HF_n-PLGLAG-PAS-RGDK).

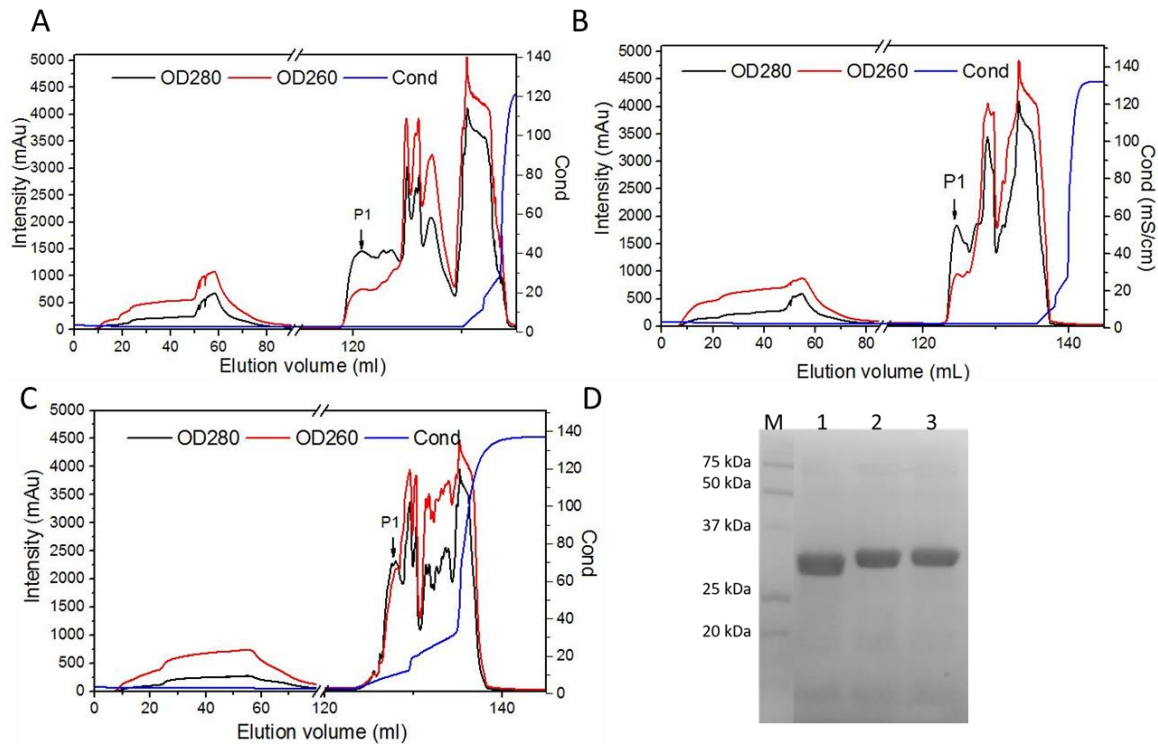


Figure 4.7 Mono Q chromatogram and SDS-PAGE image. **A**, HF_n-PAS chromatogram. **B**, HF_n-GFLG-PAS-RGDK chromatogram. **C**, HF_n-PLGLAG-PAS-RGDK chromatogram. P1 indicates the peak containing target protein. **D**, SDS-PAGE image of P1 from mono Q chromatography. Lane 1, HF_n-PAS. 2, HF_n-GFLG-PAS-RGDK. 3, HF_n-PLGLAG-PAS-RGDK.

4.3.6 Functionalisation effect on HF_n secondary and tertiary structure

Far-UV circular dichroism (CD) analysis aims to compare the secondary structures of PAS functionalised HF_ns with that of HF_n. In HF_n subunit, approximate 68 % residues form α -helix and the rest are non-structured loops.²⁴ As shown in **Figure 4.8A**, similar CD profiles indicate the protein conformation does not undergo significant secondary structural changes after fusion. Two negative bands at 208 and 222 nm prove that all 4 HF_n-based proteins are mainly α -helix. This is because flexible linker and functional peptides are random coils, and fused to one end of HF_n subunit, without disrupting subunit main structure of 4-helix bundle conformation.

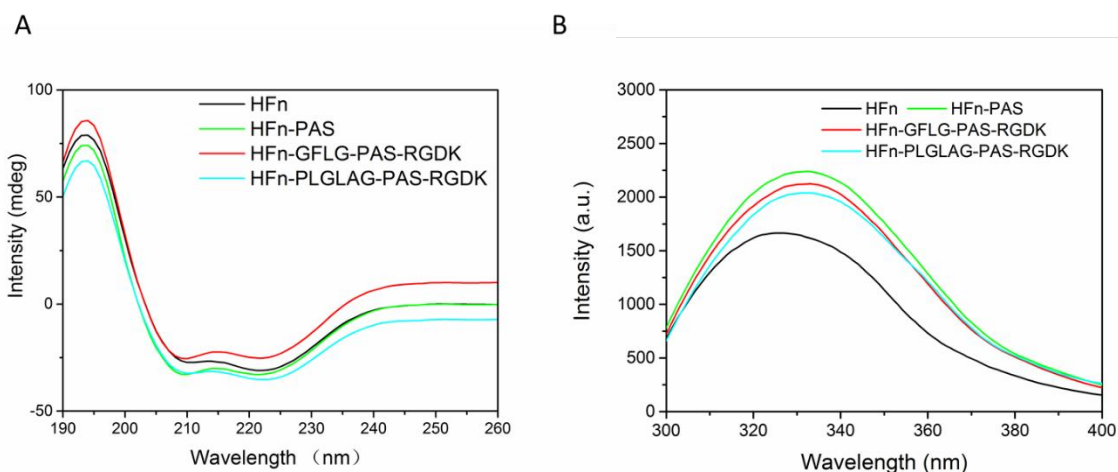


Figure 4.8 Circular dichroism spectra (A) and intrinsic fluorescence spectra of four HFn-based proteins (B).

However, in the protein intrinsic fluorescence spectra (**Figure 4.8B**), 3 PAS functionalised HFns (peaked at 333 nm) have a 7 nm red shift compared to HFn (326 nm). Protein intrinsic fluorescence spectrum is related to the microenvironment of Trp, Tyr and Phe residues.²⁶ This red shift is probably caused by the polarity increase of microenvironment, which results from a subtle conformational change in proteins.²⁷ This indicates that PAS functionalised HFns have conformational differences from HFn due to the C-terminal insertion.

4.3.7 Functionalisation effect on HFn nanoparticle structure

Protein particle sizes and zeta potentials were measured by DLS (**Figure 4.9** and **Table 4.4**). The particle sizes of HFn, HFn-PAS, HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK are 11.9, 13.7, 14.1 and 14.1 nm, which proves that 24 subunits of all 4 proteins self-assemble into nanocages. Inserted PAS and RGDK both increase protein particle size, and PAS leads to a larger size increase than RGDK. Surface charges of proteins detected in zeta potential arise primarily from ionization of surface groups, such as acidic and basic side chains of exposed amino acid residues and associated counter ions bounding to protein surface.²⁸ Under the same buffer condition, a minor zeta-potential difference between HFn and the PAS functionalised HFns was detected (**Table 4.4**).

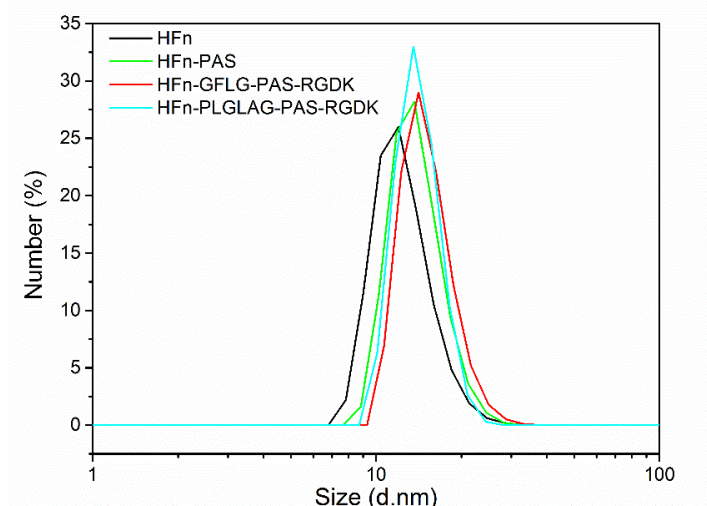


Figure 4.9 Size distributions of proteins from DLS

Table 4.4 Protein particle sizes and zeta potentials measured from DLS.

Protein	Size (nm)	Zeta potential (mV)
HFn	11.9 ± 0.1	-7.98 ± 0.23
HFn-PAS	13.3 ± 0.2	-5.73 ± 0.40
HFn-GFLG-PAS-RGDK	14.1 ± 0.2	-5.21 ± 0.22
HFn-PLGLAG-PAS-RGDK	14.1 ± 0.1	-5.20 ± 0.25

In Native-PAGE analysis, proteins remain in natural form and the migration distances depend on both particle size and surface negative charge density. HFn migrated fastest and HFn-PAS ranked second (**Figure 4.10**), followed by HFn-GFLG-PAS-RGDK and then HFn-PLGLAG-PAS-RGDK. The migration distance order is in alignment with the order of protein sizes. As detected in DLS, the charge difference at neutral pH was minor, so the size difference between all proteins was the main cause. All PAS-containing HFns show relatively fuzzy bands than HFn, which possibly be ascribed to the flexible non-structured functional peptides and linkers.

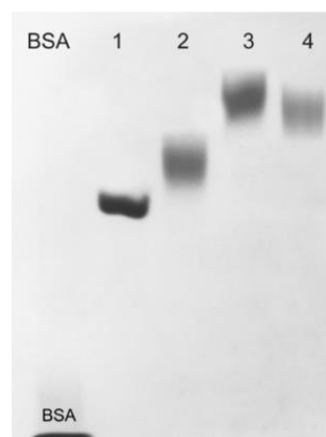


Figure 4.10 5 % Native-PAGE analysis results of purified proteins.
Lane 1: HFn, 2: HFn-PAS, 3: HFn-PLGLAG-PAS-RGDK, 4: HFn-GFLG-PAS-RGDK.

Figure 4.11 shows the HPSEC chromatograms of assembly expressed HF_n-based proteins and molecular weight distribution of main peaks in chromatograms. MALLS determined proteins molecular weights order was in accordance with theoretical order: HF_n-PLALAG-PAS-RGDK > HF_n-GFLG-PAS-RGDK > HF_n-PAS > HF_n, and average molecular weights of all proteins determined were similar to their theoretical Mw (**Table 4.5**). The order of the particle hydrodynamic radius follows the order detected in DLS.

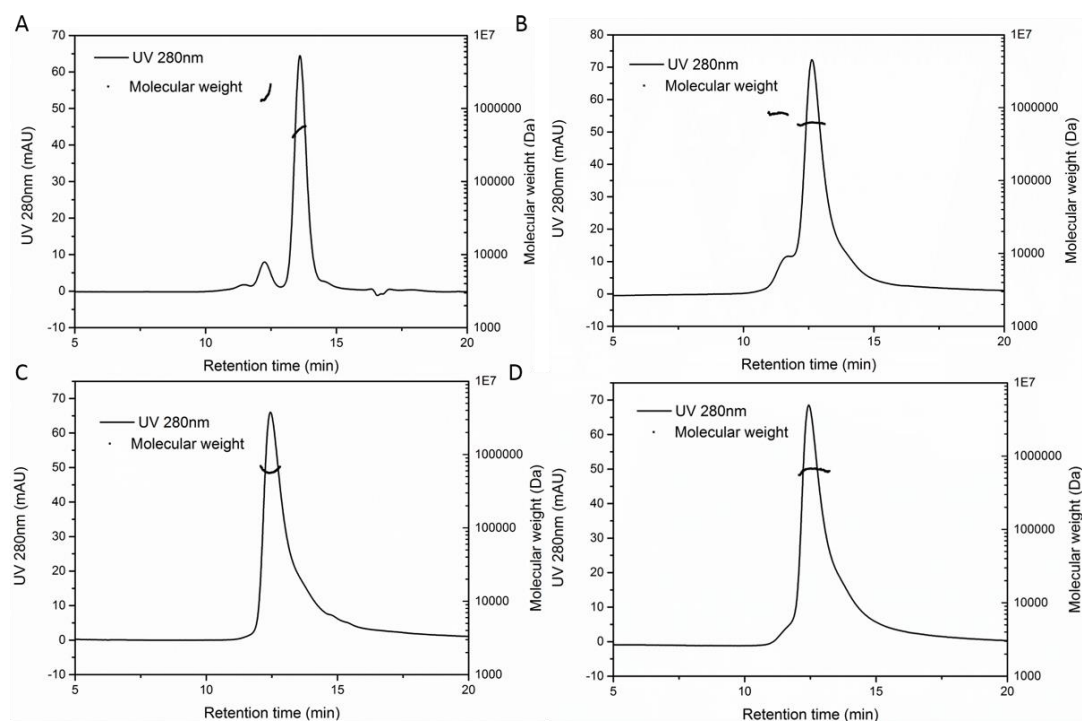


Figure 4.11 HPSEC-MALLS chromatograms of proteins. **A**, HF_n. **B**, HF_n-PAS. **C**, HF_n-GFLF-PAS-RGDK. **D**, HF_n-PLGLAG-PAS-RGDK.

Table 4.5 High performance size exclusion chromatography and multi-angle laser light scattering (HPSEC-MALLS) analysis result.

Protein	Particle hydrodynamic radius (nm)	Measured average molecular weight (kDa)	Theoretical molecular weight (kDa)
HF _n	6.31 (± 0.53%)	493.3(± 0.07%)	506.0
HF _n -PAS	7.72 (± 0.54%)	608.6(± 2.18%)	625.1
HF _n -GFLG-PAS-RGDK	8.06 (± 0.55%)	630.8(± 0.12%)	636.1
HF _n -PLGLAG-PAS-RGDK	8.22 (± 0.55%)	638.7(± 0.11%)	639.3

TEM analysis demonstrates that all PAS functionalised HFns are assembled hollow spheres, same as HFn. Cages of all proteins are around 12 nm in diameter no matter with or without functionalisation (**Figure 4.12**).

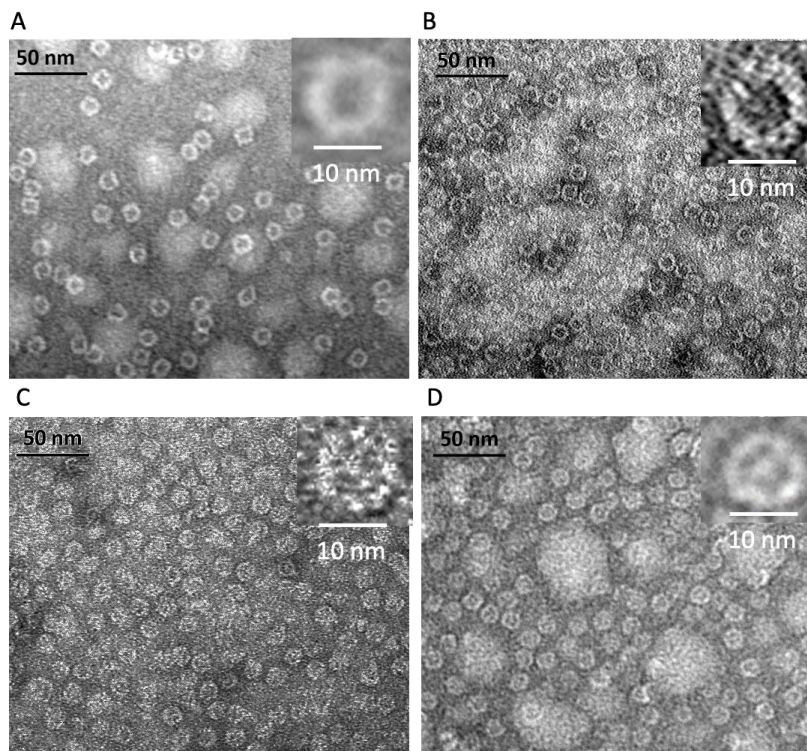


Figure 4.12 TEM images of purified proteins. **A**, HFn. **B**, HFn-PAS. **C**, HFn-GFLG-PAS-RGDK. **D**, HFn-PLGLAG-PAS-RGDK.

Combining the results in purification and structural characterisations, there emerged 2 interesting findings. 1) The 12 nm of functionalised HFns and HFn spheres observed under TEM contradict to the different hydrodynamic sizes detected in DLS and HPSEC-MALLS. 2) Binding strengths of modified HFns to Q FF column were generally higher than HFn although their predicted PI is higher than HFn PI, as the recovery yield (step) of Q FF eluted peaks of modified HFns was greater than that of HFn. We, therefore, reasonably speculate that some or all of the inserted peptides together with the E-helix are exposed outside HFn nanocage. In this case, the exposed PAS peptides can bind to surrounding water molecules, so the hydrodynamic size of modified HFns were enlarged whilst the nanocage diameters observed under TEM remained the same as HFn. Binding strength to Q FF column has to do with the local negative charge density of the potential binding sites. The exposed E-helices in modified HFns are probably responsible for the increase of the binding strength to Q FF column, because predicted PI of E-helix is 3.84. Predicted PI values of E-helix with PAS and E-helix with PAS-RGDK are 3.84 and 4.2, respectively. All these PI values are lower than HFn PI. This turnover of E-

helix with the C-terminal inserted peptide has been observed in previous work. HF_n with a red fluorescence protein (25.9 kDa) or 36 aa intrinsically disordered proteins at C-terminal haven been proven to be in flipped conformation.^{29, 30} The HF_n conformation with E-helix buried inside is the native conformation (as shown in **Figure 4.13A** and **B**). If all 24 E-helix are outside, HF_n is in the flipped conformation. Previously, a whole barrel-like red fluorescence protein (25.9 kDa) was fused to HF_n subunit at C-terminal and proven to be exposed outside HF_n cage.²⁹ If all 24 modified subunits in functionalised protein are extruded outside, theoretical structures of subunits and assemblies would be as shown in **Figure 4.13C** and **D**.

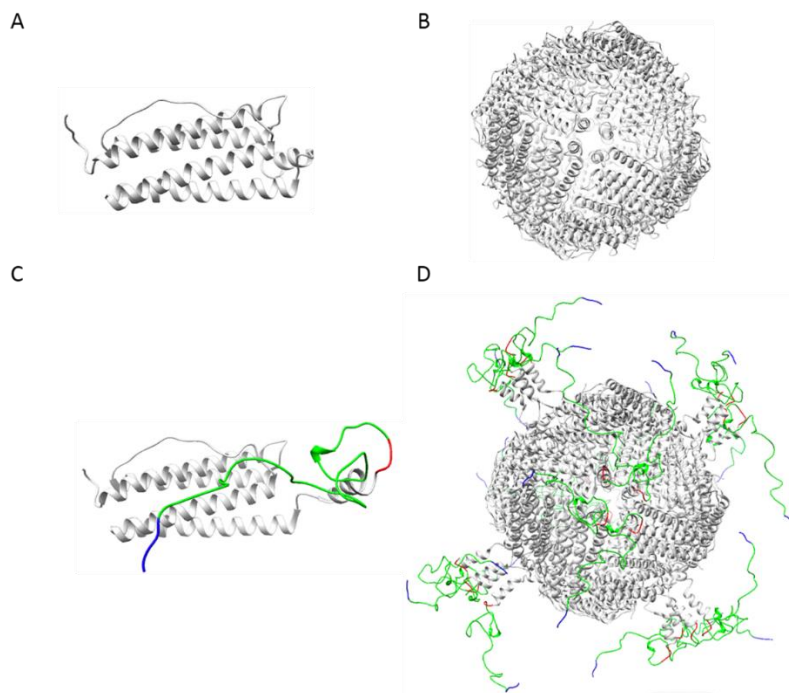


Figure 4.13 Schematic illustration of assumed protein structures.

A, native HF_n subunit. **B**, native HF_n assembly. **C**, flipped functionalised subunit (HF_n-GFLG-PAS-RGDK/HF_n-PLGLAG-PAS-RGDK/HF_n-PAS). **D** fully flipped assembly. Light grey parts represent HF_n, red parts represent enzyme-cleavable site GFLG or PLGLAG, green parts are PAS peptides and blue parts are RGDK peptides. Structure of HF_n was originally obtained from PDB (5N27) and then modified and visualized using Chimera.²⁴

Average volume of residues is about 100 Å³ and average residue molecular weight is about 100 Da. The inner cavity of ferritin is 8 nm across and thus it can only accommodate 24 of foreign peptides with maximum Mw of 9.2 kDa, when peptide structure and repulsive forces between peptides are not considered. Although in our case, inserted PAS-RGDK or PAS only weights 5-6 kDa, the water-binding ability and repulsive forces between PAS peptides may make it impossible for HF_n cavity to accommodate 24 inserted peptides simultaneously.

This turnover of the E-helix with the inserted functional peptide could be the cause for the structural difference detected in tertiary structure and the decrease of stability discovered in heat-acid precipitation. This is because, in native conformation, the hydrophobic interactions between 4 E-helices surrounding the hydrophobic channels have been proven to contribute to the stability of HF_n assembly.³¹ In the flipped conformation, however, the 4 E-helices are not physically confined as a bundle, but more likely to be floating around. Additionally, the inserted hydrophilic PAS peptide can reduce the hydrophobic attractions between E-helix.

4.4 Conclusions

Based on the purification and structural characterisation results of all proteins, the following 5 important findings emerge:

1. A two-step purification pathway (heat-acid precipitation + butyl FF HIC) was established for HF_n. The first step aims for the removal of HCPs and, the second focuses on nucleic acid separation.

2. Purity process of HF_n failed to achieve purity above 90 % in 3 PAS functionalised HF_ns. Pure PAS functionalised HF_ns were obtained through heat-acid precipitation followed by mono Q IEC.

3. A significant decline in stability against heat and acidic pH of all PAS functionalised HF_ns was found in heat-acid precipitation in contrast with HF_n. Q FF binding strengths of HF_n was lower than PAS functionalised HF_n because PAS peptide was negatively charged in tested pH (7, 8 and 9). In HIC comparison, there was no significant difference between HF_n and PAS functionalised HF_ns.

4. PAS peptide insertion at HF_n C-terminal has caused the turnover of E-helix. Three PAS functionalised HF_ns are in fully or partially flipped conformation according to the differences in purification and structure in comparison with HF_n. The turnout of E-helix with functional peptide could be the reason for the decrease of stability in PAS functionalised HF_ns in heat-acid precipitation and the difference in tertiary structure observed in IF analysis.

5. RGDK peptide did not make any significant difference in purification process and HF_n conformation.

After the obtaining of pure and correctly structured HF_n and PAS-functionalised HF_ns, the next step is anti-tumour drug loading. Thermally induced passive drug loading was applied to HF_n and functionalised HF_ns for the first time and optimisation of drug loading conditions is reported in the following **Chapter 5**.

4.5 References

1. Young, C. L.; Britton, Z. T.; Robinson, A. S., Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnology Journal* 2012, 7 (5), 620-634.
2. Li, L.; Fang, C. J.; Ryan, J. C.; Niemi, E. C.; Lebrón, J. A.; Björkman, P. J.; Arase, H.; Torti, F. M.; Torti, S. V.; Nakamura, M. C., Binding and uptake of H-ferritin are mediated by human transferrin receptor-1. *Proceedings of the National Academy of Sciences of the United States of America* 2010, 107 (8), 3505-3510.
3. Rucker P, T. F., Torti SV, Recombinant ferritin: modulation of subunit stoichiometry in bacterial expression systems. *Protein Engineering* 1997, 10 (8), 967-973.
4. 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. *Journal of Biological Chemistry* 2010, 285 (6), 4049-4059.
5. Mosca, L.; Falvo, E.; Ceci, P.; Poser, E.; Genovese, I.; Guarguaglini, G.; Colotti, G., Use of ferritin-based metal-encapsulated nanocarriers as anticancer agents. *Applied Sciences* 2017, 7 (1), 101.
6. Falvo, E.; Tremante, E.; Fraioli, R.; Leonetti, C.; Zamparelli, C.; Boffi, A.; Morea, V.; Ceci, P.; Giacomini, P., Antibody–drug conjugates: targeting melanoma with cisplatin encapsulated in protein-cage nanoparticles based on human ferritin. *Nanoscale* 2013, 5 (24), 12278-12285.
7. Cheng, X.; Fan, K.; Wang, L.; Ying, X.; Sanders, A. J.; Guo, T.; Xing, X.; Zhou, M.; Du, H.; Hu, Y., TfR1 binding with H-ferritin nanocarrier achieves prognostic diagnosis and enhances the therapeutic efficacy in clinical gastric cancer. *Cell Death & Disease* 2020, 11 (2), 1-13.
8. Fan, K.; Cao, C.; Pan, Y.; Lu, D.; Yang, D.; Feng, J.; Song, L.; Liang, M.; Yan, X., Magnetoferritin nanoparticles for targeting and visualizing tumour tissues. *Nature Nanotechnology* 2012, 7 (7), 459-464.
9. Santambrogio, P.; Cozzi, A.; Levi, S.; Rovida, E.; Magni, F.; Albertini, A.; Arosio, P., Functional and Immunological Analysis of Recombinant Mouse H- and L-Ferritins from *Escherichia coli*. *Protein Expression and Purification* 2000, 19 (1), 212-218.
10. Uchida, M., Flenniken, M.L., Allen, M., Willits, D.A., Crowley, B.E., Brumfield, S., Willis, A.F., Jackiw, L., Jutila, M., Young, M.J., Douglas, T., Targeting of cancer cells with ferrimagnetic ferritin cage nanoparticles. *Journal of the American Chemical Society* 2006, 128, 16626-16633.
11. Su, W.; Tan, H.; Janowski, R.; Zhang, W.; Wang, P.; Zhang, J.; Zhai, H.; Li, J.; Niessing, D.; Sattler, M.; Zou, P., Ferritin-displayed GLP-1 with improved pharmacological activities and pharmacokinetics. *Molecular Pharmaceutics* 2020, 17 (5), 1663-1673.

12. Cioloboc, D.; Kurtz, D. M., Targeted cancer cell delivery of arsenate as a reductively activated prodrug. *JBIC Journal of Biological Inorganic Chemistry* 2020, 1-9.
13. Zhang, S.; Zang, J.; Zhang, X.; Chen, H.; Mikami, B.; Zhao, G., "Silent" amino acid residues at key subunit interfaces regulate the geometry of protein nanocages. *ACS Nano* 2016, 10 (11), 10382-10388.
14. Schlapschy, M.; Binder, U.; Borger, C.; Theobald, I.; Wachinger, K.; Kisling, S.; Haller, D.; Skerra, A., PASylation: a biological alternative to PEGylation for extending the plasma half-life of pharmaceutically active proteins. *Protein Engineering Design & Selection* 2013, 26 (8), 489-501.
15. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W., NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 2012, 9 (7), 671-675.
16. Cordes, A. A.; Platt, C. W.; Carpenter, J. F.; Randolph, T. W., Selective domain stabilization as a strategy to reduce fusion protein aggregation. *Journal of Pharmaceutical Sciences* 2012, 101 (4), 1400-1409.
17. Fast, J. L.; Cordes, A. A.; Carpenter, J. F.; Randolph, T. W., Physical instability of a therapeutic Fc fusion protein: domain contributions to conformational and colloidal stability. *Biochemistry* 2009, 48 (49), 11724-11736.
18. Kulakova, A.; Indrakumar, S.; Sønderby Tuelung, P.; Mahapatra, S.; Streicher, W. W.; Peters, G. H. J.; Harris, P., Albumin-nepriylisin fusion protein: understanding stability using small angle X-ray scattering and molecular dynamic simulations. *Scientific Reports* 2020, 10 (1).
19. Gerstweiler, L.; Bi, J.; Middelberg, A., Virus-like particle preparation is improved by control over capsomere DNA interactions during chromatographic purification. *Biotechnology and Bioengineering* 2021.
20. Kim, M.; Rho, Y.; Jin, K. S.; Ahn, B.; Jung, S.; Kim, H.; Ree, M., pH-dependent structures of ferritin and apoferritin in solution: disassembly and reassembly. *Biomacromolecules* 2011, 12 (5), 1629-1640.
21. Worwood, M., Serum ferritin. *Clinical Science* 1986, 70 (3), 215-220.
22. Heller, M.; Wimbish, R.; Gurgel, P. V.; Pourdeyhimi, B.; Carbonell, R. G., Reducing diffusion limitations in Ion exchange grafted membranes using high surface area nonwovens. *Journal of Membrane Science* 2016, 514, 53-64.
23. Ramos-de-la-Peña, A. M.; González-Valdez, J.; Aguilar, O., Protein A chromatography: Challenges and progress in the purification of monoclonal antibodies. *Journal of Separation Science* 2019, 42 (9), 1816-1827.
24. Ferraro, G.; Ciambellotti, S.; Messori, L.; Merlino, A., Cisplatin binding sites in human H-chain ferritin. *Inorganic Chemistry* 2017, 56 (15), 9064-9070.
25. Nasrollahi, F.; Sana, B.; Paramelle, D.; Ahadian, S.; Khademhosseini, A.; Lim, S., Incorporation of graphene quantum dots, iron, and doxorubicin in/on ferritin nanocages for bimodal imaging and drug delivery. *Advanced Therapeutics* 2020, 3 (3), 1900183.
26. Liu, X.; Powers, J. R.; Swanson, B. G.; Hill, H. H.; Clark, S., Modification of whey protein concentrate hydrophobicity by high hydrostatic pressure. *Innovative Food Science & Emerging Technologies* 2005, 6 (3), 310-317.
27. Ghisaidoobe, A. B.; Chung, S. J., Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Forster resonance energy transfer techniques. *International Journal of Molecular Sciences* 2014, 15 (12), 22518-22538.
28. Bowen, W. R.; Hall, N. J.; Pan, L. C.; Sharif, A. O.; Williams, P. M., The relevance of particle size and zeta-potential in protein processing. *Nature Biotechnology* 1998, 16, 785-787.

29. Lee, B. R.; Ko, H. K.; Ryu, J. H.; Ahn, K. Y.; Lee, Y. H.; Oh, S. J.; Na, J. H.; Kim, T. W.; Byun, Y.; Kwon, I. C.; Kim, K.; Lee, J., Engineered human ferritin nanoparticles for direct delivery of tumor antigens to lymph node and cancer immunotherapy. *Scientific Reports* 2016, 6, 35182.
30. Lee, N. K.; Lee, E. J.; Kim, S.; Nam, G.-h.; Kih, M.; Hong, Y.; Jeong, C.; Yang, Y.; Byun, Y.; Kim, I.-S., Ferritin nanocage with intrinsically disordered proteins and affibody: A platform for tumor targeting with extended pharmacokinetics. *Journal of Controlled Release* 2017, 267, 172-180.
31. A.Luzzago, G. C., Isolation of point mutations that affect the folding of the H chain of human ferritin in E.coli. *The EMBO Journal* 1989, 8 (2), 569-576.

Chapter 5

THERMALLY INDUCED DOXORUBICIN LOADING OF HUMAN HEAVY-CHAIN FERRITIN AND FUNCTIONALISED HUMAN HEAVY-CHAIN FERRITIN

5.1 Introduction

As was reviewed in **Chapter 2**, because of human heavy-chain ferritin (HF_n)'s unique structure and high biocompatibility, it has been widely recognised as a potent drug nanocarrier.¹ Through decades of efforts, diverse drug loading approaches have been developed, including: 1) disassembly/reassembly; 2) and passive diffusion.

Disassembly/reassembly approach involves a dissociation/re-association of HF_n assembly induced by pH or 8 M urea. This approach is suitable for loading drugs either smaller or larger than ferritin channels. However, the disassembly-reassembly process in pH-induced pathway has been criticized by Kim et al. Their study has proven that the process damaged ferritin structure and led to random aggregation, in which small aggregates were soluble and huge ones became precipitates.² This results in 2 problems in drug loading performance: 1) precipitation causes the unsatisfactory protein recovery yield; 2) soluble drug-ferritin aggregates can affect drug performance. Condition optimisation was often required to mitigate these problems. For example, Kilic et al. and Ruozi et al. critically investigated the pH adjustment course in drug loading and used a stepwise pH adjustment or optimisation of final pH to boost protein recovery yield to 55 %.^{3,4} The 8 M urea-based approach is less frequently used in contrast with the pH-induced one. In 2 previous studies, it showed a doxorubicin (DOX) loading ratio comparable with that of optimised pH-induced approach (around 33 DOX per HF_n nanocage), but the protein recovery yield was still undesirable, around 64.8 %.^{5,6}

Passive diffusion loads drugs through the hydrophobic or hydrophilic channels on ferritin shell, after incubating ferritin and drugs together under suitable mixing conditions. Different stressors such as high temperature, additives, and pressure have been introduced to expand the channels and facilitate drug loading. This approach poses minor effects on ferritin structure. Even though this approach causes relatively low ferritin aggregation and loss compared with the disassembly-reassembly approach, the loading efficiency is low.⁷ In the study of Yang et al, soybean ferritin was heated with rutin at 60 °C for 1 h, resulting in a loading ratio of 10.5 rutin molecules per ferritin.⁸ They also used chaotropic chemicals, urea and guanidine chloride, to expand soybean ferritin channels and load molecules.^{9,10} To boost passive loading performance, Wang et.al has successfully applied high hydrostatic pressure, explored different levels of variables, such as: pressure values, buffer pH and additives, to finally achieve a 99 % of HF_n

recovery and high doxorubicin (DOX) loading ratio (32 DOX per HFn nanocage).¹¹ However, the high pressurised device is expensive and possesses a number of potential safety risks in operation. Therefore, it is challenging to achieve concomitantly a desirable drug loading ratio and a protein recovery yield in ferritin drug loading process.

In theory, after drug enters ferritin, it retains in ferritin either by physical entrapment or chemical interaction or both. For small molecule drugs such as DOX (molecular weight < 600 Da), the chemical interaction type and strength are critical to the stability of drug loaded ferritin to prevent drug leakage from ferritin channel. Currently, the interactions have not been investigated in detail. An investigation on these interactions can help understand the drug loading mechanism, interpret the findings in drug loading and eventually lead to an improvement of drug loading performance. In the investigation of protein-ligand binding mechanism, computational tools, molecular docking and molecular dynamic (MD) simulation are significantly regarded and widely used. Molecular docking provides multiple reliable models of protein-ligand complexes, based on a searching algorithm, whilst MD simulation can help assess the validity of these complexes by stability evaluation.^{12, 13} Shahwan et al. used AutoDock Vina, a molecular docking service, to find out the most possible human ferritin (PDB ID: 3AJO)-enzyme inhibitor Donepezil complex, and ran a MD simulation of the complex to assess its stability.¹⁴ These 2 tools are potentially capable of analysing the chemical interactions between ferritin and drug in loading process.

In **Chapter 4**, pure control HFn and 3 PAS functionalised HFns were obtained. In this chapter, a thermally induced passive diffusion was introduced to load DOX to 3 PAS functionalised HFns and HFn. It is expected to obtain desirable loading results. Condition optimisation in drug loading for HFn and HFn-GFLG-PAS-RGDK were conducted. Size-exclusion chromatography (SEC) and transmission electron microscopy (TEM) were used to detect the structures of proteins after drug loading. DOX loaded proteins (protein/DOX) stability test was performed to check whether drug loaded HFn-based proteins would be stable during storage. For the first time, computational analysis, molecular docking and MD simulation, were adopted to analyse chemical interactions contributing to drug loading and aggregation in thermally induced DOX loading process.

5.2 Materials and methods

5.2.1 Materials

All HF_n-based proteins (HF_n, HF_n-PAS, HF_n-GFLG-PAS-RGDK, HF_n-PLGLAG-PAS-RGDK) were purified in our laboratory using the optimal conditions, as is described in **Chapter 4**. Doxorubicin hydrochloride (DOX) was purchased from Dalian Meilun Biotechnology (China). Chromatography columns were bought from GE Healthcare (USA). All other reagents of analytical grade (Chem-Supply, Australia) were used directly. Milli Q water was obtained from a Millipore purification system (Merck, USA).

5.2.2 Thermally induced passive loading of DOX into HF_n and HF_n-GFLG-PAS-RGDK

DOX was loaded to HF_n-based nanocages through thermally induced passive diffusion. Temperature, buffer pH and incubation time are the 3 main factors affecting drug loading. An orthogonal test was designed to optimise thermally induced drug loading condition for HF_n and HF_n-GFLG-PAS-RGDK. Variables and levels tested are listed in **Table 5.1**.

Table 5.1 Levels of variables used in the orthogonal tests for optimisation of thermally induced DOX loading to HF_n and HF_n-GFLG-PAS-RGDK.

Variables	Level
Temperature	45, 50 and 60 °C
Phosphate buffer pH	7.0 and 7.5
Incubation time length	2, 4 and 6 h

Initial protein concentration (1 mg mL⁻¹) and DOX concentration (0.2 mg mL⁻¹) were used in all conditions to ensure DOX is excessive. Sample buffer was 20 mM phosphate buffer with 5 mM guanidinium chloride, pH 7 or 7.5. After thermal incubation of DOX and HF_n-based nanocages, samples were centrifuged at 1,000 rpm 10 min at 4 °C to remove precipitates. Protein recovery yields were calculated as in following equation.

$$\begin{aligned} & \text{Protein recovery yield (\%)} \\ &= \frac{\text{protein amount in supernatant after centrifugation}}{\text{protein amount before thermal treatment}} \times 100\% \end{aligned} \tag{5.1}$$

Unloaded DOX was removed using Hitrap Sephadex G-25 column (GE healthcare, USA) and DOX loaded HF_n-based protein (protein/DOX) were collected. All protein/DOX peaks then

underwent SEC by Superose 6 increase 10/300 GL column (GE Healthcare, USA) to detect if any soluble HF_n-DOX aggregates existed.

SEC can separate DOX loaded in HF_n-based nanocages (DOX loaded in nanocage) from soluble HF_n-DOX aggregates. Peak areas (absorbance at 480 nm) can be used to determine the proportion of DOX loaded in nanocage.

Proportion of DOX loaded in nanocage (%)

$$= \frac{OD480 \text{ area of nanocage peak}}{OD480 \text{ area of HF}_n - \text{DOX aggregate peak} + OD 480 \text{ area of nanocage peak}} \times 100\% \quad (5.2)$$

Drug loading ratio, protein recovery yield and the proportion of DOX loaded in nanocages under various conditions were compared to find the optimal condition. Drug loading ratio, i.e., the number of DOX per HF_n or functionalised HF_n nanocage (N), was determined by measuring the OD480 and OD280 of protein/DOX peaks collected from the Sephadex G25 column, as is described in **Equation 5.3**.

$$N = \frac{\text{number of DOX}}{\text{number of nanocage}} = \frac{C_{DOX} \cdot Mw_{nanocage}}{C_{nanocage} \cdot Mw_{DOX}} \quad (5.3)$$

Where C is the concentration (mg mL⁻¹), and Mw is molecular weight. The DOX concentration (C_{DOX}) and the concentration of protein nanocage (C_{nanocage}) were determined by measuring absorbance at 480 and 280 nm, respectively. DOX has absorbance at both 280 and 480 nm, and protein only has absorbance at 280 nm. Therefore, 2 assumptions were made: 1) OD480_{nanocage/DOX} = OD480_{DOX}; 2) OD280_{nanocage/DOX} = OD280_{DOX} + OD280_{nanocage}. Standard OD vs. C linear curves of DOX and HF_n-based nanocages were determined by serial concentrations of DOX (1 - 40 µg mL⁻¹) and protein nanocages (0.1 - 1.2 mg mL⁻¹).

5.2.3 Thermally induced passive loading of DOX into HF_n-PLGLAG-PAS-RGDK and HF_n-PAS

For HF_n-PAS and HF_n-PLGLAG-PAS-RGDK, DOX loading was conducted at the optimal loading condition of HF_n-GFLG-PAS-RGDK. Briefly, 0.2 mg mL⁻¹ DOX was added to protein solutions (1 mg mL⁻¹) and mixed thoroughly. Mixtures were then heated at 50 °C for 6 h,

followed by Hitrap Sephadex G25 (GE Healthcare, USA) chromatography to remove excessive free drugs. Drug loading ratio, proportion of DOX loaded in nanocage, and protein recovery yield were calculated using **Equation 5.1** and **5.2**. DOX loaded HF_n-PAS and HF_n-PLGLAG-PAS-RGDK were also analysed using SEC.

5.2.4 TEM analysis of protein/DOX

TEM was conducted to analyse the conformation of HF_n/DOX, HF_n-PAS/DOX, HF_n-GFLG-PAS-RGDK/DOX and HF_n-PLGLAG-PAS-RGDK/DOX. A DOX loaded HF_n aggregate collected from Superose 6 chromatography was also analysed. TEM procedure was the same as described in **Chapter 3, Section 3.2.6**.

5.2.5 Stability of DOX loaded HF_n and PAS functionalised HF_ns

After drug loading, the buffer of DOX loaded HF_n and PAS functionalised HF_ns samples were exchanged into either phosphate-buffered saline (PBS) pH 7.4 or 20 mM sodium acetate- acetic acid (NaAc-HAc) pH 5.5 buffer. Buffer exchanged samples were placed at 37 °C or 4 °C. Aliquots were taken from samples at certain time points (0, 2, 4, 8, 24, 72, 120, 168, 336 h) and desalted using Hitrap Sephadex G25 column (GE Healthcare, USA) to remove leaked DOX, followed by N value calculation.

5.2.6 Computational study of interactions of HF_n and DOX in thermally induced drug loading

PyRx molecular docking and Gromacs MD simulation analysis were used to identify the potential HF_n and DOX interactions to explain the formation of HF_n/DOX and soluble HF_n-DOX aggregates. Molecular docking was performed to analyse the possible poses of HF_n subunit and DOX interactions. MD simulation of the docking complexes aimed to find out the most stable HF_n subunit-DOX complex structures.

Computational analysis was based on 2 assumptions: 1) HF_n subunit was used as a representative of HF_n assembly. This is because the assembly was theoretically 24 repetitions of the monomer. DOX is very small, even smaller than HF_n channels, which makes it unlikely to simultaneously interact with more than one subunit of the same HF_n assembly. 2) The computational analysis focus was on the interactions between DOX and the residues located

on HF_n assembly outer surface and inner surface. All interactions with interface residues of HF_n assembly were ignored.

In molecular docking analysis, PyRx software was used. DOX 3D structure was from Pub Chem and HF_n subunit structure file (PDB file) from RCSB PDB (ID: 2FHA). DOX was energy minimized before conducting docking. Top 9 docking HF_n-DOX complexes were obtained and saved as PDB files. PDB files from docking results underwent Gromacs MD simulation and analysed using Gromacs 2018.

In MD simulation, CHARMM36 force field was used. The HF_n-DOX complex structure was solvated in a dodecahedral box of size 460.73 nm³ with water molecules and the box was charge neutralised by replacing 8 water molecules with 8 Na⁺ ions. Energy minimization was conducted using the steepest descent integrator for 50,000 steps, until a tolerance of 10 kJ mol⁻¹. After this, temperature (NVT) and pressure equilibration (NPT) of the full system were performed at 323 K (approximate 50 °C). Finally, 10 ns 323 K simulation were conducted with 5,000,000 steps and 2 fs each step. Lincs constraint algorithm, Verlet cut-off scheme, Particle Mesh Ewald coulomb type were used in this MD simulation. Root-mean-square deviation (RMSD) and short-range non-bonded interaction energy of each complex in MD simulation were analysed for the stability assessment. 3D structures of 9 complexes after MD simulation were saved and the interactions of HF_n subunit and DOX within were visualised by Discovery studio visualizer. Interactions analysed include hydrogen bond, salt bridge and Pi (π) effects. Possible hydrophobic interaction was evaluated by analysing the residue hydrophobicity in DOX binding area.

5.3 Results and Discussion

5.3.1 Optimisation of HF_n thermally induced passive loading to increase drug loading

Thermally induced strategy mainly takes advantage of the thermal energy mediated structural perturbation of selective hydrophilic channel areas. Theil and co-workers used Circular Dichroism to analyse the α -helix content change of HF_n following heat treatment at different temperatures, and found that a small amount of secondary structure began to transition into random coil when temperature is greater than 45 °C, and it is very likely to take place in channel areas and expand channels.¹⁵ Heating also accelerates Brownian motion of proteins and drug

molecules so that greater efficiency could be achieved than in non-heated passive diffusion. In this work, pH 7.0 and 7.5 were chosen to ensure that DOX carries positive charge (DOX pKa 8.3) and HFn exterior surface has the opposite charge (HFn pI 4.8). Temperature conditions are selected based on thermal stability of HFn.

Figure 5.1 summarises the changes of drug loading ratios, proportions of DOX loaded in nanocage and HFn recovery yields with varying thermal induction time, temperature and buffer pH. Drug loading ratios (Ns) were calculated using the standard curves in **Figure B3** (**Appendix B**). **Table 5.2** lists all drug loading ratios (Ns), proportions of DOX loaded into nanocage and protein recovery yields for HFn. As is shown in **Figures 5.1A, 5.1C, 5.1E** and **Table 5.2**, with the increase of thermal induction duration from 2 to 6 h, N increased at all tested temperatures. At 45, 50 and 60 °C, the highest N was 30.3, 41.6 and 56.7, respectively. N at pH 7.5 was slightly greater than that at pH 7 in most of the time regardless of temperature.

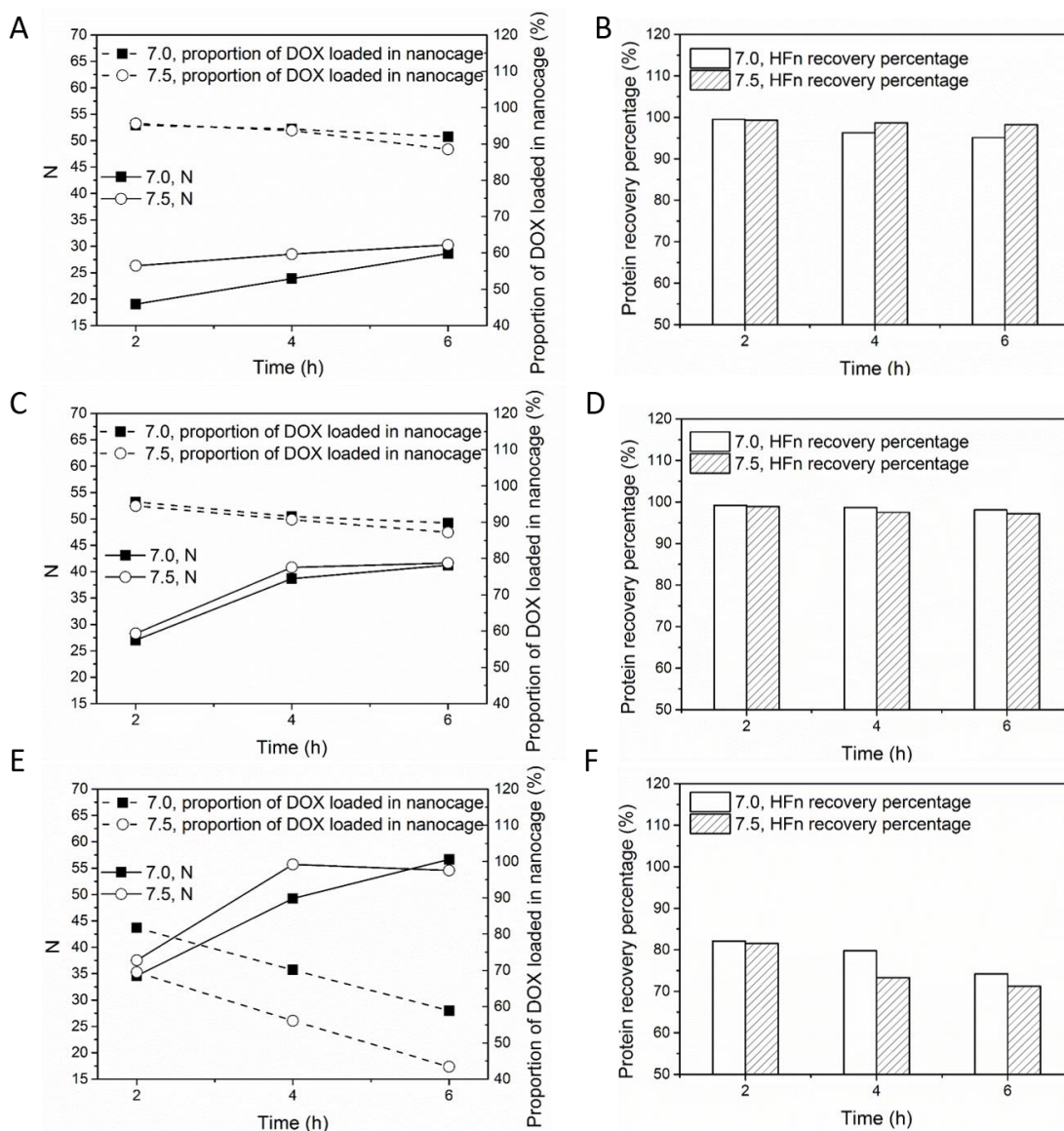


Figure 5.1 Thermally induced DOX loading to HFn under different experimental conditions **A**, drug loading ratios (Ns) and proportions of DOX loaded in nanocage at 45 °C. **B**, HFn recovery yields at 45 °C. **C**, loading ratios (Ns) and proportions of DOX loaded in nanocage at 50 °C. **D**, HFn recovery yields at 50 °C. **E**, loading ratios (Ns) and proportions of DOX loaded in nanocage at 60 °C. **F**, HFn recovery yields at 60 °C.

Table 5.2 Drug loading ratios (Ns), proportions of DOX loaded in nanocage and protein recovery yields in HFn thermally induced drug loading optimisation.

Temperature (°C)	Thermal induction duration (h)	Buffer pH	N	Proportion of DOX loaded in nanocage (%)	Protein recovery yield (%)
45	2	7.0	19.06	95.2	99.5
		7.5	26.30	95.6	99.3
	4	7.0	23.88	94.1	96.3
		7.5	28.51	93.7	98.7

	6	7.0	28.62	92.0	95.1
		7.5	30.26	88.5	98.2
50	2	7.0	27.04	95.6	99.2
		7.5	28.28	94.5	98.9
	4	7.0	38.69	91.6	98.7
		7.5	40.8	90.7	97.5
	6	7.0	41.24	89.8	98.1
		7.5	41.64	87.2	97.2
60	2	7.0	34.60	81.7	82.1
		7.5	37.53	69.5	81.5
	4	7.0	49.25	70.1	79.8
		7.5	55.69	56.1	73.3
	6	7.0	56.67	58.9	74.2
		7.5	54.56	43.4	71.2

In terms of proportions of DOX loading, at 45 and 50 °C, the values were at least 85 % whilst at 60 °C values were below 85 %. At all 3 temperatures, proportion of DOX loaded into nanocage decreased with the duration of thermal induction. The proportions are significantly pH-dependent, and the extent of pH-dependency is subject to temperature, hence they decreased by 0.4 - 3.5 % at 45 and 50 °C, and by 10 - 15 % at 60 °C as pH increased from 7.0 to 7.5.

For the HF_n recovery yields shown in **Figures 5.1B, 5.1D, 5.1F** and **Table 5.2**, at 45 and 50 °C, they were above 90 %, and at 60 °C, they were mostly below 85 %. These results suggest that in the thermally induced drug loading process, DOX loaded in individual HF_n nanocages, soluble HF_n-DOX aggregates and HF_n-DOX precipitates were simultaneously produced as in previous research using disassembly/reassembly drug loading approaches. At 45 °C, proportion of drug loaded in nanocage and HF_n recovery yield decreased slowly and N increased slowly over time, whilst at 60 °C, proportion of drug loaded in nanocage, HF_n recovery yield and N behaved in the opposite manner. These results confirm that 45 °C may not be effective to accelerate drug loading. In addition, at 60 °C, the local structures of HF_n nanocages undergo excessive changes, resulting in massive formation of aggregates of HF_n with DOX.

Considering N, proportion of DOX loaded in nanocage and HFfn recovery yield together, 50 °C, pH 7.5, 6 h is the best drug loading condition (N of 41.6, proportion of DOX loaded in nanocage of 87.2 % and HFfn recovery yield of 97.2 %).

Table 5.3 indicates the comparison of HFfn drug loading performance between this work and published studies. In this study, N of HFfn (41.6) is greater than previously published results, which adopted 8 M urea or optimised stepwise pH-induced disassembly-reassembly approaches. Recovery yield of HFfn in this study is 97.2 %, similar to high hydrostatic pressure passive diffusion approach (99 %) and significantly greater than the pH-induced (25 %, 55 %) or 8 M urea-based approach (64.8 %). Disassembly/reassembly approach has been questioned to be not fully reversible because 2 holes were detected by synchrotron small-angle X-ray scattering, and the author argued that this structural damage may result in protein loss and aggregation in the drug loading process.² To the contrary, at 50 °C, HFfn nanocage remains intact throughout the thermally induced drug loading process, which involves less structural changes.

Table 5.3 Comparison on DOX loading to HFfn in this work and previous studies.

Protein	Loading approach	Average number of DOX per protein cavity (N)	Protein recovery yield (%)	Reference
HFfn	Thermal induction	41.6	97.2	
Horse spleen ferritin	Stepwise pH induction	28	55 ± 7	Ref ³
HFfn	pH induction	28.3	/	Ref ¹⁶
Equine spleen ferritin	pH induction	22 ± 1	25	Ref ¹⁷
Ferritin	Urea-based	32.5	64.8	Ref ⁶
HFfn	Urea-based	33	/	Ref ⁵
HFfn	High hydrostatic pressure	32 ± 2	99	Ref ¹¹

5.3.2 Optimisation of HFfn-GFLG-PAS-RGDK thermally induced passive loading

For HFfn-GFLG-PAS-RGDK, DOX loading optimization results are shown in **Figure 5.2** and **Table 5.4**. **Figure B7, 8** and **9 (Appendix B)** list size-exclusion chromatograms of HFfn-GFLG-PAS-RGDK/DOX obtained under 45, 50 and 60 °C. The relations between drug loading performance indicators (N, proportion of DOX loaded in nanocage and HFfn-GFLG-PAS-RGDK recovery yield) and experimental variables (induction time, pH and temperature) are similar to those in HFfn.

In **Figures 5.2A, 5.2C, 5.2E** and **Table 5.4**, N positively responded to thermal induction duration and temperature. pH 7.5 showed greater Ns than pH 7.0 in most of time. Proportion of DOX loaded in nanocage was negatively related to temperature and incubation time. At 45 and 50 °C, proportions of DOX loaded in nanocage were greater than 75 %. At 60 °C, they were lower than 70 %. As is shown in **Figures 5.2B, 5.2D, 5.2E** and **Table 5.4**, HF_n-GFLG-PAS-RGDK recovery yields were greater than 75 % except at 60 °C 4 h and 6 h. The best DOX loading condition is obtained at 50 °C, pH 7.5 and 6 h, with an N of 45.2, proportion of DOX loaded in nanocage of 78.5 % and HF_n-GFLG-PAS-RGDK recovery yield of 76.0 %.

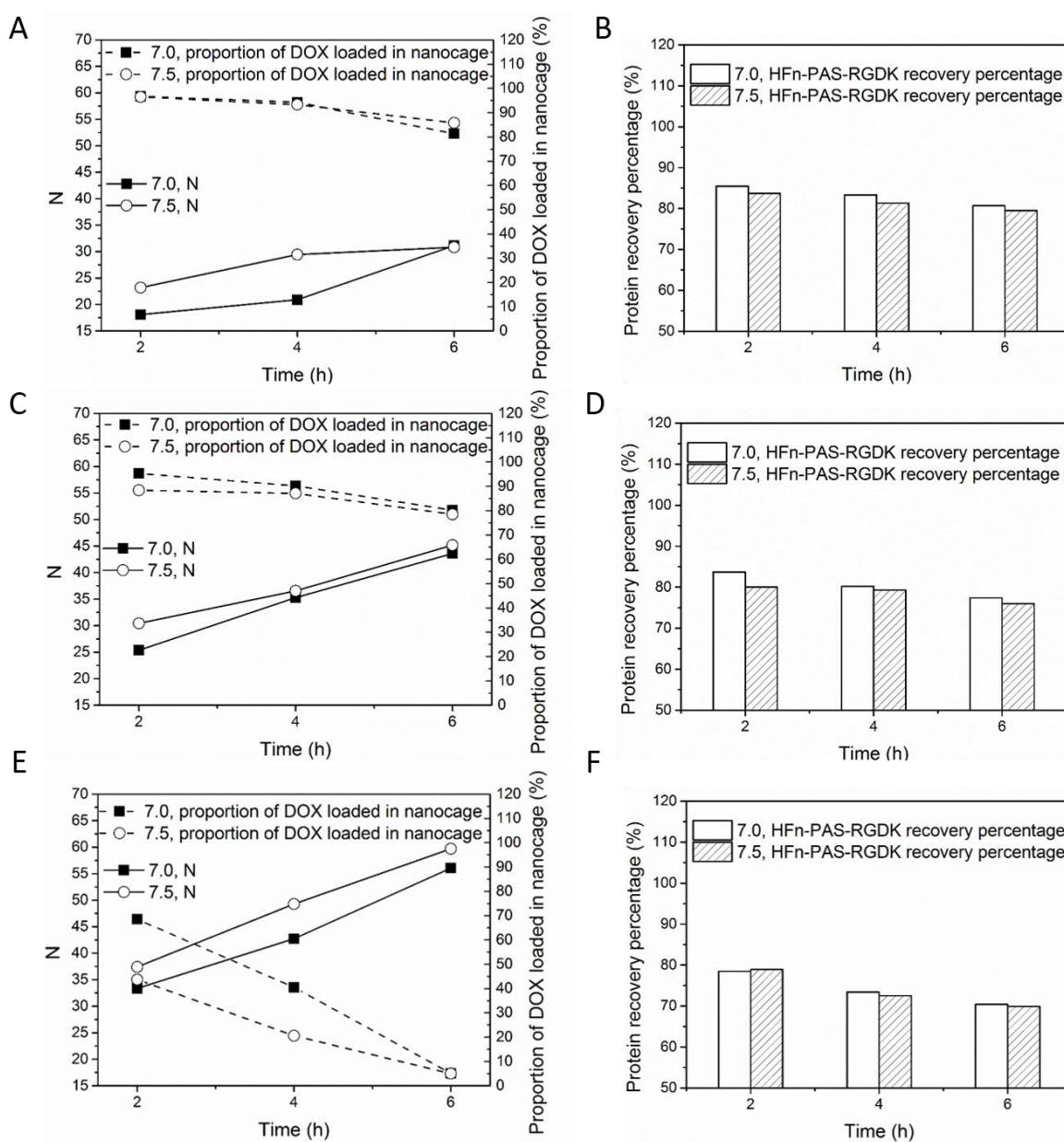


Figure 5.2 Thermally induced DOX loading to HF_n-GFLG-PAS-RGDK under different experimental conditions.

A, loading ratios (Ns) and proportions of DOX loaded in nanocage at 45 °C. **B**, HF_n-GFLG-PAS-RGDK recovery yields at 45 °C. **C**, loading ratios (Ns) and proportions of DOX loaded in nanocage at 50 °C. **D**, HF_n-GFLG-PAS-RGDK recovery yields at 50 °C. **E**, loading ratios (Ns) and proportions of DOX loaded in nanocage at 60 °C. **F**, HF_n-GFLG-PAS-RGDK recovery yields at 60 °C.

Table 5.4 Ns, proportions of DOX loaded in nanocage and protein recovery yields in HF_n-GFLG-PAS-RGDK thermally induced drug loading optimisation

Temperature (°C)	Thermal induction duration (h)	Buffer pH	N	Proportion of DOX loaded in nanocage (%)	Protein recovery yield (%)
45	2	7.0	18.10	96.7	85.5
		7.5	23.20	96.5	83.7
	4	7.0	20.92	94.3	83.3
		7.5	29.43	93.3	81.3
	6	7.0	31.13	81.3	80.7
		7.5	30.84	85.8	79.5
50	2	7.0	25.35	95.3	83.7
		7.5	30.44	88.4	80.0
	4	7.0	35.29	90.2	80.2
		7.5	36.54	87.1	79.3
	6	7.0	43.60	80.2	77.4
		7.5	45.16	78.5	76.0
60	2	7.0	33.37	68.5	78.5
		7.5	37.41	43.7	78.9
	4	7.0	42.73	40.5	73.4
		7.5	49.28	20.6	72.5
	6	7.0	56.03	5.1	70.4
		7.5	59.68	5.0	69.9

Compared with HF_n, under most experimental conditions, especially at 50 and 60 °C, HF_n-GFLG-PAS-RGDK shows relatively low protein recovery yields and low proportions of DOX loaded in nanocage, implying a lower stability caused by the functional peptide insertion.

5.3.3 Conformation characterisation of optimal protein/DOX

Figure 5.3 shows the size-exclusion chromatograms of HFn/DOX and functionalised HFns/DOX obtained from 50 °C, pH 7.5 6 h condition. All of these have low proportions of DOX loaded aggregates. As is shown in **Table 5.5**, HFn-PAS, HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK have reduced proportions of DOX loaded in nanocage and protein recovery yields, in contrast to HFn at the same condition. Therefore, all 3 PAS functionalised HFns have been more prone to aggregation in the heating process, suggesting the functionalised HFns are less stable. This is in agreement with the conformational changes of HFn after functional peptide insertion and stability decrease, as was discovered in **Chapter 4**. The E-helix turnover and the impact on hydrophobic interactions of 4 helices E are probably the reasons.

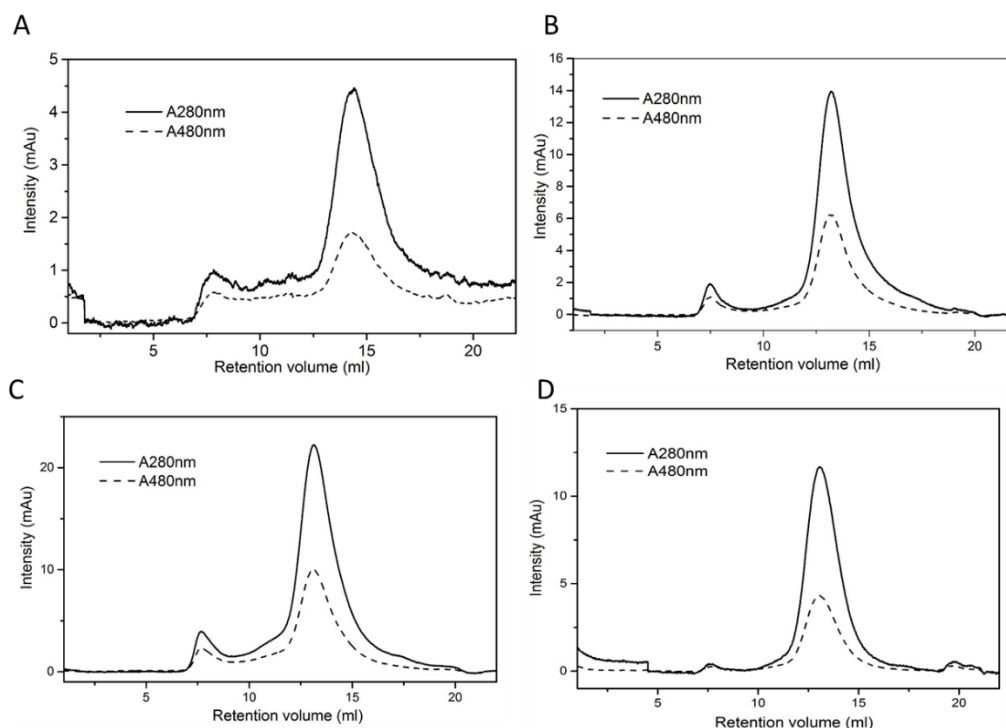


Figure 5.3 Size-exclusion chromatograms of HFn and functionalised HFns/DOX. **A**, HFn/DOX. **B**, HFn-PAS/DOX. **C**, HFn-GFLG-PAS-RGDK/DOX. **D**, HFn-PLGLAG-PAS-RGDK/DOX.

Table 5.5 DOX loading efficiency of HFn and functionalised HFns

Protein	N	Proportion of DOX loaded in nanocage (%)	Protein recovery yield (%)
HFn	41.6	87.2	97.2
HFn-PAS	38.4	73.4	75.1
HFn-GFLG-PAS-RGDK	45.2	78.5	76.0

After drug loading, TEM images in **Figure 5.4** demonstrate that the heating process did not affect ferritin nanocage. Most of the protein/DOX were hollow spheres. Nanocage sizes were still around 12 nm, which are the same as before thermally induced passive drug loading process.

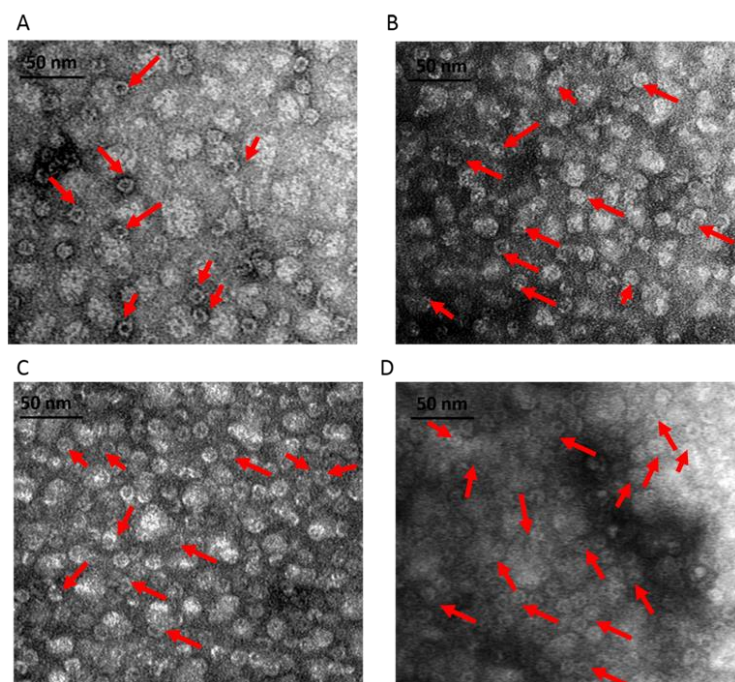


Figure 5.4 TEM images of protein/DOX. **A**, HF_n/DOX. **B**, HF_n-PAS/DOX. **C**, HF_n-GFLG-PAS-RGDK/DOX. **D**, HF_n-PLGLAG-PAS-RGDK/DOX. Red arrows indicate some of the protein/DOX structures.

5.3.4 Stability study of DOX loaded proteins

Protein/DOX stability test was designed to reflect the stability of protein/DOX in the storage condition (4 °C, pH 7.4) and blood circulation environment (37 °C). In storage condition, drug leakage profiles for all protein/DOX were consistent, where around 20 % of loaded drug leaked over 2 weeks (**Figure 5.5**). At pH 5.5, 4 °C, around 25 % of DOX leaked out in 2 weeks. At 37 °C, protein/DOX were less stable, with around 30 % of drug loss detected in HF_n/DOX and 35 % of drug leaking in other 3 groups for 2 weeks. In all 4 DOX loaded proteins, drug leaked fast during the initial 12 h, then slowed down. Perhaps some of the loaded drugs are just loosely attached or physically trapped inside protein nanocages. Hence these drugs are more prone to dissociation from protein, while drugs strongly interact with HF_n remain within ferritin. The drug loss difference detected at 37 °C between functionalised HF_ns and HF_n can be explained by the conformation difference.

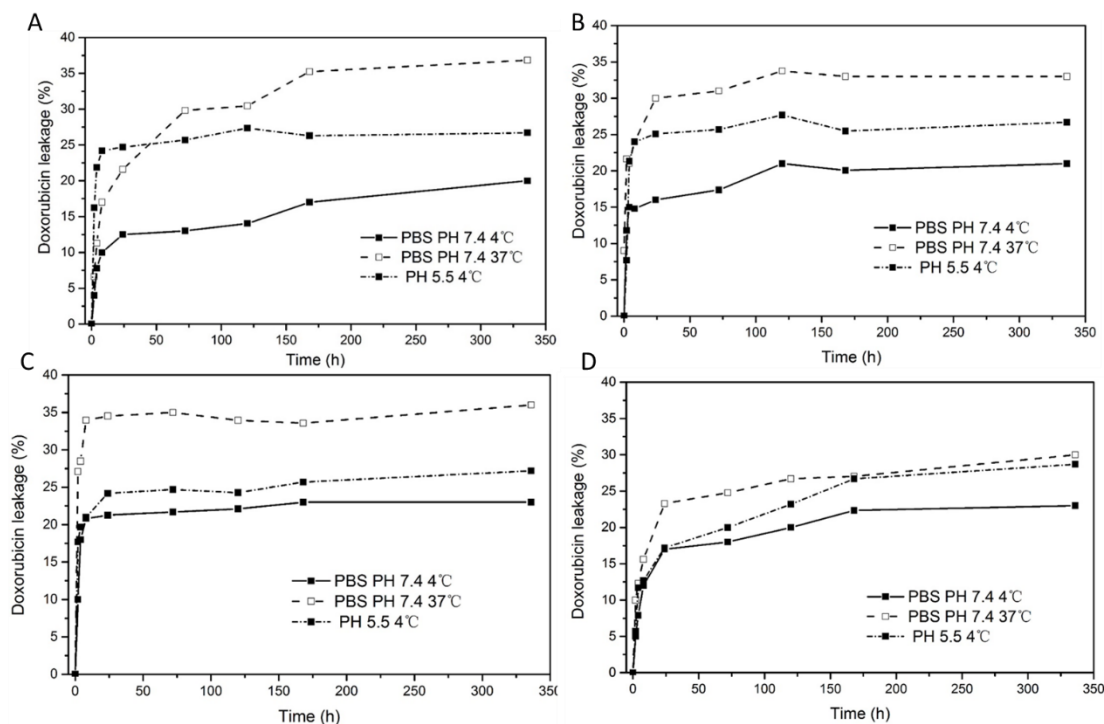


Figure 5.5 DOX leakage over time at different conditions. **A**, HFnanoparticle-GFLG-PAS-RGDK/DOX. **B**, HFnanoparticle-PAS/DOX. **C**, HFnanoparticle-PLGLAG-PAS-RGDK/DOX. **D**, HFnanoparticle/DOX.

5.3.5 Interactions between HFnanoparticle and DOX in thermally induced drug loading by computational analysis

From molecular docking results, 9 different HFnanoparticle subunit-DOX complexes were obtained. They underwent 10 ns 50 °C MD simulation for stability assessment. 3D structures of 9 complexes after simulation are shown in **Figure 5.6**. Among them, only in Complex 1, the location DOX binds to is the inner surface in HFnanoparticle assembly. In the other 8 complexes, DOX binds to areas corresponding to the outer surface in HFnanoparticle assembly. This implies that Complex 1 is very likely to be the structure of DOX loaded in individual HFnanoparticle nanocage, while the interaction ways in the other complexes could form drug loading, soluble HFnanoparticle-DOX aggregates, and HFnanoparticle-DOX precipitates in thermally induced drug loading process. This is because when DOX is bound to the residues on inner surface of HFnanoparticle assembly, it can only interact with one HFnanoparticle assembly. However, when DOX is bound on the surface of HFnanoparticle assembly, it is possible to work as a cross-linker between HFnanoparticle assemblies to induce HFnanoparticle aggregation.

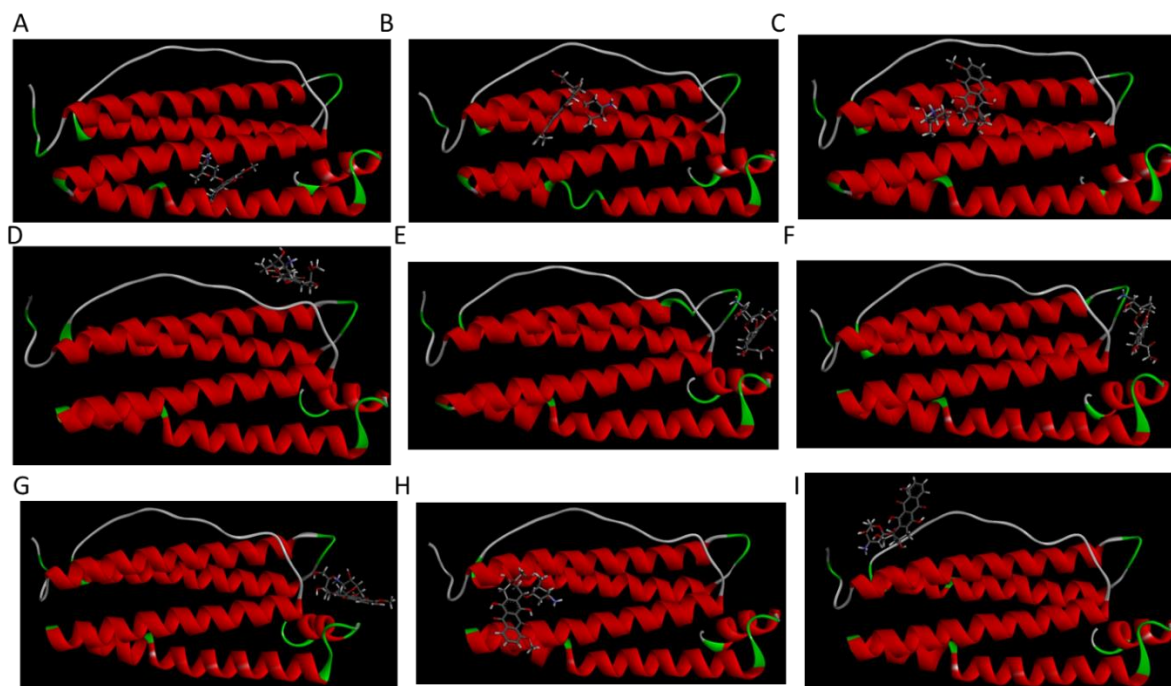


Figure 5.6 3D structures of 9 complexes after 10 ns MD simulation. **A-I** are Complex 1 to Complex 9.

To evaluate the stabilities of these structures, the RMSD and the short-range non-bonded interaction energy of HF_n subunit and DOX molecule in 9 complexes during simulation were monitored and are presented in **Figure 5.7** and **5.8**. The smaller the RMSD and the lower the energy is, the more stable the complex structure is. The stability orders of structures demonstrated in RMSD and energy are consistent.

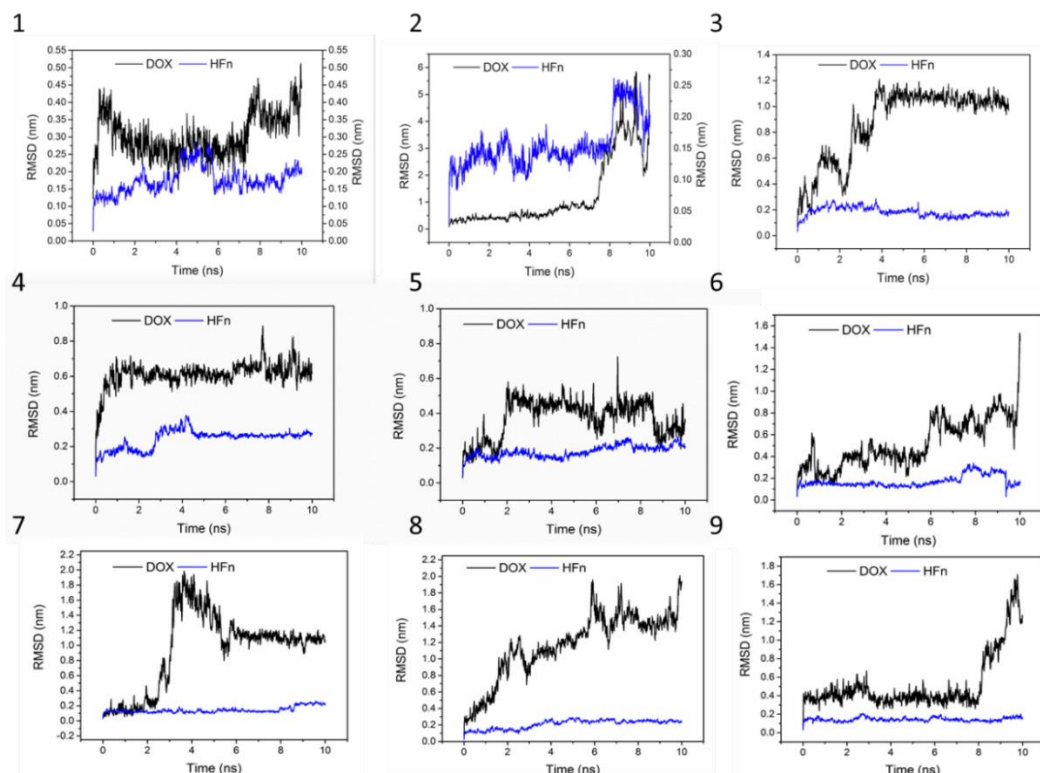


Figure 5.7 RMSD of HFn subunit and DOX in complexes during MD simulation

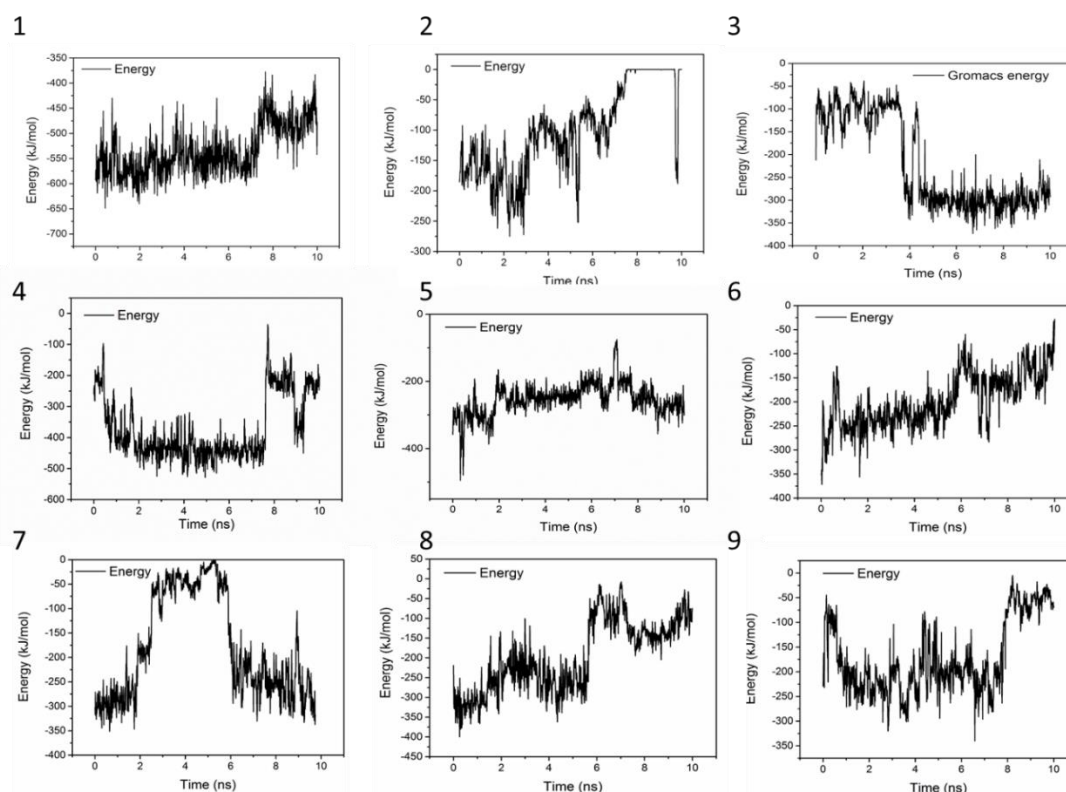


Figure 5.8 Short-range non-bonded interaction energy of HFn subunit and DOX in complexes during MD simulation.

Complex 1 was the most stable structure in 50 °C MD simulation, with RMSD lower than 1 nm and energy below -350 kJ mol^{-1} . This result is in accordance with the experiment result

conducted under 50 °C, that more DOX was being loaded to individual HF_n nanocage than forming soluble aggregates and precipitates. Complex 4 and 5 were the second most stable structures, of which the RMSD were below 1 nm and the energies were below -200 kJ mol⁻¹ most of the time. Complex 3 was the third most stable structure. Its RMSD was lower than 1.2 nm. RMSD of Complex 7, 6, 8, 9 and 2 were greater than 1.5 nm and energies of them were above -50 kJ mol⁻¹, indicating relatively unstable structures and possibly weak interactions. Based on the stability results, Complex 1, 4, 5 and 3 are the focus in interaction analysis.

Table 5.6 lists the hydrogen bond, salt bridge and Pi effect interactions between HF_n subunit and DOX in Complex 1, 4, 5 and 3 at 10 ns of the simulation. **Figure B9 (Appendix B)** lists the 2D diagrams of 9 complexes of HF_n subunit with DOX after 10 ns MD simulation. Hydrogen bonds and salt bridges are strong non-covalent bonds, in contrast with van der Waals interaction such as Pi effects. The more of them there are, the more stable the complex structure is. Complex 1 has the most hydrogen bonds and salt bridges.

Table 5.6 List of residues in HF_n subunit that interacts with DOX in complexes and the interaction type.

Complex	Residues forming hydrogen bond with DOX	Residues forming salt bridge/attractive charge with DOX	Residues have Pi effects with DOX
1	GLN58, GLU62, HIS65, GLN141	GLU27, GLU62 (2), GLU107 (2)	HIS57 (Pi-Pi stacked), TYR54(Pi-alkyl)
4	ARG43, ASP91	ASP91	TYR39 (Pi-Pi stacked), TYR39 and PRO88 (Pi-alkyl)
5	TYR40, ASP45, GLU94, GLU167	ASP45	/
3	/	/	TYR29 (Pi-Pi T shaped), LEU26 (Pi-alkyl)

Numbers in the brackets after the residue are the number of the interactions involving the residue.

Regarding the possible hydrophobic interactions between HF_n subunit and DOX in Complex 1, 4, 5 and 3, the 5 residue average hydrophobicity to reflect the hydrophobicity level of residues in DOX binding area was used. This is because in a protein, the hydrophobicity of residues can be affected by the nearby residues. Local area hydrophobicity reflects the possibly of hydrophobic interaction better than considering individual residue hydrophobicity. The calculation of 5 residue average hydrophobicity has considered the impact of nearby residues

and its value demonstrates how hydrophobic the local area of the residue is. Hydrophobicity values in **Table 5.7** were calculated using Discovery Studio Visualiser. The greater the value is, the more likely it would interact with the hydrophobic DOX molecule. In all four (4) complexes, there are at least three (3) residues at the binding pocket available for hydrophobic interactions with DOX.

Table 5.7 Five (5) residue average hydrophobicity of residues in DOX binding area in complex 1, 4, 5 and 3.

Complex	5-residue average hydrophobicity values of hydrophobic residues in DOX binding pocket
1	TYR34 (0.92), TYR54 (0.64), LYS143 (0.62), ALA144 (0.54), GLU147 (0.1).
4	TYR32 (0.52), SER36 (0.56), TYR39 (0.26).
5	VAL46 (0.56), ALA47 (0.48), LEU48 (0.48).
3	LEU26 (1.02), GLN83 (0.82), GLN112 (0.04), GLU116 (0.94).

Numbers in brackets following the residue are the five (5) residue average hydrophobicity value of the residue.

According to the computational analysis, in DOX loaded individual HF_n nanocage, DOX was mostly bound to HF_n subunit as in Complex 1. Relatively weak binding ways found in Complex 4, 5 and 3 and physically trapped DOX also existed. Therefore, the loading ratio could reach above 24. However, DOX remained in HF_n by these weaker ways are more prone to be dissociated during storage. Physically trapped DOX probably accounts for the burst release of DOX in the initial 12 h, and the weakly bounded DOX on HF_n surface in Complex 4, 5 and 3 would gradually be released, as is shown in **Figure 5.5**.

Because no aggregates nor precipitates were found in 50 °C 6 h heated HF_n, it is reasonable to infer that the interaction of DOX and HF_n assembly has led to HF_n and DOX aggregation. Small aggregates are still soluble while huge ones turn into HF_n-DOX precipitates. TEM image in **Figure 5.9A** demonstrates that the HF_n in soluble HF_n-DOX aggregates were still intact spheres but clumped into a large particle. Interaction ways in Complex 4, 5 and 3 and others, except Complex 1, are theoretically possible to cause aggregation in a way that DOX works as a cross linker (**Figure 5.9B**). In each HF_n assembly, there are 24 subunits for DOX to bind to, and the large aggregate can stem from multiple DOX molecules.

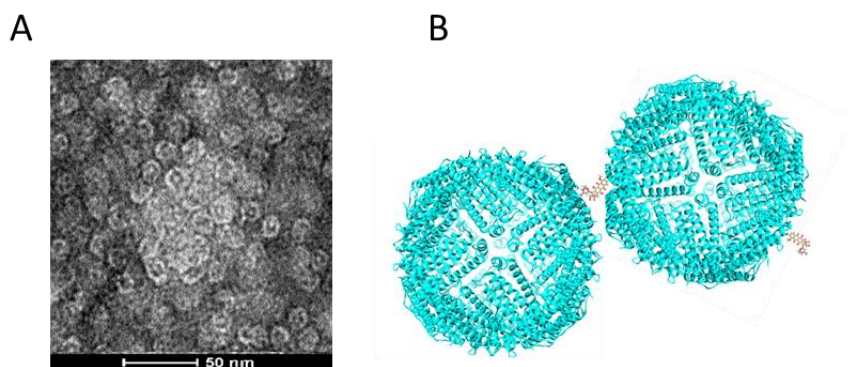


Figure 5.9 TEM image of soluble HFn-DOX aggregates (**A**) and schematic of conformation of soluble HFn-DOX aggregates (**B**). Cyan part is HFn assembly and brown part is DOX molecule.

5.4 Conclusion

Main findings from this chapter include:

1. Thermally induced passive drug loading approach has achieved a desirable DOX loading ratio and a greater protein recovery for HFn than disassembly-reassembly approach used in previous approaches.
2. Thermally induced passive drug loading also demonstrates desirable results in functionalised HFns, HFn-PAS, HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK. This approach holds the potential to be a widely applicable alternative of drug loading for HFn-based proteins.
3. The protein recovery yields of functionalised HFns were lower in comparison with that of HFn due to the E-helix turnover and the stability decrease, as was described in **Chapter 4**.
4. According to molecular docking and MD simulation analysis, hydrogen bonds, and slat bridges are the main interactions between HFn and DOX molecules contributing to DOX loading. Interactions as in Complex 1 are the most dominant. Interactions in other complexes together with physical entrapment also contributes to DOX loading.
5. Multiple DOX molecules work as a cross-linker to form aggregates of DOX and HFn.

Based on the results of previous chapters and this chapter, HFn and 3 PAS functionalised HFns were successfully prepared and loaded with anti-tumour drug DOX. The evaluation of anti-tumour performances is necessitated to examine if the PAS and RGDK peptide can improve

HF_n as in theory. **Chapter 6** focuses on employing *in vitro* and *in vivo* assays to compare DOX loaded HF_n and PAS functionalised HF_ns.

5.5 References

1. Truffi, M.; Fiandra, L.; Sorrentino, L.; Monieri, M.; Corsi, F.; Mazzucchelli, S., Ferritin nanocages: A biological platform for drug delivery, imaging and theranostics in cancer. *Pharmacological Research* 2016, 107, 57-65.
2. Kim, M.; Rho, Y.; Jin, K. S.; Ahn, B.; Jung, S.; Kim, H.; Ree, M., pH-dependent structures of ferritin and apoferritin in solution: disassembly and reassembly. *Biomacromolecules* 2011, 12 (5), 1629-1640.
3. Kilic, M. A.; Erdal, O.; Calis, S., A novel protein-based anticancer drug encapsulating nanosphere: Apoferritin-doxorubicin complex. *Journal of Biomedical Nanotechnology* 2012, 8 (3), 508-514.
4. Ruozi, B.; Veratti, P.; Vandelli, M. A.; Tombesi, A.; Tonelli, M.; Forni, F.; Pederzoli, F.; Belletti, D.; Tosi, G., Apoferritin nanocage as streptomycin drug reservoir: Technological optimization of a new drug delivery system. *International Journal of Pharmaceutics* 2017, 518 (1-2), 281-288.
5. Liang, M.; Fan, K.; Zhou, M.; Duan, D.; Zheng, J.; Yang, D.; Feng, J.; Yan, Y., H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection. *Proceedings of National Academy of Sciences of the United States of America* 2014, 111 (41), 14900-14905.
6. Lei, Y.; Hamada, Y.; Li, J.; Cong, L.; Wang, N.; Li, Y.; Zheng, W.; Jiang, X., Targeted tumor delivery and controlled release of neuronal drugs with ferritin nanoparticles to regulate pancreatic cancer progression. *Journal of Controlled Release* 2016, 232, 131-142.
7. Kuruppu, A. I.; Zhang, L.; Collins, H.; Turyanska, L.; Thomas, N. R.; Bradshaw, T. D., An apoferritin-based drug delivery system for the tyrosine kinase inhibitor Gefitinib. *Advanced Healthcare Materials* 2015, 4 (18), 2816-2821.
8. Yang, R.; Tian, J.; Liu, Y.; Yang, Z.; Wu, D.; Zhou, Z., Thermally induced encapsulation of food nutrients into phytoferritin through the flexible channels without additives. *Journal of Agriculture and Food Chemistry* 2017, 65 (46), 9950-9955.
9. Yang, R.; Liu, Y.; Meng, D.; Chen, Z.; Blanchard, C. L.; Zhou, Z., Urea-driven epigallocatechin gallate (EGCG) permeation into the ferritin cage, an innovative method for fabrication of protein-polyphenol co-assemblies. *Journal of Agricultural and Food Chemistry* 2017, 65 (7), 1410-1419.
10. Yang, R.; Liu, Y.; Blanchard, C.; Zhou, Z., Channel directed rutin nano-encapsulation in phytoferritin induced by guanidine hydrochloride. *Food Chemistry* 2018, 240, 935-939.
11. Wang, Q.; Zhang, C.; Liu, L.; Li, Z.; Guo, F.; Li, X.; Luo, J.; Zhao, D.; Liu, Y.; Su, Z., High hydrostatic pressure encapsulation of doxorubicin in ferritin nanocages with enhanced efficiency. *Journal of Biotechnology* 2017, 254, 34-42.
12. Pagadala, N. S.; Syed, K.; Tuszynski, J., Software for molecular docking: a review. *Biophysical Reviews* 2017, 9 (2), 91-102.
13. Subramanian, V.; Evans, D. G., A molecular dynamics and computational study of ligand docking and electron transfer in ferritins. *The Journal of Physical Chemistry B* 2012, 116 (31), 9287-9302.

14. Shahwan, M.; Khan, M. S.; Husain, F. M.; Shamsi, A., Understanding binding between donepezil and human ferritin: molecular docking and molecular dynamics simulation approach. *Journal of Biomolecular Structure and Dynamics* 2020, 1-9.
15. Liu, X.; Jin, W.; Theil, E. C., Opening protein pores with chaotropes enhances Fe reduction and chelation of Fe from the ferritin biomineral. *Proceedings of National Academy of Sciences of the United States of America* 2003, 100 (7), 3653-3658.
16. Blazkova; Nguyen, H. V.; Dostalova, S.; Kopel, P.; Stanisavljevic, M.; Vaculovicova, M.; Stiborova, M.; Eckschlager, T.; Kizek, R.; Adam, V., Apoferritin modified magnetic particles as doxorubicin carriers for anticancer drug delivery. *International Journal of Molecular Sciences* 2013, 14 (7), 13391-13402.
17. Falvo, E.; Tremante, E.; Arcovito, A.; Papi, M.; Elad, N.; Boffi, A.; Morea, V.; Conti, G.; Toffoli, G.; Fracasso, G.; Giacomini, P.; Ceci, P., Improved doxorubicin encapsulation and pharmacokinetics of ferritin-fusion protein nanocarriers bearing proline, serine, and alanine elements. *Biomacromolecules* 2016, 17 (2), 514-522.

Chapter 6

ENGINEERED HUMAN HEAVY-CHAIN FERRITINS PERFORMANCE AS AN ANTI-TUMOUR DRUG DELIVERY PLATFORM *IN VITRO* AND *IN VIVO*

6.1 Introduction

It has been reviewed in detail in **Chapter 2**, that human heavy-chain ferritin (HFn) is intensively used as an anti-tumour drug delivery system in research.¹⁻³ Its hollow sphere structure, high stability, readily available drug loading approaches and capability of functionalisation have laid a solid foundation as a highly promising anti-tumour drug delivery platform.⁴

However, in tumour therapy, its short half-life in circulation and limited tumour targeting ability have drawn back this practical application. Its half-life in circulation in Sprague Dawley rats was 2-3 h, thus requiring a frequent administration in real practice.⁵ Human transferrin receptor 1 (TfR1), which HFn selectively binds to, is ubiquitous in healthy cells such as liver tissue and blood brain barrier, which perhaps causes unwanted drug accumulation and therefore side-effects.⁶

To address 1) the short half-life, and 2) the limited tumour targeting ability of HFn, 2 functional peptides were fused to HFn subunit C-terminal to construct 3 PAS functionalised HFns (HFn-PAS, HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK). As has been mentioned in **Chapter 3** introduction, PAS peptide binds to the surrounding water molecule and enlarges particle size to extend half-life, and RGDK peptide achieves tumour targeted delivery through the selective affinity towards neuropilin-1 and integrin $\alpha\beta3/5$.^{7, 8} The GFLG and PLGLAG in HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK are enzyme cleavable sites corresponding to cathepsin B and matrix metalloproteinase-2 (MMP-2), respectively.^{9, 10} Both enzymes have been found overexpressed in tumours.

In **Chapter 3**, HFn and 3 PAS functionalised HFns were expressed in soluble self-assemblies. In **Chapter 4**, they were further purified through a two-step reproducible and scalable purification process to remove host cell proteins (HCPs) and host cell nuclear acid in high efficiency. In **Chapter 5**, a wide spectrum anti-tumour drug, doxorubicin hydrochloride (DOX) was loaded onto 3 PAS functionalised HFns and HFn by thermally induced passive diffusion. The impacts of PAS and RGDK functionalisation on HFn in expression, purification, conformation and drug loading process have already been studied in detail. However, the

engineered HFns performance as an anti-tumour drug delivery platform in *vitro* and *in vivo* remains uncertain.

Therefore, the aims of this chapter are to 1) investigate the mechanism behind the anti-tumour performance difference, and 2) compare HFn and functionalised HFns anti-tumour performance after being loaded with DOX. *In vitro* and *in vivo* tests were employed.

To investigate the mechanism behind RGDK peptide impact on HFn anti-tumour drug delivery performance, intracellular distribution assay and cytotoxicity assay were conducted in 3 tumour cell lines, 4T1, MDA-MB-231 and MCF7. 4T1 is BALB/c mice breast tumour cell line. It has been validated to overexpress neuropilin-1 and integrin $\alpha\beta3/5$, but does not express human TfR1.^{11, 12} MDA-MB-231 and MCF7 are human breast tumour cells and have been proven to overexpress human TfR1, express neuropilin-1 and integrin $\alpha\beta3/5$, with MDA-MB-231 possessing greater expression levels of neuropilin-1 and integrin $\alpha\beta3/5$ than MCF7.¹³⁻¹⁵ The different expression levels of related receptors in cell lines would result in performance differences between HFn-based proteins, because of the different cellular internalisation efficiency. Comparing results from MDA-MB-231 and MCF7 cell line can assess whether RGDK peptide can improve tumour targeting ability. Comparing 4T1 with MDA-MB-231 cell line can confirm whether RGDK functionalisation is able to compensate in tumour cell lines that HFn fails to target and achieve desirable anti-tumour efficacy. Intracellular distribution assay detected whether DOX carried by HFn-based proteins could enter nucleus to kill tumour cells and roughly compared *in vitro* drug accumulation in cell nucleus. IC₅₀ against 3 tumour cell lines was determined in cytotoxicity test to compare *in vitro* anti-tumour proliferation ability of drug carried by each HFn-based protein. Cellular uptake test was conducted in 4T1 cell line. It aimed to prove RGDK peptide's function, enhancement of tumour cell uptake efficiency.

To compare HFn and functionalised HFns anti-tumour performance *in vivo* after being loaded with DOX, pharmacokinetic study was conducted to examine the function of PAS peptide by determination of half-lives in circulation of DOX carried by HFn and PAS functionalised HFns. 4T1 tumour cell line was selected in biodistribution and *in vivo* anti-tumour inhibition assay to

mimic tumours where HF_n cannot achieve desirable targeting ability, and test RGDK and PAS peptide function. Biodistribution study assessed tumour targeting ability by monitoring the accumulation of HF_n and functionalised HF_ns in tumour area over time after intravenously injection. *In vivo* anti-tumour test was conducted by comparing the tumour growth inhibition efficacy of DOX carried HF_n and functionalised HF_ns.

6.2 Materials and methods

6.2.1 Materials

4T1 cell line was purchased from ATCC (USA), MDA-MB-231 cells was purchased from Cellbank (Australia). MCF7 cell line was kindly offered by Dr Nicholas Eyre from the University of Adelaide. Doxorubicin hydrochloride (DOX) was purchased from Dalian Meilun Biotechnology (China). RPMI-1640 medium, L-15 medium, penicillin-streptomycin solution (100 ×) (PS), fetal bovine serum (FBS), 0.25 % trypsin-EDTA (1 ×) solution, Hoechst 33258 reagent and MTT reagent were purchased from Invitrogen (Thermo Fisher Scientific, USA). RGDK peptide was synthesised by Gene Script (China). Propidium iodide and trypan blue solution were bought from Sigma-Aldrich (USA). Sulfo-cy5 NHS ester was synthesised by Lumiprobe (USA). All other reagents of analytical grade were obtained from Chem-Supply (Australia). Milli Q water was obtained from a Millipore purification system (Merck, USA).

Four HF_n-based proteins (HF_n, HF_n-PAS, HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK) were purified by heat-acid precipitation followed by chromatography, as is described in **Chapter 4**. DOX loaded HF_n-based proteins (protein/DOX) were prepared as in **Chapter 5**.

6.2.2 Intracellular distribution analysis in tumour cells

DOX is an anthracycline topoisomerase inhibitor and exerts its function mainly inside cell nucleus.⁵ Free DOX directly diffuses into nucleus and disrupts cell division after internalisation, whilst the protein/DOX are supposed to first be broken down by enzymes in lysosome to release loaded DOX and then DOX reaches nucleus. Intracellular distribution test was designed to check if DOX loaded on HF_n-based proteins could enter cell nucleus to kill tumour cells. Three tumour cell lines, 4T1, MDA-MB-231 and MCF7, were selected to

compare the effect of receptor expression levels on drug distribution to understand inserted peptide function.

4T1 and MCF7 cells were cultured in RPMI-1640 medium supplemented with 10 % FBS and 1 % PS. MDA-MB-231 cells were cultured in L-15 medium with 10 % FBS and 1 % PS. All three types of cells were cultivated in Contherm Miltre 4000 culture incubator (Thermo Fisher Scientific, USA) at 37 °C in a 5 % CO₂ atmosphere. All 3 types of tumour cells in exponential growth phase were placed on 6-well plates at a density of 4×10^5 cells per well and cultured for 24 h. One cover-glass slide was put in each well prior to seeding. The medium was then discarded, and cells were treated with fresh media containing protein/DOX or free DOX (20 µM DOX-equivalent) in 2 mL per well for 3 h. After that, drugs were removed, and the cells were washed 3 times using $1 \times$ phosphate-buffered saline (PBS) and then cultured with fresh complete medium for another 36 h. Subsequently, the cells were washed 3 times with $1 \times$ PBS and fixed with 4 % paraformaldehyde for 10 min at 25 °C. Following another 3 washes with $1 \times$ PBS, cells were stained with $0.5 \mu\text{g mL}^{-1}$ Hoechst 33258 for 5 min and washed 5 times with $1 \times$ PBS, then visualized under a ZOE fluorescence cell imager (Bio-Rad, USA). Images of cells under green channel (Excitation: 480/17 nm, Emission: 517/23 nm) and blue channel (Excitation: 355/40 nm, Emission: 433/36 nm) were then captured.

6.2.3 Cytotoxicity against tumour cells

The cytotoxicity of 4 protein/DOX and free DOX against 4T1, MDA-MB-231 and MCF7 cell lines was evaluated by MTT assay. Exponential growth-phase cells were digested by 0.25 % trypsin-0.05 % EDTA, and the cell density was adjusted to 1×10^5 cells per mL by complete medium. 100 µL of cells were seeded to each well in 96-well plates. After incubation for 24 h, previous medium was replaced with fresh complete medium separately containing either free DOX or protein/DOX, whose concentrations of DOX were 0.0096, 0.048, 0.24, 1.2, 6 and 30 µg mL⁻¹. Each concentration was repeated in 4 wells. After incubation with drug for 60 h, the media were removed, and cells were washed 3 times by $1 \times$ PBS. Then, 90 µL of fresh complete medium with 10 µL of MTT solution was added to each well and incubated for another 4 h. 100 µL dimethyl sulfoxide (DMSO) was added to wells to ensure complete solubilisation of the formed formazan crystals. Finally, the absorbance of each well was measured at 595 nm (background: 630 nm) by a Microplate Reader (Biotek, USA). 4 wells with 0 µg mL⁻¹ DOX

equivalent incubation were regarded as cell control wells and 4 wells without cells were regarded as blank control wells. Absorbance and cell viability of each well was calculated by the **Equation 6.1** and **6.2**:

$$A_{\text{well}} = A_{595} - A_{630} \quad (6.1)$$

$$\text{Cell viability (\%)} = (A_{\text{well}} - A_{\text{blank}})/(A_{\text{cell}} - A_{\text{blank}}) \times 100 (\%) \quad (6.2)$$

6.2.4 Cellular uptake assay of 4T1 cell line

Cellular uptake test procedure has been adapted from a previous paper on 4T1 cell line.¹⁶ 4T1 cell line was selected because it does not express human Tfr1 and the results specifically reflect RGDK related receptor effect. For each protein/DOX group, 3 different treatments were done to obtain 3 fluorescence intensities, total fluorescence, internalised fluorescence, fluorescence after RGDK peptide pre-incubation.

The procedure was as followings: 1) Cell seeding. 4T1 Cells in the exponential growth phase were seeded in 24-well plates at a density of 1×10^5 cells per well and cultured for 48 h for attachment. 2) RGDK peptide pre-incubation in wells for fluorescence after RGDK peptide pre-incubation determination. To investigate the impact of fused RGDK on cellular uptake characteristics, 500 μM free RGDK peptide was pre-incubated with the cell for 1 h at 37 °C to saturate RGDK specific receptors. 3) Drug incubation. The media with or without RGDK peptide in all wells were discarded and cells were washed with PBS 3 times, prior to adding 100 μL serum-free culture medium containing free DOX or protein/DOX (15 $\mu\text{g mL}^{-1}$ DOX-equivalent). Then the cells were incubated for 90 min at 37 °C and washed 3 times with PBS to remove drugs. 4) Trypan blue quenching. In wells for internalised fluorescence and fluorescence after RGDK peptide pre-incubation determination, cells were incubated with trypan blue (0.25 % in 0.85 % NaCl) for 5 min at 25 °C, and then washed 5 times with PBS to remove trypan blue. 5) Detachment of cells for flow cytometry analysis. 400 μL of 0.25 % trypsin-0.05 % EDTA solution was added to all wells for digestion for 5 min at 37 °C and 2 mL of complete medium was added to stop the digestion. Detached cells were spun at 1,000 rpm for 3 min at 4 °C and resuspended in 1 mL PBS. 5 μL working propidium iodide (PI) was added to incubate with cells for 10 min at 25 °C for differentiation of alive and dead cells in flow cytometry detection. 6) Flow cytometry analysis. A Csampler flow cytometry (Becton

Dickinson, USA) was employed to determine the mean fluorescence of 5,000 cells in each sample. A cell control underwent PI staining but without drug incubation, trypan blue and RGDK peptide treatment was used for gating and parameter setting prior to sample detection. PE channel (Excitation laser light: 488 nm, Emission: 578 nm) was utilised for DOX fluorescence detection. Mean fluorescence intensity of each sample was recorded.

6.2.5 Pharmacokinetic study

All animal experiments were performed with the approval of the medical ethics committee of Shanxi University of Chinese Medicine (Approval Number 2019LL137). 18 Male Sprague Dawley rats weighing 230-250 g (specific-pathogen free, SPF Biotechnology Co., Ltd. Beijing, China) were randomly assigned to 6 groups, and administrated with PBS, free DOX and protein/DOX (3.0 mg kg⁻¹ DOX equivalent) separately via intravenous injection at tail vein. After injection, blood samples were collected from the retro orbital sinus at fixed time points (10, 30 min, 1, 2, 4, 8, 12, 24, 36, 48 h) and followed by clotting for at least 0.5 h at 37 °C. Serum was obtained by centrifugation at 6,000 rpm for 30 min at 4 °C. Finally, 100 µL of serum of each sample was transferred to a 96-well microplate, and the DOX contents were determined using SpectraMax i3x microplate reader (Molecular devices, USA). Excitation and emission wavelength were set at 480 nm and 580 nm, respectively. Meanwhile, the standard curve of the fluorescence intensity with varying concentration of DOX in rat serum was also measured for quantitative analysis.

6.2.6 Biodistribution study in 4T1 tumour bearing mice

As 4T1 is a BALB/c breast tumour cell line, female BALB/c mice were chosen to establish tumour-bearing animal model. 1×10^6 4T1 cells in 100 µL of PBS were injected into right armpit of 8-week-old female BALB/c mice (specific-pathogen free, SPF Biotechnology Co., Ltd. Beijing, China) to form mice tumour model for both biodistribution study and anti-tumour growth study. All 4 HFn-based proteins were first labelled by Sulfo-cy5 NHS ester with a molar ratio of 1:30 (Protein to Cy5) and the uncoupled Cy5 was removed by Hitrap G25 desalting column (GE Healthcare, USA). When tumour volume reached about 300 mm³, a 150 µL sample of Cy5 or protein-Cy5 conjugates (0.2 mg kg⁻¹ Cy5 equivalent) was intravenously injected into the tumour-bearing mice via tail vein. After treatment, the mice were anesthetized using isoflurane at 2, 4, 6.5, 12, 24, and 52 h and fluorescence images were taken under

excitation wavelength of 646 nm and emission wavelength of 662 nm using FX Pro *in vivo* imaging system (Bruker BioSpin, USA).

6.2.7 Anti-tumour study in 4T1 tumour bearing mice

1×10^6 4T1 cells in 100 μL of PBS were injected into right armpit of 8-week old female BALB/c mice (specific-pathogen free, SPF Biotechnology Co., Ltd. Beijing, China). For *in vivo* tumour growth inhibition assessment, female BALB/c mice bearing 4T1 tumours of approximate 250 mm^3 in size were randomly assigned to 6 groups ($n = 6$ mice in each group) and treated with protein/DOX (3 mg kg^{-1} DOX equivalent), free DOX (3 mg kg^{-1}), or PBS via 200 μL intravenous injection. The drug injection was carried out every 4 days for 2 doses. The volumes of tumours were measured every other day. Mice were monitored for up to 17 days post-implantation and then sacrificed. Primary tumours were harvested for *ex vivo* imaging.

6.2.8 Statistical analysis

All data are expressed as mean \pm standard deviation (SD). The statistical significance was assessed *via* unpaired T test and defined as $P > 0.05$, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. IC_{50} values were calculated using Origin 9.0 software. Half-life of all protein/DOX was calculated using DAS 2.0 software and fitting data in a single-compartment mode.

6.3 Results and discussion

6.3.1 DOX intracellular distribution after cellular internalisation in 3 cell lines

Intracellular distribution analysis was conducted to prove whether DOX could accumulate in tumour cell nucleus. In **Figure 6.1**, blue colour indicates cell nucleus and green colour represents the fluorescence from DOX. Clearly, the cyan colour in merged images demonstrates that in all 4 protein/DOX groups, DOX has entered and accumulated inside cell nucleus of all 3 tumour cell lines. Amongst all groups in all cell lines, the drug fluorescence detected in free DOX group was the strongest. This is attributed to the advantage of the highly efficient small molecule passive diffusion in direct incubation with tumour cells. This advantage only occurs in *in vitro* tests because of the lack of biodistribution, metabolism and elimination process. In contrast, protein/DOX theoretically enter tumour cells through receptor-mediated mechanism, which is relatively slow *in vitro*.

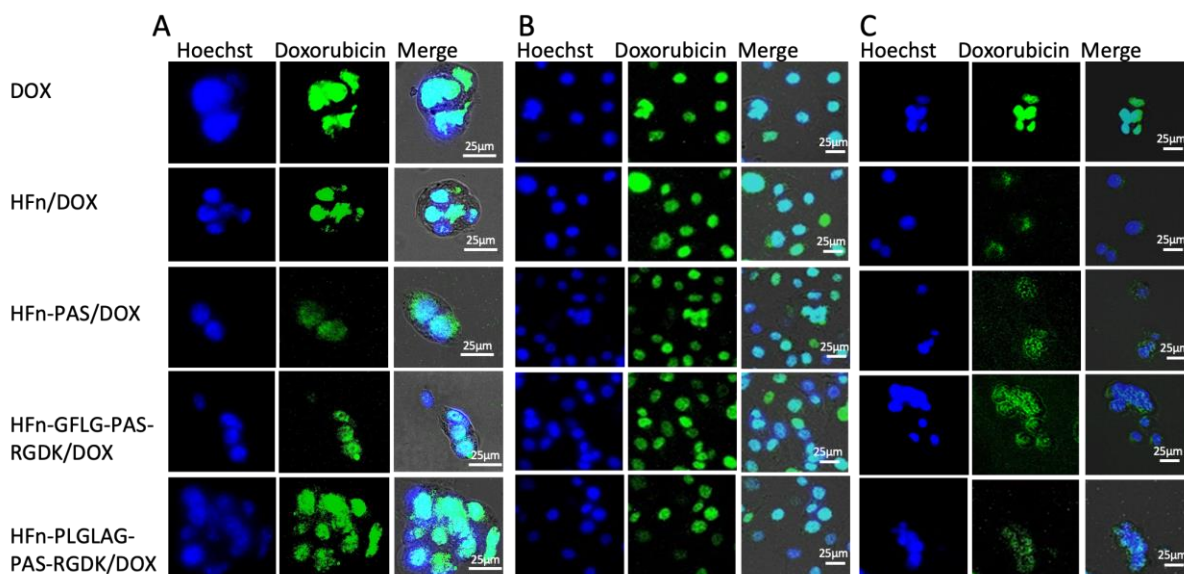


Figure 6.1 Intracellular distribution of DOX in tumour cells.

Under the Bio-Rad Zoe cell imager, blue: nucleus after being stained with Hoechst 33258. Green: DOX because of its intrinsic fluorescence. Cyan: merged fluorescence signal.

A, 4T1 cells. **B**, MDA-MB-231 cells. **C**, MCF7 cells.

Comparing 3 cell lines, in 4T1 and MDA-MB-231 cell lines, which both have greater expression level of integrin $\alpha\beta3/5$ and neuropilin-1, drug fluorescence intensity detected in HFn-GFLG-PAS-RGDK/DOX and HFn-PLGLAG-PAS-RGDK/DOX were significantly greater (**Figure 6.1A** and **B**) than those in MCF7 cell line (**Figure 6.1C**), which has lower expression level. It could be inferred that the uptake efficiency of protein/DOX depends on the expression level of integrin $\alpha\beta3/5$ and neuropilin-1.

6.3.2 Functionalisation impact on anti-proliferation efficacy of protein/DOX *in vitro*

In order to test the inhibition of 4 HFn-based proteins/DOX on tumour cell proliferation, MTT assay was adopted. As is shown in **Figure 6.2**, all protein/DOX groups demonstrate anti-proliferation abilities against 3 types of tumour cells. In all cell lines, the order of IC_{50} value of each group are the same (**Table 6.1**). Free DOX group has the lowest IC_{50} and t-test show it was significantly lower than all the rest groups in all cell lines. This is due to its relatively high cellular internalisation efficiency through passive diffusion *in vitro*, as is also shown in intracellular distribution result. Inhibition impact of HFn-GFLG-PAS-RGDK/DOX and HFn-PLGLAG-PAS-RGDK/DOX on all tumour cells growth were the second strongest. T-test shows that significant differences exist when comparing 2 PAS-RGDK functionalised HFn/DOX groups with HFn/DOX and HFn-PAS/DOX groups ($P < 0.05$). This demonstrates

the RGDK peptide in HF_n-GFLG-PAS-RGDK and HF_n-PLGLAG-PAS-RGDK has significantly enhanced HF_n performance in terms of drug cytotoxicity against all 3 cell lines. No statistical difference was found in 2 PAS-RGDK fused HF_ns in all cell lines. HF_n-PAS/DOX ranked the third, and HF_n/DOX showed the worst anti-proliferation effect in all cell lines. In MDA-MB-231 and MCF7 cell lines, no statistical difference was found between HF_n-PAS/DOX and HF_n/DOX. This could be explained by the mechanism of PAS peptide. It requires a blood circulation to achieve half-life extension.

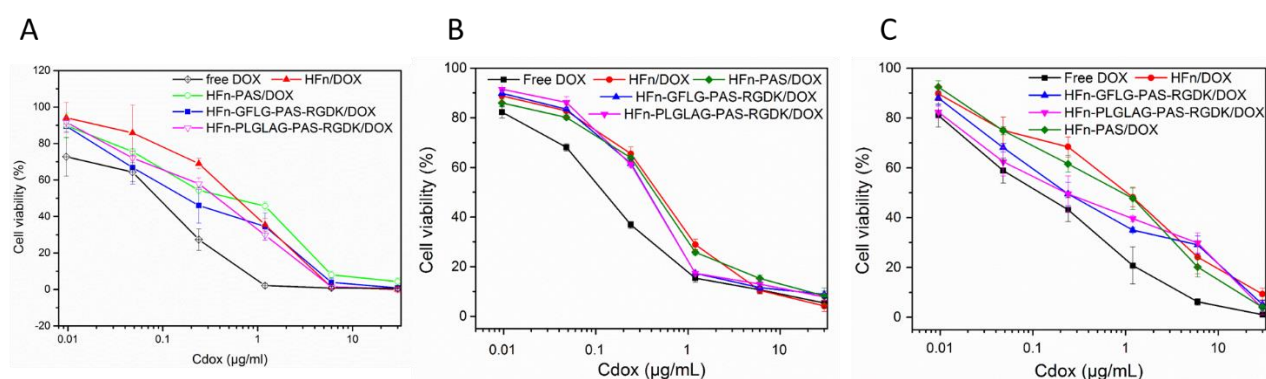


Figure 6.2 Proliferation inhibition effects on tumour cells.

Data were mean \pm standard deviation (n = 4).

A, 4T1 cells. B, MDA-MB-231 cells. C, MCF7 cells.

Table 6.1 IC₅₀ values of all groups.

Group	IC ₅₀ ($\mu\text{g mL}^{-1}$) against 4T1	IC ₅₀ ($\mu\text{g mL}^{-1}$) against MDA-MB-231	IC ₅₀ ($\mu\text{g mL}^{-1}$) against MCF7
DOX	0.08 \pm 0.03	0.15 \pm 0.01	0.20 \pm 0.03
HF _n /DOX	0.49 \pm 0.11	0.57 \pm 0.02	1.31 \pm 0.22
HF _n -PAS/DOX	0.38 \pm 0.09	0.46 \pm 0.01	0.98 \pm 0.14
HF _n -GFLG-PAS-RGDK/DOX	0.17 \pm 0.01	0.34 \pm 0.01	0.59 \pm 0.05
HF _n -PLGLAG-PAS-RGDK/DOX	0.18 \pm 0.04	0.33 \pm 0.01	0.55 \pm 0.10

IC₅₀ values of DOX and protein/DOX were different in 3 cell lines (**Table 6.1**). 4T1 cell line had the lowest free DOX IC₅₀, showing it was the most sensitive cell line to DOX. MDA-MB-231 was the second sensitive cell line, and MCF7 was the least. In comparison of protein/DOX IC₅₀, 4T1 cell line has the lowest values in all 4 groups, and MBA-MD-231 cell line values

were lower than those of MCF7. HFn/DOX and HFn-PAS/DOX IC₅₀ values in MCF7 were significantly greater than those in the other cell lines. The IC₅₀ values of 2 PAS-RGDK functionalised HFn/DOX in 4T1 cell line were approximate a half of those in MDA-MB-231 and a third of those in MCF7. The protein/DOX cytotoxicity difference between cell lines stems from a combinative effect of their expression levels of integrin $\alpha\beta 3/5$, neuropilin-1 and TfR1 receptors, and their sensitivity towards DOX. The greater the receptor expression level the cell line has and the more sensitive it is to DOX, the lower the IC₅₀ of protein/DOX would be. Overall, 2 PAS-RGDK functionalised HFn/DOX and HFn-PAS/DOX achieved the best anti-proliferation effect on 4T1 cell line.

6.3.3 RGDK functionalisation impact on 4T1 uptake efficiency of protein/DOX *in vitro*

In cellular uptake test, the DOX uptake efficiencies carried by 4 HFn-based proteins by 4T1 cells to study the RGDK functionalisation effect were compared. Total fluorescence of DOX measured in flow cytometry came from 2 sources, DOX internalised by cells and un-specifically bound to cell membranes. Trypan blue treatment quenched the signal from membrane bound DOX, and lower mean fluorescence intensities are observed in all groups (**Figure 6.3**).

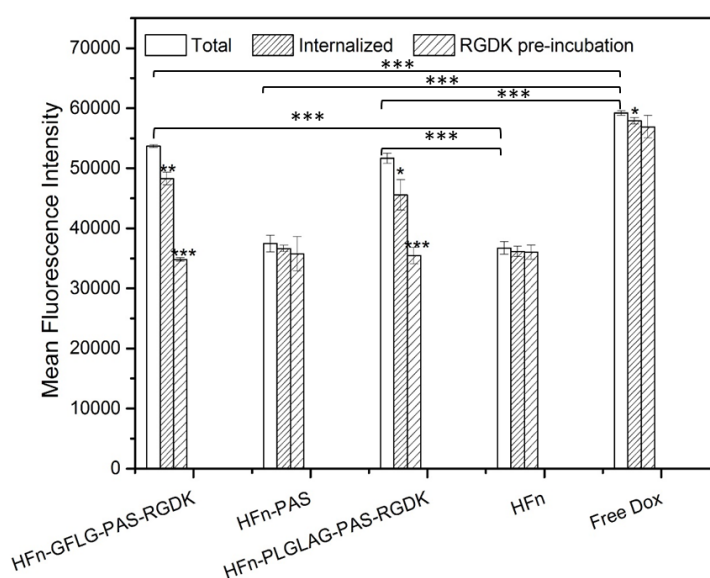


Figure 6.3 Mean DOX fluorescence intensity in cellular uptake test.

Data were represented as mean \pm standard deviation (n = 3), *P < 0.05, **P < 0.01, ***P < 0.001. *on top of the column represents the significance of P value between this column and the white column (total DOX fluorescence) in the same group.

With regard to total fluorescence and internalised fluorescence intensity, free DOX shows significantly greater internalised cellular uptake than others, in accordance with the findings in intracellular distribution and MTT assays. DOX is a small molecule and enters cells via passive diffusion. HFn-GFLG-PAS-RGDK/DOX and HFn-PLGLAG-PAS-RGDK had the second highest efficiencies and were significantly different from the HFn/DOX and HFn-PAS/DOX. Because 4T1 does not express human TFR1, HFn/DOX and HFn-PAS/DOX enter the cell through pinocytosis.

RGDK peptide pre-incubation treatment, with the use of excessive amount of free RGDK, is intended to mask RGDK-specific receptors, integrin $\alpha\text{v}\beta\text{3}/\text{5}$ and neuropilin-1, on cells to hamper RGDK-related cellular uptake. As is demonstrated in **Figure 6.3**, the uptake of 2 PAS-RGDK functionalised HFns groups, were significantly inhibited by RGDK peptide pre-incubation while in other groups no obvious difference occurred. This proves that it was the RGDK peptide that helped improve 4T1 cells internalisation of HFn-GFLG-PAS-RGDK/DOX and HFn-PLGLAG-PAS-RGDK/DOX.

6.3.4 PAS Functionalisation impact on pharmacokinetic profile *in vivo*

Pharmacokinetic profile of all 4 protein/DOX and free DOX were obtained through tail vein injection of healthy Sprague Dawley rats. Standard curve for DOX concentration calculation is in **Appendix B, Figure B11**. Line chart of DOX concentrations in plasma over time (10 min - 48 h) is shown in **Figure 6.4**, and half-life in circulation of all protein/DOX are listed in **Table 6.2**. Plasma drug concentrations of free DOX group rats reduced rapidly right after administration (**Figure 6.4**). At 10 min, average plasma drug concentration was $15 \mu\text{g mL}^{-1}$. 8 h later, almost all free DOX was eliminated from circulation. In terms of HFn/DOX group rats, the plasma drug cleaning out speed ranked the second, with average drug concentration of $21.6 \mu\text{g mL}^{-1}$ at 10 min and below $5 \mu\text{g mL}^{-1}$ after 12 h. Average plasma drug concentrations in 3 functionalised HFn/DOX group rats (approximate $35 \mu\text{g mL}^{-1}$) were more than double of those in free DOX group at 10 min. 48 h after administration, more than $5 \mu\text{g mL}^{-1}$ drug still remained in plasma of functionalised HFn/DOX group rats.

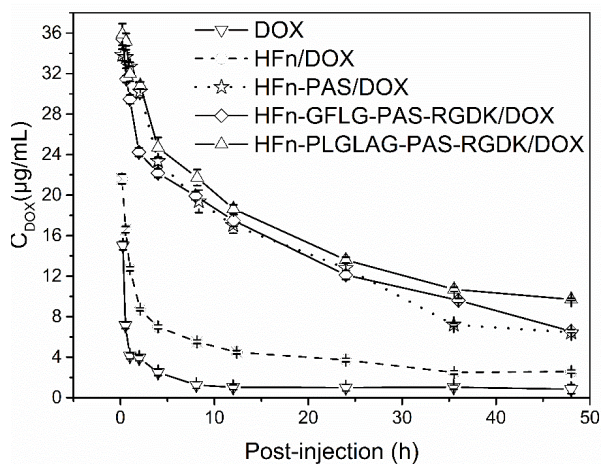


Figure 6.4 Plasma concentrations of protein/DOX and free DOX in SD rats of different groups. Data were expressed as mean \pm SD (n = 3).

Table 6.2 Half-life in circulation of each protein/DOX in SD Rats (n = 3)

Sample	$T_{1/2}$ (h)
DOX	0.42 \pm 0.03
HFn/DOX	3.07 \pm 0.06
HFn-PAS/DOX	14.96 \pm 0.29
HFn-GFLG-PAS-RGDK/DOX	17.61 \pm 0.39
HFn-PLGLAG-PAS-RGDK/DOX	18.93 \pm 0.61

Single compartment fitting of the drug concentration-time curve was applied to evaluate drug half-life in circulation. The results showed free DOX only had about 25 min of half-life in circulation (**Table 6.2**), similar to reported value in previous research.¹⁷ HFn/DOX half-life, approximate 3 h, was 7.3 times of free DOX. PAS peptide in HFn-PAS/DOX has increased half-life almost 4.9 times (14.96 h) compared with HFn/DOX and the extra residues in HFn-GFLG-PAS-RGDK/DOX further extended the half-life to 17.61 h. HFn-PLALGA-PAS-RGDK/DOX possessed the longest half-life, 18.93 h. Differences in half-life of all PAS functionalised HFns compared with HFn and free DOX are statistically significant in t-test ($P < 0.001$).

Comparing the differences of half-life in circulation between groups, the increase of half-life mainly came from the insertion of PAS peptide. Constituted by repetitive P, A and S residues,

PAS peptide is hydrophilic and uncharged in neutral solutions and plasma. Circular dichroism shows it is a flexible random coil.⁷ Its properties are similar to PEG, but it is advantageous in terms of biodegradability and biocompatibility. When it is attached to HFn, it can attract more water molecules to increase molecule hydrodynamic volume. PAS peptide usually has to be over 200 residues to achieve half-life extension, but because of the repetitive and organized presentation manner on ferritin shell, 40 aa and 75 aa PAS peptides have been proven to be long enough to significantly increase half-life when fused onto N-terminal of ferritin subunit.¹⁷

6.3.5 Functionalisation impact on protein biodistribution *in vivo*

To monitor biodistribution of all HFn-based proteins in 4T1 tumour-bearing mice after tail vein administration over time, *in vivo* imaging was used. In this analysis, fluorescence label Cy5 was attached to all employed proteins and free cy5 worked as control. Real-time biodistribution of Cy5 attached proteins and free Cy5 were visualised in BALB/c mice with 4T1 tumour in right armpit, and fluorescence intensities of tumour areas were recorded. Two control groups, mice injected with free cy5 and HFn-cy5 were scanned at the same time, and mice in other 3 groups (HFn-PAS, HFn-PLALGA-PAS-RGDK and HFn-GFLG-PAS-RGDK) were scanned together.

As is shown in **Figure 6.5A**, free cy5 preferred to accumulate in liver, perhaps because liver is the main organ for metabolism. After 12 h, fluorescence could not be captured by camera. Theoretically, as a nanoparticle, HFn has passive tumour targeting ability. However, from the results in **Figure 6.5A**, signal of fluorescence was observed in liver rather than in tumour at 2 to 12 h. It seems that HFn-cy5 did not show desirable tumour targeting ability and it preferred liver. Probably the particle size of HFn is still too small to achieve desirable passive tumour targeting ability. HFn-PAS-cy5 did not show any obvious accumulation in mice at all times in **Figure 6.5A**, but as is shown in **Figure 6.5B**, there actually was fluorescence detected in tumour area. Perhaps because of the sharp contrast between signal intensities of HFn-PAS-cy5 and 2 PAS-RGDK fused proteins, lower intensity of HFn-PAS-cy5 failed to be captured by the machine camera under the same exposure time. In 2 PAS-RGDK fused HFns, proteins were distributed in both tumour and liver, but tumour area always had stronger signals (**Figure 6.5A**). In addition, 52 h after injection, whilst HFn-cy5 and free cy5 were almost completely

eliminated, PAS-RGDK fused proteins were still detectable in region of tumour site, implying functionalised HFns were retained in tumour by longer and stronger accumulation.

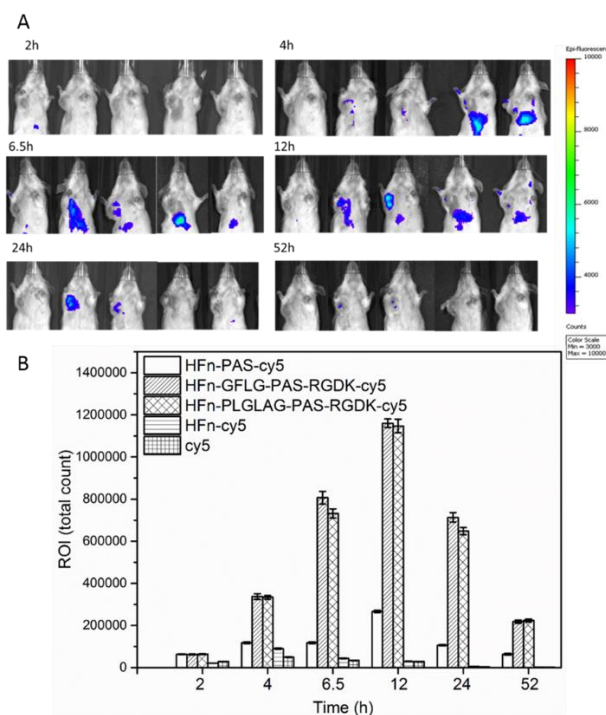


Figure 6.5 Biodistribution results of HFn and PAS functionalised HFns in 4T1 tumour bearing mice. **A**, in vivo fluorescence imaging of tumour-bearing mice at different time points, from left to right: HFn-PAS-cy5, HFn-GFLG-PAS-RGDK-cy5, HFn-PLGLAG-PAS-RGDK-cy5, HFn-cy5 and free cy5. **B**, the sum fluorescent intensity of region of interest (ROI, tumour area) at each time point.

As is presented in **Figure 6.5B**, free cy5 and HFn-cy5 always had the lowest tumour fluorescence intensity. At 2 h, free cy5 had a greater intensity than HFn-cy5 but was surpassed by HFn-cy5 afterwards. Free cy5 tumour area fluorescence intensity peaked at 4 h and decreased rapidly after that, suggesting a fast clearance. HFn-cy5 achieved the highest concentration in tumour at around 4 h after injection (**Figure 6.5B**). The difference in free cy5 and HFn-cy5 is likely to be due to a quicker distribution and a shorter half-life of small molecule cy5 than HFn nanoparticle. HFn-PAS-cy5 demonstrated significantly stronger and longer lasting tumour intensities than HFn-cy5 at all detected time points ($P < 0.001$). As proven in the pharmacokinetic study, the insertion of the PAS peptide could lead to a longer half-life in circulation and probably result in the slower clearance of HFn-PAS-cy5 than HFn-cy5. The best drug targeting delivery results are from 2 PAS-RGDK functionalised HFns. They had significantly greater signal intensities in tumour area at all times than all the other groups ($P < 0.001$). This shows that the RGDK peptide can significantly improve HFn biodistribution. A previous study of RGDK fused Albumin binding domain has also proven the tumour

targeting ability improvement of RGDK peptide *in vivo*.⁵ Overall, both PAS and RGDK functionalisation, and particularly RGDK functionalisation, improved the tumour biodistribution of HFn.

6.3.6 Functionalisation impact on protein/DOX anti-tumour efficacy *in vivo*

To compare tumour treatment efficacies of all protein/DOX and free DOX, 4T1 tumour bearing BALB/c mice model was built and 36 mice with around 250 mm³ tumour were randomly assigned into 6 groups. Intravenous injections of 4 HFn-based protein/DOX, free DOX and PBS were conducted at day 0 and day 5. As is shown in **Figure 6.6A**, the fastest mice tumour growth rate was observed in PBS control group rats which underwent no drug treatment. The average tumour volume reached 2,030 mm³ after 17 days. The second fastest tumour growth rate was in free DOX group mice and their average group tumour volume were 1,667 mm³ at day 17. HFn/DOX showed a better tumour growth inhibition and at day 17, tumour volume grew to 1,521 mm³. In HFn-PAS/DOX group, mice tumour volume reached 1,432 mm³ in the end. 2 PAS-RGDK functionalised HFn/DOX treated group had the strongest tumour-growth inhibition. In spite of just 2 administrations, average tumour volume of these 2 group mice at day 17 were just around 1,100 mm³, close to half of the volume of PBS group tumour.

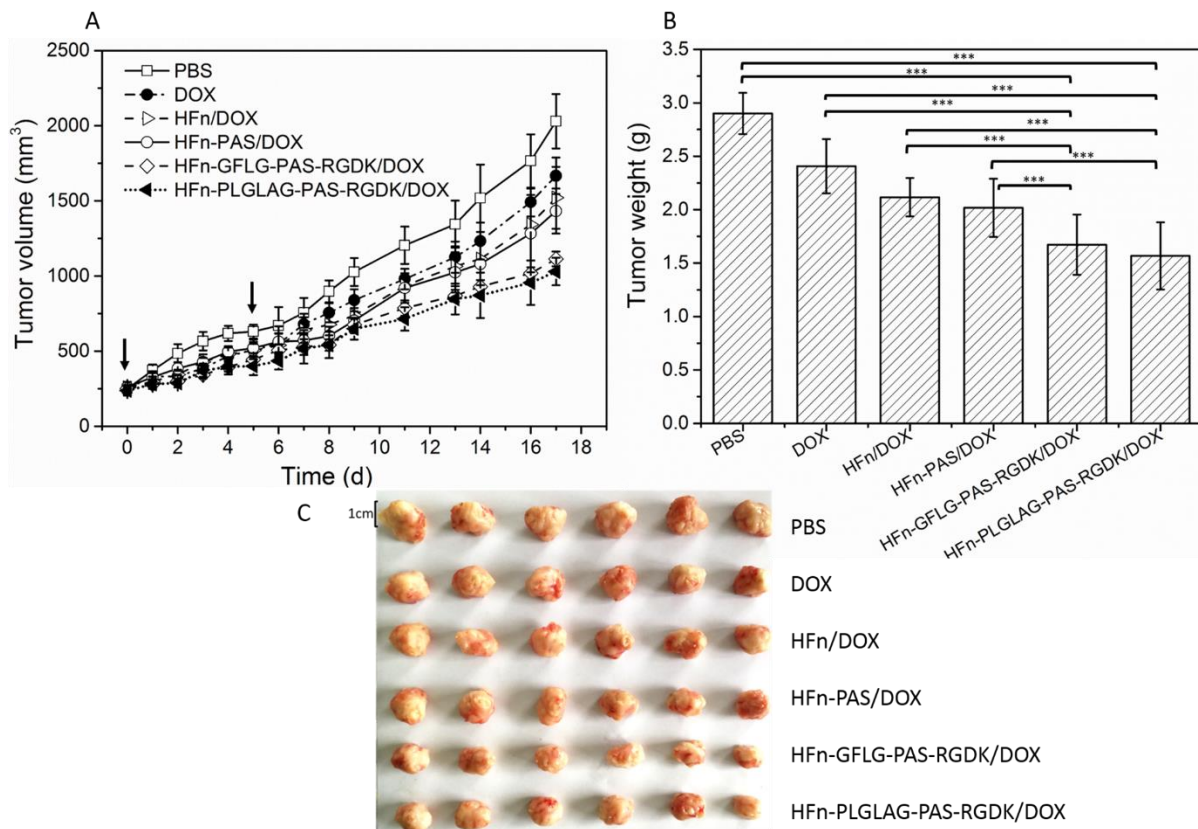


Figure 6.6 *In vivo* tumour inhibitory effects on 4T1 tumour-bearing mice. **A**, tumour volume change over time, arrows indicated the injection days. **B**, the average tumour masses on Day17. **C**, the photo of excised tumour tissues on Day17. Data are mean \pm RSD (n = 6). **P < 0.01, ***P < 0.001.

On day 17, the corresponding average tumour weights were measured (**Figure 6.6B**) and the photo of excised tumour tissues were captured (**Figure 6.6C**). T-test results demonstrate that there were significant statistical differences between final tumour masses of protein/DOX and free DOX group. Both HFn and functionalised HFns had significantly increased DOX anti-tumour efficacy ($P < 0.001$). Compared with HFn/DOX group, HFn-PAS/DOX did not show statistical distinction ($P = 0.468$), showing that PAS functionalisation alone was not enough to significantly improve anti-tumour efficacy. Masses of tumours from 2 PAS-RGDK protein/DOX groups, however, were significantly lower than those of both HFn/DOX group and HFn-PAS/DOX group. This indicates RGDK functionalisation primarily accounts for the significant improvement of growth inhibition efficacy of 4T1 tumour. Difference between HFn-GFLG-PAS-RGDK/DOX and HFn-PLGLAG-PAS-RGDK/DOX ($P = 0.977$) was not significant in statistical analysis. Two different enzyme-cleavable sites did not make a difference in anti-tumour efficacy.

Based on all 6 assessments in this chapter, PAS and RGDK functionalisation have both improved HFn anti-tumour performance. PAS functionalisation impacts HFn mainly by extension of half-life in circulation and it was because of an enlargement of the hydrodynamic volume, as was characterised in **Chapter 4**. RGDK functionalisation improves tumour targeted delivery through its specific affinity to integrin $\alpha\beta3/5$ and neuropilin-1 receptor. A schematic illustration of molecular mechanism behind RGDK functionalisation impact on anti-tumour efficacy is presented in **Figure 6.7**. It is the assumed tumour cell internalisation pathways of HFn and functionalised HFns in human tumour tissues. Free DOX enters tumour cells via passive diffusion due to its small size (**Figure 6.7A**). Because 4T1 tumour cell does not overexpress human TfR1, HFn and HFn-PAS were likely to enter cells through non-specific pinocytosis as in other nanoparticles (**Figure 6.7B**).¹⁸

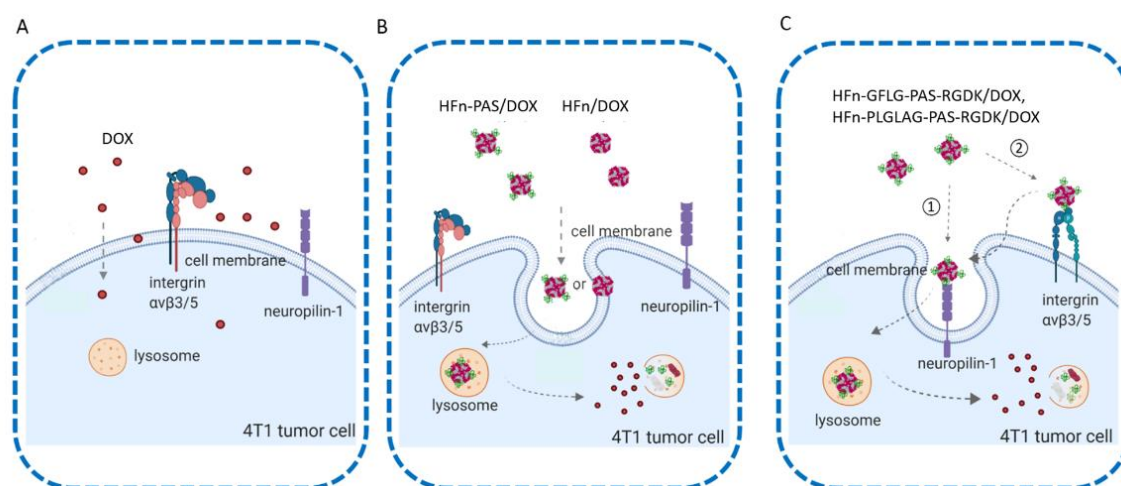


Figure 6.7 Schematics of different tumour cellular internalization mechanisms of DOX and protein/DOX.

A, free DOX passive diffusion. **B**, HFn/DOX and HFn-PAS/DOX pinocytosis internalization. **C**, two possible receptor-mediated internalization pathways of HFn-GFLG-PAS-RGDK/DOX and HFn-PLGLAG-PAS-RGDK. (Created with Biorender).

HFn-GFLG-PAS-RGDK/DOX and HFn-PLGLAG-PAS-RGDK/DOX have different drug internalisation mechanisms from free DOX, HFn/DOX and HFn-PAS/DOX. In theory, GFLG responds to cathepsin B, which is located in perinuclear vesicles and vesicles, functional peptide PAS-RGDK is supposed to be removed after HFn-GFLG-PAS-RGDK/DOX enters tumour cell.^{9, 19} The other enzyme cleavable linker PLGLAG is specifically cleaved by MMP-2/9, which are secreted outside tumour cells.^{20, 21} However, the two enzyme-cleavable sites, GFLG and PLGLAG, did not show any significant differences in cellular uptake efficiency,

and in any other *in vitro* and *in vivo* tests. This suggests the PLGLAG enzyme-cleavable site was probably not digested by MMP-2/9 before cell internalisation. It could be caused by the insufficient activity of MMP-2/9 *in vitro* and *in vivo* and/or the low accessibility of PLGLAG to enzymes. In this case, HFn-GFLG-PAS-RGDK/DOX and HFn-PLGLAG-PAS-RGDK were very likely to follow the same cellular internalisation pathways, as shown in **Figure 6.7C**.

Both made use of the RGDK peptide to be internalised by tumour cells. The cellular internalisation mechanism of iRGD, RGDK's cyclic analog, has been well studied and it is reasonable to assume RGDK follow a similar pattern. Since RGDK has RGD motif and an exposed free C-terminal K residue, it can be directly recognized by both integrin $\alpha\beta3/5$ and neuropilin-1. As shown in **Figure 6.7C**, when HFn-GFLG/PLGLAG-PAS-RGDK/DOX reaches tumour tissue, there are 2 possible internalisation pathways. RGDK can either directly bind to NRP1 or firstly interact with integrin $\alpha\beta3/5$ and then be transferred to NRP1, followed by endocytosis. After that, some of the HFn-GFLG/PLGLAG-PAS-RGDK/DOX inside tumour cells would be digested in lysosome while some will travel to other cells nearby via paracellular pathway or transcytosis.²²

Regarding the improved tumour targeting ability and anti-tumour efficacy of PAS-RGDK dually functionalised HFns, it was reckoned that some or all of functional peptides (PAS-RGDK) at subunit C-terminal should be accessible to receptors and are extruded outside of HFn cage instead of buried inside. This is in accordance with the flipped ferritin conformation characterised in **Chapter 4**. However, fully flipped or partially flipped conformation of functionalised HFns need to be further confirmed.

6.4 Conclusion

Conclusions from this chapter include:

1. DOX carried by HFn and 3 PAS functionalised HFns could be released and accumulate in cell nucleus as free DOX to kill tumour cells.

2. Both PAS and RGDK functionalisation have improved HF_n anti-tumour performance. RGDK peptide has significantly promoted DOX internalisation efficiency to tumour cells because of its selective affinity of integrin $\alpha\beta 3/5$ and neuropilin-1 receptor. It has significantly decreased the IC₅₀ values of DOX in 4T1, MDA-MB-231 and MCF7 cell lines, enhanced tumour biodistribution, and also significantly inhibited tumour growth in 4T1 tumour bearing BALB/c mice.

3. PAS has improved HF_n performance by increasing HF_n half-life around 4.9 times. It also improved HF_n biodistribution because of the half-life extension.

4. PAS functionalised HF_ns are probably in flipped conformation based on the assessment results, as was also shown in **Chapter 4**.

5. Two enzyme-cleavable sites GFLG and PLGLAG did not make any significant differences in all *in vitro* and *in vivo* tests. Results suggests that PLGLAG might not be cleaved by MMP-2/9. A further detailed investigation of the cleavage of PLGLAG is needed.

6. Overall, 3 PAS functionalised HF_n constructs (HF_n-PAS, HF_n-GFLG-PAS-RGDK, HF_n-PLGLAG-PAS-RGDK), especially 2 PAS-RGDK fused HF_ns, hold great potentials as anti-tumour drug delivery nanoparticles.

6.5 References

1. A.Luzzago, G. C., Isolation of point mutations that affect the folding of the H chain of human ferritin in *E.coli*. The EMBO Journal 1989, 8 (2), 569-576.
2. Jutz, G.; van Rijn, P.; Santos Miranda, B.; Boker, A., Ferritin: a versatile building block for bionanotechnology. Chemical Reviews 2015, 115 (4), 1653-1701.
3. Lee, B. R.; Ko, H. K.; Ryu, J. H.; Ahn, K. Y.; Lee, Y. H.; Oh, S. J.; Na, J. H.; Kim, T. W.; Byun, Y.; Kwon, I. C.; Kim, K.; Lee, J., Engineered human ferritin nanoparticles for direct delivery of tumor antigens to lymph node and cancer immunotherapy. Scientific Reports 2016, 6, 35182.
4. He, D.; Marles-Wright, J., Ferritin family proteins and their use in bionanotechnology. Nature Biotechnology 2015, 32 (6), 651-657.

5. Liu, L.; Zhang, C.; Li, Z.; Wang, C.; Bi, J.; Yin, S.; Wang, Q.; Yu, R.; Liu, Y.; Su, Z., Albumin binding domain fusing R/K-X-X-R/K sequence for enhancing tumor delivery of doxorubicin. *Molecular Pharmaceutics* 2017, 14 (11), 3739-3749.
6. Shen, Y.; Li, X.; Dong, D.; Zhang, B.; Xue, Y.; Shang, P., Transferrin receptor 1 in cancer: a new sight for cancer therapy. *American Journal of Cancer Research* 2018, 8 (6), 916.
7. Schlapschy, M.; Binder, U.; Borger, C.; Theobald, I.; Wachinger, K.; Kisling, S.; Haller, D.; Skerra, A., PASylation: a biological alternative to PEGylation for extending the plasma half-life of pharmaceutically active proteins. *Protein Engineering Design & Selection* 2013, 26 (8), 489-501.
8. Zhang, X.; Yao, S.; Liu, C.; Jiang, Y., Tumor tropic delivery of doxorubicin-polymer conjugates using mesenchymal stem cells for glioma therapy. *Biomaterials* 2015, 39, 269-281.
9. Aggarwal, N.; Sloane, B. F., Cathepsin B: multiple roles in cancer. *Proteomics Clinical Applications* 2014, 8 (5-6), 427-437.
10. Olson, E. S.; Jiang, T.; Aguilera, T. A.; Nguyen, Q. T.; Ellies, L. G.; Scadeng, M.; Tsien, R. Y., Activatable cell penetrating peptides linked to nanoparticles as dual probes for in vivo fluorescence and MR imaging of proteases. *Proceedings of Natational Academy of Sciences of the United States of America* 2010, 107 (9), 4311-4316.
11. Roth, L.; Agemy, L.; Kotamraju, V. R.; Braun, G.; Teesalu, T.; Sugahara, K. N.; Hamzah, J.; Ruoslahti, E., Transtumoral targeting enabled by a novel neuropilin-binding peptide. *Oncogene* 2012, 31 (33), 3754-3763.
12. He, Z.; Tessier-Lavigne, M., Neuropilin is a receptor for the axonal chemorepellent semaphorin III. *Cell* 1997, 90, 739-751.
13. Cai, Y.; Cao, C.; He, X.; Yang, C.; Tian, L.; Zhu, R.; Pan, Y., Enhanced magnetic resonance imaging and staining of cancer cells using ferrimagnetic H-ferritin nanoparticles with increasing core size. *International Journal of Nanomedicine* 2015, 10, 2619.
14. Tian, Y.; Li, S.; Song, J.; Ji, T.; Zhu, M.; Anderson, G. J.; Wei, J.; Nie, G., A doxorubicin delivery platform using engineered natural membrane vesicle exosomes for targeted tumor therapy. *Biomaterials* 2014, 35 (7), 2383-2390.
15. Gao, F.; Zhang, J.; Fu, C.; Xie, X.; Peng, F.; You, J.; Tang, H.; Wang, Z.; Li, P.; Chen, J., iRGD-modified lipid-polymer hybrid nanoparticles loaded with isoliquiritigenin to enhance anti-breast cancer effect and tumor-targeting ability. *International Journal of Nanomedicine* 2017, 12, 4147.
16. Wang, K.; Zhang, X.; Liu, Y.; Liu, C.; Jiang, B.; Jiang, Y., Tumor penetrability and anti-angiogenesis using iRGD-mediated delivery of doxorubicin-polymer conjugates. *Biomaterials* 2014, 35 (30), 8735-8747.
17. Falvo, E.; Tremante, E.; Arcovito, A.; Papi, M.; Elad, N.; Boffi, A.; Morea, V.; Conti, G.; Toffoli, G.; Fracasso, G.; Giacomini, P.; Ceci, P., Improved doxorubicin encapsulation and pharmacokinetics of ferritin-fusion protein nanocarriers bearing proline, serine, and alanine elements. *Biomacromolecules* 2016, 17 (2), 514-522.
18. Kettler, K.; Veltman, K.; van de Meent, D.; van Wezel, A.; Hendriks, A. J., Cellular uptake of nanoparticles as determined by particle properties, experimental conditions, and cell type. *Environmental Toxicology and Chemistry* 2014, 33 (3), 481-492.
19. Zhong, Y. J.; Shao, L. H.; Li, Y., Cathepsin B-cleavable doxorubicin prodrugs for targeted cancer therapy (Review). *International Journal of Oncology* 2013, 42 (2), 373-383.

20. Mei, L.; Zhang, Q.; Yang, Y.; He, Q.; Gao, H., Angiopep-2 and activatable cell penetrating peptide dual modified nanoparticles for enhanced tumor targeting and penetrating. *International Journal of Pharmaceutics* 2014, 474 (1-2), 95-102.
21. Gao, H.; Zhang, S.; Cao, S.; Yang, Z.; Pang, Z.; Jiang, X., Angiopep-2 and activatable cell-penetrating peptide dual-functionalized nanoparticles for systemic glioma-targeting delivery. *Molecular Pharmaceutics* 2014, 11 (8), 2755-2763.
22. Feron, O., Tumor-penetrating peptides: A from magic bullets to magic guns. *Science Translational Medicine* 2010, 2 (34), 34ps26-34ps26.

Chapter 7

CONCLUSION AND FUTURE RESEARCH

7.1 Conclusions

Table 7.1 lists a brief summary of procedure and results in this thesis.

Key conclusions from this thesis work include:

1. Three PAS functionalised HFns were expressed successfully as self-assemblies in *Escherichia coli* (*E. coli*). By contrast, pro-apoptotic peptide (P) has negatively affected HFn self-assembling because of its high density of positive charges, its rigid helix structure and its hydrophobicity. In the design of functionalised HFn, the avoidance of functional peptide with high positive charge and rigid structure is necessary.
2. A two-step purification process has been developed for recombinant HFn and PAS functionalised HFns to remove host cell proteins (HCPs) and nucleic acids. Functional peptide has impacted HFn conformation, leading to difference in performance. PAS peptide insertion has caused an E-helix turnover of HFn, resulting in a significant stability decrease in 3 functionalised HFns. Therefore, purification performance differed from HFn.
3. The thermally induced passive diffusion approach applies to both HFn and PAS functionalised HFns. It can be a desirable alternative for loading other small molecule drugs to HFn and modified HFns. Molecular docking combined with molecular dynamic (MD) simulation has shed light on the drug loading mechanism by study of the interactions between drug and HFn.
4. Both PAS and RGDK peptide have improved HFn anti-tumour drug delivery performance. PAS has extended HFn half-life in circulation through its hydration ability to enlarge HFn hydrodynamic volume, RGDK has significantly improved tumour targeted delivery and anti-tumour growth efficacy. Two dually functionalised HFns, HFn-GFLG-PAS-RGDK and HFn-PLGLAG-PAS-RGDK hold greatest practical potential as anti-tumour drug carriers.

Table 7.1 Experiment procedure and brief result summary of all HFn-based proteins in this thesis.

HFn-based protein	Aim	Expression conformation	Purification pathway	Drug loading study	Anti-tumour drug delivery performance
HFn	Control for both functionalisation strategies	Self-assembly and IBs	Heat-acid precipitation +HIC	1. Thermally induced passive loading 2. Condition optimisation 3. Computation analysis of drug loading mechanism	Significantly better anti-tumour performance than free DOX <i>in vivo</i>
HFn-PAS	PAS Functionalised strategy, DOX loading	Self-assembly and IBs	Heat-acid precipitation +Mono Q IEC	1. Thermally induced passive loading	Significantly better anti-tumour performance than DOX loaded HFn and DOX <i>in vivo</i>
HFn-GFLG-PAS-RGDK	PAS Functionalised strategy, DOX loading	Self-assembly and IBs	Heat-acid precipitation + Mono Q IEC	1. Thermally induced passive loading 2. Condition optimisation	Significantly better anti-tumour performance than all the other groups <i>in vivo</i>
HFn-PLGLAG-PAS-RGDK	PAS Functionalised strategy, DOX loading	Self-assembly and IBs	Heat-acid precipitation + Mono Q IEC	1. Thermally induced passive loading	Significantly better anti-tumour performance than all the other groups <i>in vivo</i>
P-HFn	Peptide drug P functionalised strategy	IBs, computation analysis	/	/	/
sHFn-P	Peptide drug P functionalised strategy	Monomer and IBs, computational analysis	/	/	/
sHFn-P-RGDK	Peptide drug P functionalised strategy	Monomer and IBs	/	/	/
sHFn	Control for Peptide drug P functionalised strategy	Self-assembly and IBs	/	/	/

IBs: inclusion bodies; HIC: hydrophobic interaction chromatography; IEC: ion exchange chromatography; DOX: doxorubicin hydrochloride.

7.2 Future research directions

Future research directions will necessarily include:

1. Although pro-apoptotic peptide has a significantly negative impact on HF_n, different length of linkers and perhaps the addition of other functional peptide are worth trying to form assemblies with pro-apoptotic peptide. Hybridisation of sHF_n-P with HF_n by disassembly/re-assembly has been tried but unsuccessful. However, hybridisation by co-expression or the change of expression host to mammalian cells can be alternative means to obtain assemblies with pro-apoptotic peptide.
2. Molecular docking combined with MD simulation analysis can be applied to screen suitable drugs to be loaded to HF_n. A database of small molecule anti-tumour drugs (molecular weight < 600 Da) can be employed as ligand in a high throughput screening. The drugs that can form a dominant and strongly interacted complex with HF_n inner surface would be the most likely ones to achieve a desirable loading performance.
3. MD simulation can be used to assess the impact of additives and stressors (temperature, pH, pressure) on HF_n drug loading by analysing the stability of HF_n subunit-drug complex obtained from molecule docking screening. The result can potentially guide experimental optimisation of drug loading.
4. Protein-peptide molecular docking combined with MD simulation can be used to assess functional peptide impact on HF_n, and screen important factors in design of functionalised HF_n, such as the linker length, functional peptide type, insertion site.
5. Ferritin in different species have distinct properties from HF_n and are worth exploration as drug delivery platform. A combination of advantages of ferritin from other species with HF_n can be achieved by gene technology or hybridisation.

Appendix A-- Supporting tables

Table A1 lists the primary sequence of all designed HFn-based proteins in **Chapter 3**.

Table A1 Primary sequences of 8 HFn-based protein subunits.

Protein	Subunit primary sequence and molecular weight
sHFn	TTAST SQVRQ NYHQD SEAAI NRQIN LELYA SYVYL SMSYY FDRDD VALKN FAKYF LHQSH EEREH AEKLM KLQSQ RGGRI FLQDI KKPDC DDWES GLNAMECALH LEKNV NQSLLELHKL ATDKN DPHLC DFIET HYLNE QVKAI KELGD HVTNL RKMGA (18.69 kDa)
sHFn-P	TTAST SQVRQ NYHQD SEAAI NRQIN LELYA SYVYL SMSYY FDRDD VALKN FAKYF LHQSH EEREH AEKLM KLQSQ RGGRI FLQDI KKPDC DDWES GLNAMECALH LEKNV NQSLLELHKL ATDKN DPHLC DFIET HYLNE QVKAI KELGD HVTNL RKMGA GGSGGGGGSGGGSGG KLAKLAKKLAKLAK (21.14 kDa)
sHFn-P-RGDK	TTAST SQVRQ NYHQD SEAAI NRQIN LELYA SYVYL SMSYY FDRDD VALKN FAKYF LHQSH EEREH AEKLM KLQSQ RGGRI FLQDI KKPDC DDWES GLNAMECALH LEKNV NQSLLELHKL ATDKN DPHLC DFIET HYLNE QVKAI KELGD HVTNL RKMGA GGSGGGGGSGGGSGG KLAKLAKKLAKLAK GGG RGDK (21.77 kDa)
HFn	TTAST SQVRQ NYHQD SEAAI NRQIN LELYA SYVYL SMSYY FDRDD VALKN FAKYF LHQSH EEREH AEKLM KLQSQ RGGRI FLQDI KKPDC DDWES GLNAMECALH LEKNV NQSLLELHKL ATDKN DPHLC DFIET HYLNE QVKAI KELGD HVTNL RKMGA PESGL AEYLF DKHTL GDSDN ES (21.08 kDa)
P-HFn	KLAKLAKKLAKLAK GGSGGGGTGGSGGG GGSGGGGTGGSGGG TTAST SQVRQ NYHQD SEAAI NRQIN LELYA SYVYL SMSYY FDRDD VALKN FAKYF LHQSH EEREH AEKLM KLQSQ RGGRI FLQDI KKPDC DDWES GLNAMECALH LEKNV NQSLLELHKL ATDKN DPHLC DFIET HYLNE QVKAI KELGD HVTNL RKMGA PESGL AEYLF DKHTL GDSDN ES (24.52 kDa)
HFn-PAS	TTAST SQVRQ NYHQD SEAAI NRQIN LELYA SYVYL SMSYY FDRDD VALKN FAKYF LHQSH EEREH AEKLM KLQSQ RGGRI FLQDI KKPDC DDWES GLNAMECALH LEKNV NQSLLELHKL ATDKN DPHLC DFIET HYLNE QVKAI KELGD HVTNL RKMGA PESGL AEYLF DKHTL GDSDN ES GGSGG GGTGG GSGGG GFLG ASPAA PAPASPAAPASAPAASPAAPAPASPAAPASAPA GGSGG (26.05 kDa)

<p>HFn-GFLG-PAS- RGDK</p>	<p>TTAST SQVRQ NYHQD SEAAI NRQIN LELYA SYVYL SMSYY FDRDD VALKN FAKYF LHQSH EEREH AEKLM KLQNQ RGGRI FLQDI KKPDC DDWES GLNAMECALH LEKNV NQSL ELHKL ATDKN DPHLC DFIET HYLNE QVKAI KELGD HVTNL RKMGA PESGL AEYLF DKHTL GDSDN ES GGGSG GGTGG GSGGG GFLG <u>ASPAA PAPAS</u> <u>PAAPAPSAPAASPAAPAPASPAAPAPSAPA</u> GGSGG RGDK (26.50 kDa)</p>
<p>HFn-PLGLAG- <u>PAS-RGDK</u></p>	<p>TTAST SQVRQ NYHQD SEAAI NRQIN LELYA SYVYL SMSYY FDRDD VALKN FAKYF LHQSH EEREH AEKLM KLQNQ RGGRI FLQDI KKPDC DDWES GLNAMECALH LEKNV NQSL ELHKL ATDKN DPHLC DFIET HYLNE QVKAI KELGD HVTNL RKMGA PESGL AEYLF DKHTL GDSDN ES GGGSG GGTGG GSGGG PLGLAG <u>ASPAA PAPAS</u> <u>PAAPAPSAPAASPAAPAPASPAAPAPSAPA</u> GGSGG RGDK (26.64 kDa)</p>

Residues in green are **flexible linkers**, residues in bold are **enzyme-cleavable sites**, residues underlined are **PAS peptide**, residues in red are **RGDK**.

Table A2 lists the sequencing results of bacterial colonies picked from agarose plates in heat-shock transformation in **Chapter 3**. It proves that the grown bacterial strains contained pET30a expressing all HFn-based proteins.

Table A2 Sequencing results from **BGI Company (China)**

<p>HFn sense strand sequencing result:</p> <p>GTTGGCATGTGAGCGGAACATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATAT <u>ACATATG</u>ACTACTGCTTCTACCTCTCAAGTTCGTCAAAACTATCATCAGGATTCTGAAGCAGCAA <u>TTAATCGTCAGATTAACCTGGAGCTGTATGCTTCCTACGTTTATCTGTCTATGTCCTACTACTTCG</u> <u>ATCGTGATGATGTAGCGCTGAAAAACTTCGCGAAATATTTTCTGCACCAGTCCCACGAGGAACG</u> <u>CGAACACGCCGAAAACTGATGAAACTGCAGAACCAGCGCGGTGGTCGTATCTTTCTGCAGGA</u> <u>CATCAAGAAACCGGACTGCGACGACTGGGAAAGCGGTCTGAACGCTATGGAATGCGCGCTGCA</u> <u>TCTGGAGAAAAACGTGAACCAGAGCCTGCTGGAAGTGCATAAACTGGCGACCGATAAAAAACGA</u> <u>CCCTCACCTGTGTGACTTCATTGAAACCCACTACCTGAATGAACAGGTAAAGCGATCAAAGAA</u> <u>CTGGGCGACCACGTTACCAACCTGCGTAAAATGGGCGCCCCGGAATCCGGCCTGGCGGAATAC</u> <u>CTGTTGACAAACACACCCTGGGCGATTCTGACAACGAAAGCTAACTCGAG</u>CACCACCACCACC ACC ACTGAGATCCGGCTGCTAACAAAGCCC GAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTG AGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGG AGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGT GTGGTGGTTACGCGCAGCGTGACCG</p>
<p>sHFn sense strand sequencing result:</p> <p>ATTGGCATGTGAGCGGAACATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATAT <u>ACATATG</u>ACTACTGCTTCTACCTCTCAAGTTCGTCAAAACTATCATCAGGATTCTGAAGCAGCAA <u>TTAATCGTCAGATTAACCTGGAGCTGTATGCTTCCTACGTTTATCTGTCTATGTCCTACTACTTCG</u> <u>ATCGTGATGATGTAGCGCTGAAAAACTTCGCGAAATATTTTCTGCACCAGTCCCACGAGGAACG</u> <u>CGAACACGCCGAAAACTGATGAAACTGCAGAACCAGCGCGGTGGTCGTATCTTTCTGCAGGA</u> <u>CATCAAGAAACCGGACTGCGACGACTGGGAAAGCGGTCTGAACGCTATGGAATGCGCGCTGCA</u> <u>TCTGGAGAAAAACGTGAACCAGAGCCTGCTGGAAGTGCATAAACTGGCGACCGATAAAAAACGA</u> <u>CCCTCACCTGTGTGACTTCATTGAAACCCACTACCTGAATGAACAGGTAAAGCGATCAAAGAA</u> <u>CTGGGCGACCACGTTACCAACCTGCGTAAAATGGGCGCCTAACTCGAG</u>CACCACCACCACCACC ACTGAGATCCGGCTGCTAACAAAGCCC GAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGC AATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGG AACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTG GTGGTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCTCCTTTTCGCTTTCTT CCTTCCTT</p>
<p>sHFn-P sense strand sequencing result:</p> <p>GGCCACGGGCCTTGCCACCATTACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGCGGAG CCCGATCTTCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGCGCCGGTG</p>

ATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCGATCTCGATCCCGC GAAATTAATAC
GACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGT TTTAACTTT
AAGAAGGAGATATACATATGACCACCGCAAGTACCAGCCAGGTTCCGCCAGAATTATCATCAGG
ATAGCGAAGCCGCCATTAATCGCCAGATTAATCTGGAAGTGTATGCAAGTTATGTGTATCTGAG
TATGAGCTATTATTTTCGATCGCGATGATGTTGCCCTGAAAAATTTTGCAAAATATTTTCCTGCACC
AGAGCCATGAAGAACGCGAACATGCAGAAAACTGATGAACTGCAGAATCAGCGCGGCGGC
CGCATTTTTCTGCAGGATATTAAGAAACCGGATTGCGATGATTGGGAAAGTGGTCTGAATGCCA
TGGAATGCGCCCTGCATCTGGAAAAGAATGTGAATCAGAGCCTGCTGGAAGTGCATAAACTGGC
CACCGATAAAAATGATCCGCATCTGTGTGATTTTATTGAAACCCATTATCTGAACGAACAGGTT
AAAGCCATTAAGGAACTGGGTGACCATGTGACCAATCTGCGTAAAATGGGCGCAGGCGGTAGT
GGTGGCGGCGGTAGTGGCGGTGGTGGCAGTGGTGGCAAACCTGGCAAACCTGGCCAAAAACTG
GCAAAGCTGGCAAATGACTCGAGCACCACCACCACCACCTGAGATCCGGCTGCTAACAAA
GCCCGAAAGAAGGCGATTTTGGCCA

sHFn-P-RGDK sense strand sequencing result:

CGGGGGTCGGCGTTAGGGGGAGCGTCATTCCCTCTAGAATAATTTTGT TTTAACTTTAAGAAGGA
GATATACATATGACCACCGCAAGTACCAGCCAGGTTCCGCCAGAATTATCATCAGGATAGCGAAG
CCGCCATTAATCGCCAGATTAATCTGGAAGTGTATGCAAGTTATGTGTATCTGAGTATGAGCTAT
TATTTTCGATCGCGATGATGTTGCCCTGAAAAATTTTGCAAAATATTTTCCTGCACCAGAGCCATGA
AGAACGCGAACATGCAGAAAACTGATGAACTGCAGAATCAGCGCGGCGGCCGCATTTTTCT
GCAGGATATTAAGAAACCGGATTGCGATGATTGGGAAAGTGGTCTGAATGCCATGGAATGCGC
CCTGCATCTGGAAAAGAATGTGAATCAGAGCCTGCTGGAAGTGCATAAACTGGCCACCATAA
AAATGATCCGCATCTGTGTGATTTTATTGAAACCCATTATCTGAACGAACAGGTTAAAGCCATT
AAGGAACTGGGTGACCATGTGACCAATCTGCGTAAAATGGGCGCAGGCGGTAGTGGTGGCGGC
GGTAGTGGCGGTGGTGGCAGTGGTGGCAAACCTGGCAAACCTGGCCAAAAACTGGCAAAGCTG
GCAAAGGTGGTGGTCCGCGGTGACAAATAACTCGAGCACCACCACCACCACCTGAGATCCG
GCTGCTAACAAAGCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCA
TAACCCCTTGGGCGCTCTAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCG
GATTGGCGAATGGGACGCGCCCTGTAGCGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGC
AGCGTGACCGCTACACTTGCCAGCGCCCTAAGCGCCGCTCCTTTTCGCTTCTTCCCTTCTTTCTCG
CACGTTCCGCGGCTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTAA

P-HFn sense strand sequencing result:

AATAATTTTGT TTTAACTTTAAGAAGGAGATATACATATGAAGCTGGCAAACCTGGCCAAAAAAC
TGGCAAATTAGCCAAAGGTGGCGGTAGCGGTGGCGGTACCGGTGGTGGTAGCGGTGGTGGCG
GCGGTGGTAGTGGTGGCGGTACAGGCGGCGGTAGTGGTGGTGGCACCACCGCAAGCACCAGCC
AGGTTTCGTCAGAATTATCATCAGGATAGTGAAGCAGCAATTAATCGCCAGATTAATCTGGAAGT
GTATGCCAGTTATGTTTATCTGAGTATGAGCTATTACTTCGATCGTGATGATGTTGCCCTGAAA
ATTTTGCAAAATATTTTCCTGCACCAGAGCCATGAAGAACGCGAACATGCCGAAAACTGATGAA
ACTGCAGAATCAGCGTGGTGGTTCGTATTTTTCTGCAGGATATTAAGAAACCGGATTGTGATGAT

TGGGAAAGTGGCCTGAATGCCATGGAATGCGCACTGCATCTGGAAAAGAATGTTAATCAGAGT
CTGCTGGAAGTGCATAAACTGGCAACCGATAAAAATGATCCGCATCTGTGCGATTTTATTGAAA
CCCATTATCTGAATGAGCAGGTGAAAGCAATTAAGGAACTGGGCGATCATGTTACCAATCTGCG
TAAAATGGGCGCCCCGAAAGCGGCCTGGCAGAATATCTGTTTGATAAACATACCCTGGGCGAT
AGTGATAATGAAAGCTAACTCGAGCACCACCACCACCACCCTGAGATCCGGCTGCTAA

HF_n-PAS sense strand sequencing result:

AATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGACCACCGCAAGCACCAgCCAGGTTC
GTCAGAATTATCATCAGGATAGTGAAGCAGCAATTAATCGCCAGATTAATCTGGAAGTGTATGC
CAGTTATGTTTATCTGAGTATGAGCTTACTTCGATCGTGATGATGTTGCCCTGAAAAATTTTG
CAAAATATTTCTGCACCAGAGCCATGAAGAACGCGAACATGCCGAAAAACTGATGAAACTGC
AGAATCAGCGTGGTGGTCGTATTTTTCTGCAGGATATTAAGAAACCGGATTGTGATGATTGGGA
AAGTGGCCTGAATGCCATGGAATGCGCACTGCATCTGGAAAAGAATGTTAATCAGAGTCTGCTG
GAACTGCATAAACTGGCAACCGATAAAAATGATCCGCATCTGTGCGATTTTATTGAAACCCATT
ATCTGAATGAGCAGGTGAAAGCAATTAAGGAACTGGGCGATCATGTTACCAATCTGCGTAAAAT
GGGCGCCCCGAAAGCGGCCTGGCAGAATATCTGTTTGATAAACATACCCTGGGCGATAGTGAT
AATGAAAGCGGTGGCGGCAGTGGCGGCGGTACCGGCGGTGGTTCTGGCGGTGGTGGCTTTCTGG
GCGCCAGCCCTGCCGCTCCTGCCCTGCAAGTCCGGCAGCACCGGCTCCTAGCGCCCCTGCAGC
AAGTCCGGCCGCTCCTGCTCCGGCAAGCCCTGCAGCTCCGGCACCTAGTGCCCCGGCATAA**CTC**
GAGCACCACCACCACCACCCTGAGATCCGGCT

HF_n-GFLG-PAS-RGDK sense strand sequencing result:

AATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGACCACCGCAAGCACCCAGGTTTC
GTCAGAATTATCATCAGGATAGTGAAGCAGCAATTAATCGCCAGATTAATCTGGAAGTGTATGC
CAGTTATGTTTATCTGAGTATGAGCTTACTTCGATCGTGATGATGTTGCCCTGAAAAATTTTG
CAAAATATTTCTGCACCAGAGCCATGAAGAACGCGAACATGCCGAAAAACTGATGAAACTGC
AGAATCAGCGTGGTGGTCGTATTTTTCTGCAGGATATTAAGAAACCGGATTGTGATGATTGGGA
AAGTGGCCTGAATGCCATGGAATGCGCACTGCATCTGGAAAAGAATGTTAATCAGAGTCTGCTG
GAACTGCATAAACTGGCAACCGATAAAAATGATCCGCATCTGTGCGATTTTATTGAAACCCATT
ATCTGAATGAGCAGGTGAAAGCAATTAAGGAACTGGGCGATCATGTTACCAATCTGCGTAAAAT
GGGCGCCCCGAAAGCGGCCTGGCAGAATATCTGTTTGATAAACATACCCTGGGCGATAGTGAT
AATGAAAGCGGTGGCGGCAGTGGCGGCGGTACCGGCGGTGGTTCTGGCGGTGGTGGCTTTCTGG
GCGCCAGCCCTGCCGCTCCTGCCCTGCAAGTCCGGCAGCACCGGCTCCTAGCGCCCCTGCAGC
AAGTCCGGCCGCTCCTGCTCCGGCAAGCCCTGCAGCTCCGGCACCTAGTGCCCCGGCAGGTGGT
AGCGGCGGTCTGTGGTGACAAA**TAACTCGAG**CACCACCACCACCACCCTGAGATCCGGCT

HF_n-PLGLAG-PAS-RGDK sense strand sequencing result:

AATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGACCACCGCAAGCACCCAGGTTTC
GTCAGAATTATCATCAGGATAGTGAAGCAGCAATTAATCGCCAGATTAATCTGGAAGTGTATGC
CAGTTATGTTTATCTGAGTATGAGCTTACTTCGATCGTGATGATGTTGCCCTGAAAAATTTTG
CAAAATATTTCTGCACCAGAGCCATGAAGAACGCGAACATGCCGAAAAACTGATGAAACTGC
AGAATCAGCGTGGTGGTCGTATTTTTCTGCAGGATATTAAGAAACCGGATTGTGATGATTGGGA

AAGTGGCCTGAATGCCATGGAATGCGCACTGCATCTGGAAAAGAATGTTAATCAGAGTCTGCTG
GAACTGCATAAACTGGCAACCGATAAAAATGATCCGCATCTGTGCGATTTTATTGAAACCCATT
ATCTGAATGAGCAGGTGAAAGCAATTAAGGAACTGGGCGATCATGTTACCAATCTGCGTAAAAT
GGGCGCCCCGAAAGCGGCCTGGCAGAATATCTGTTTGATAAACATAACCCTGGGCGATAGTGAT
AATGAAAGCGGTGGCGGCAGTGGCGGCGGTACCGGCGGTGGTTCTGGCGGTGGTCCGCTGGGT
CTGGCTGGTGCCAGCCCTGCCGCTCCTGCCCTGCAAGTCCGGCAGCACCGGCTCCTAGCGCCC
CTGCAGCAAGTCCGGCCGCTCCTGCTCCGGCAAGCCCTGCAGTCCGGCACCTAGTGCCCCGGC
AGGTGGTAGCGGCGTCTGGTGACAAA**TAA****CTCGAG**CACCACCACCACCACCCTGAGATCC
GGCT

Nucleotides in red are **Nde I and Xho I recognition sites**, codon in bold is the **stop codon** and nucleotides underlined are the codons for target protein residues.

Appendix B – Supporting figures

Figure B1 is the standard curve for Bradford protein concentration determination in **Chapter 3, 4 and 5**.

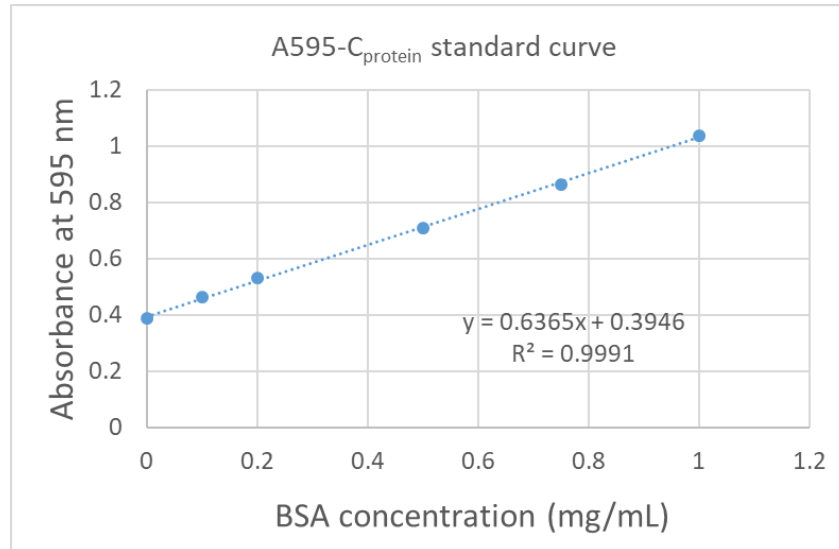


Figure B1 Standard curve for Bradford assay.
 $A_{595} = 0.6365 C_{\text{protein}} + 0.3946$, $R^2 = 0.999$.

Figure B2 is the standard curve for nucleic acid concentration determination in IEC and HIC in **Chapter 4**.

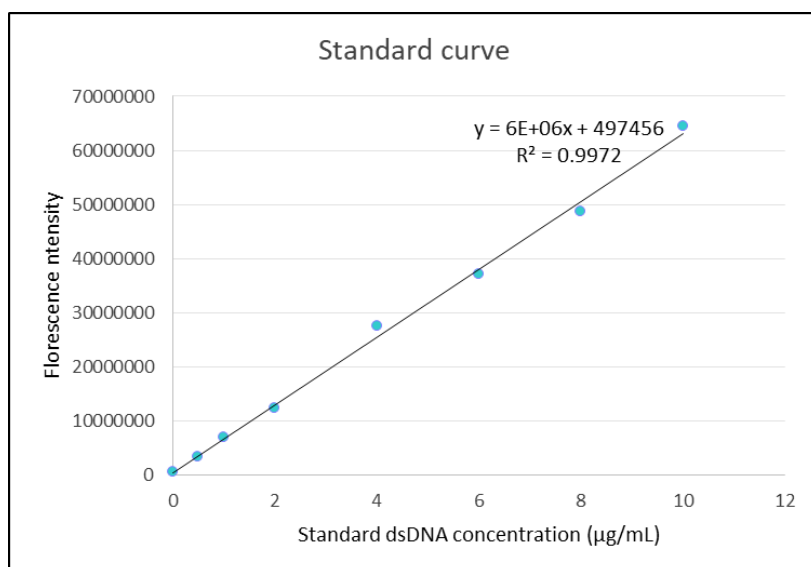


Figure B2 Standard curve for nucleic acid concentration determination.
Intensity = 6,000,000 C_{DNA} + 497456, $R^2 = 0.997$.

Figure B3 is the standard curves for concentration determination of HF_n-based protein and DOX in **Chapter 5**. All the equations obtained from standard curves are for the calculation of loading ratio in thermally induced drug loading.

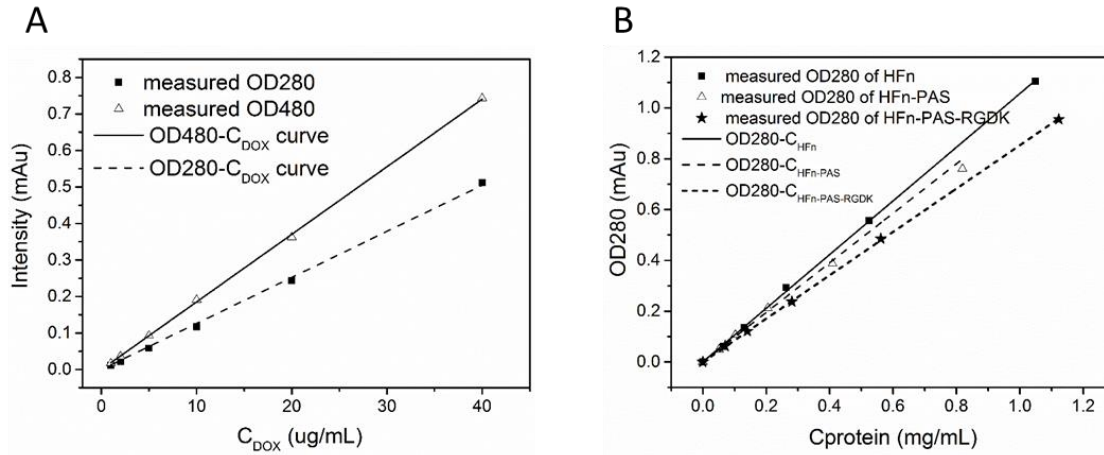


Figure B3 Standard linear curves of correlations between DOX or HF_n-based protein nanocages concentrations and optical densities.

A, OD480- C_{DOX} and OD280- C_{DOX} correlation curves.

$$OD280 = 0.0126 C_{DOX}, R^2 = 0.999,$$

$$OD480 = 0.0185 C_{DOX}, R^2 = 0.999$$

B, OD280- $C_{nanocage}$ correlation curves.

$$HF_n, OD280 = 1.0561 C, R^2 = 0.999,$$

$$HF_n-PAS, OD280 = 0.9258 C, R^2 = 0.998$$

$$HF_n-PAS-RGDK, OD280 = 0.8529 C, R^2 = 0.999$$

Figure B4, B5, and B6 show Superose 6 size-exclusion chromatograms of DOX loaded HFn (HFn/DOX) under all 18 tested conditions in **Chapter 5**. Chromatograms prove that DOX was loaded in HFn.

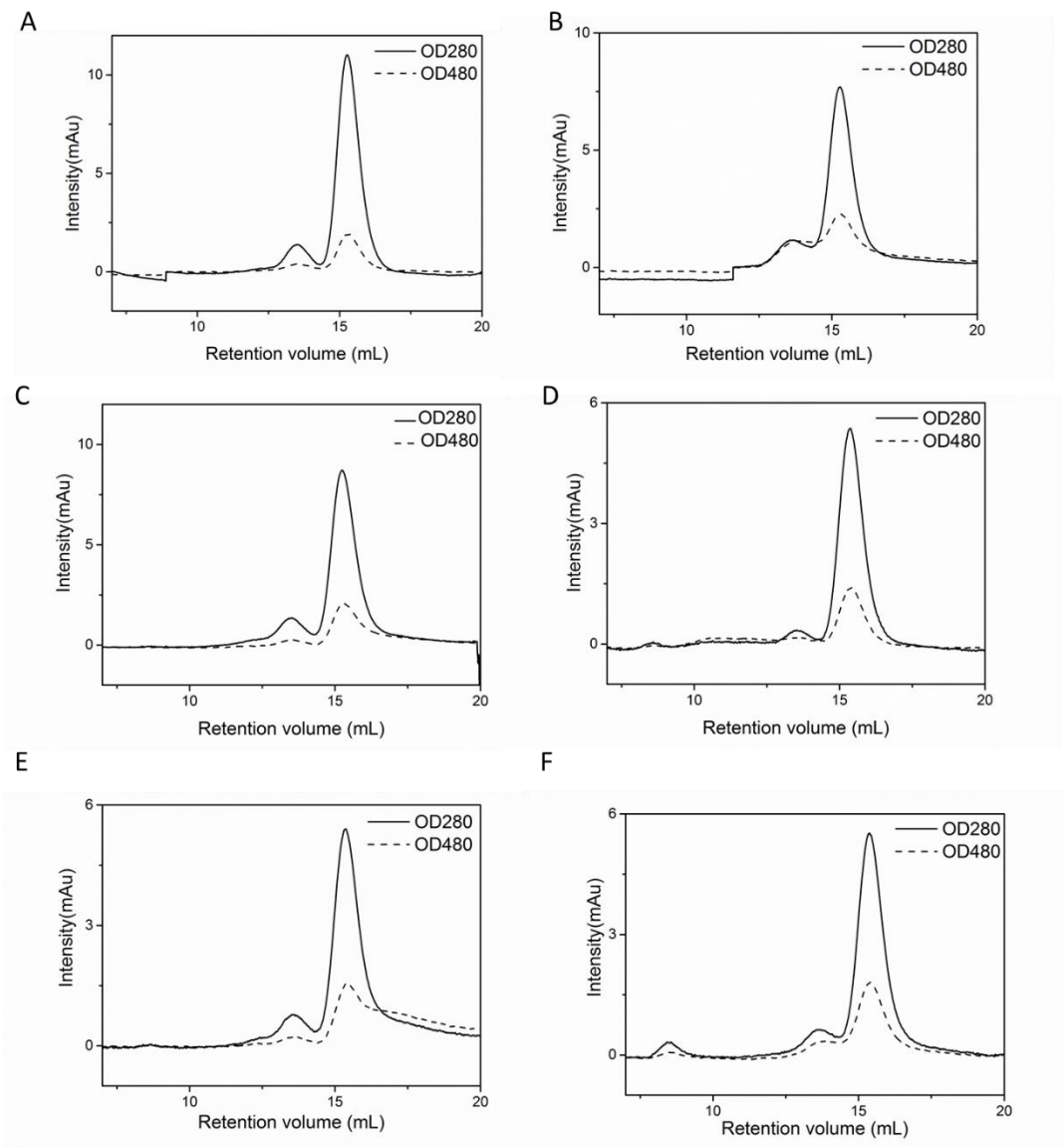


Figure B4 Size-exclusion chromatograms of 45 °C HFn/DOX samples. **A**, pH 7.0, 2 h. **B**, pH 7.5 2 h. **C**, pH 7.0, 4 h. **D**, pH 7.5 4 h. **E**, pH 7.0 6 h. **F**, pH 7.5 6 h.

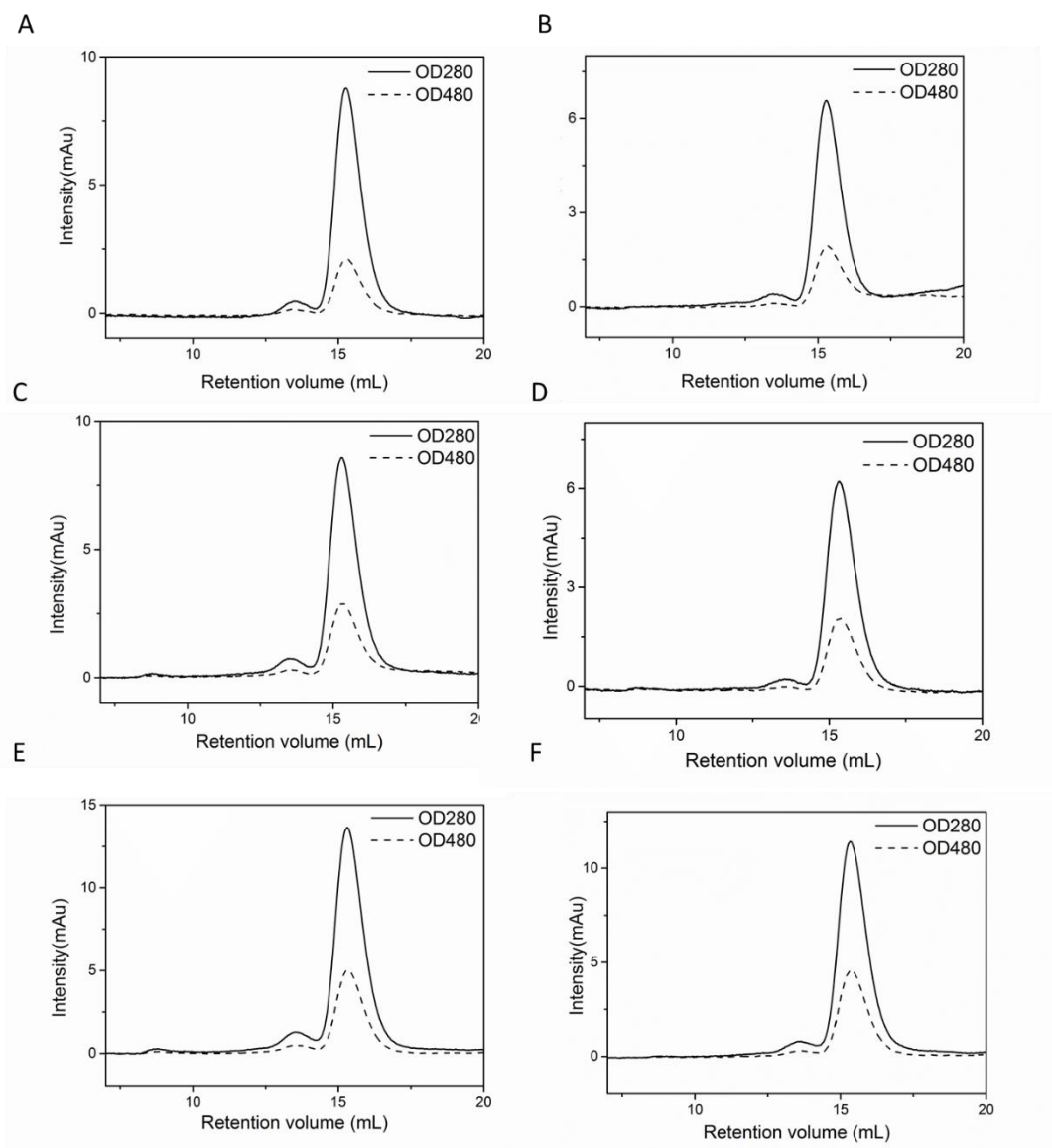


Figure B5 Size-exclusion chromatograms of 50 °C HF_n/DOX samples. **A**, pH 7.0, 2 h. **B**, pH 7.5 2 h. **C**, pH 7.0, 4 h. **D**, pH 7.5 4 h. **E**, pH 7.0 6 h. **F**, pH 7.5 6 h.

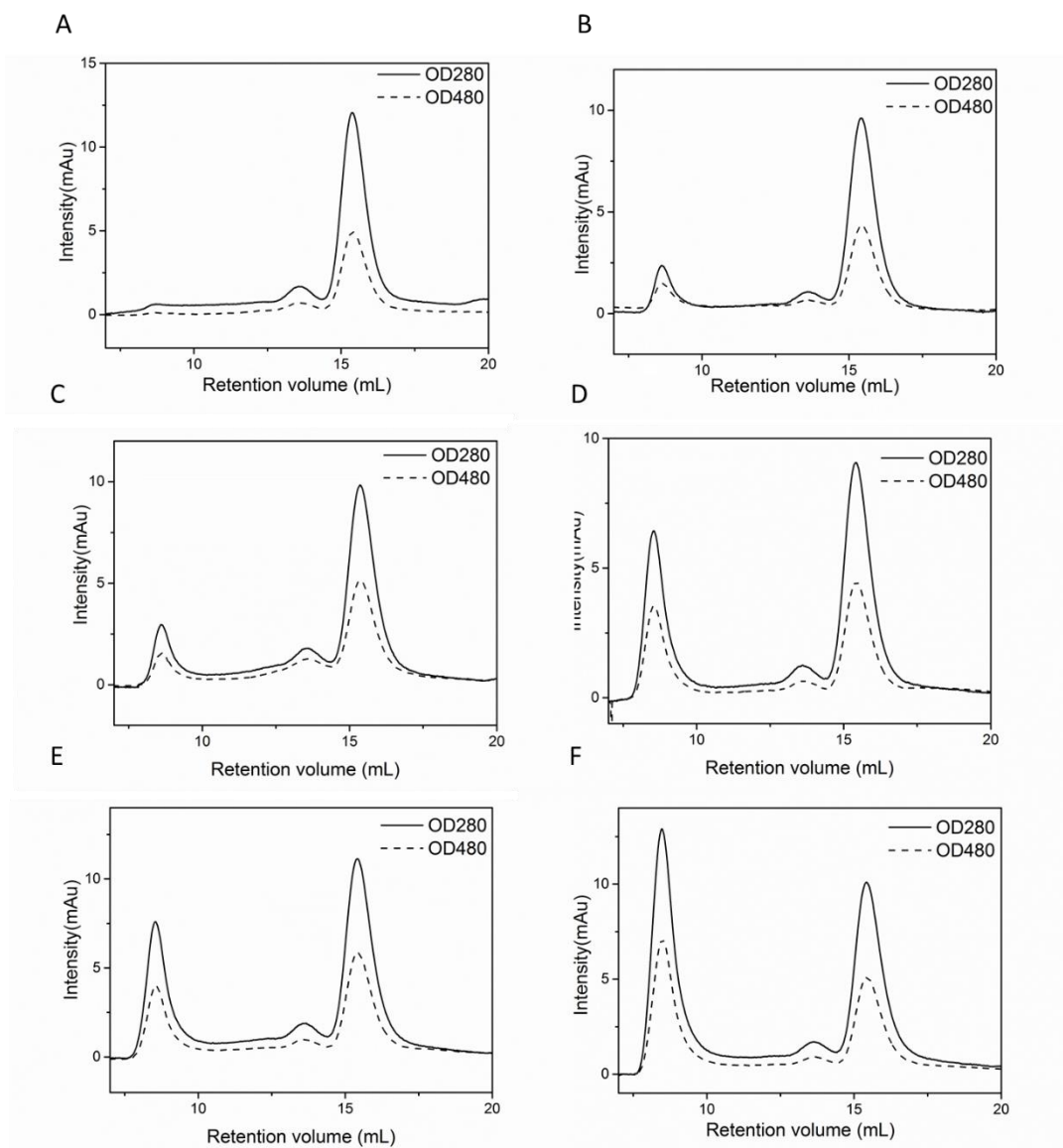


Figure B6 Size-exclusion chromatograms of 60 °C HFn/DOX samples. **A**, pH 7.0, 2 h. **B**, pH 7.5 2 h. **C**, pH 7.0, 4 h. **D**, pH 7.5 4 h. **E**, pH 7.0 6 h. **F**, pH 7.5 6 h.

Figure B7, B8, and B9 show Superose 6 size-exclusion chromatograms of DOX loaded HF_n-GFLG-PAS-RGDK (HF_n-GFLG-PAS-RGDK/DOX) under 18 tested conditions in **Chapter 5**. Chromatograms prove that DOX was loaded in HF_n-GFLG-PAS-RGDK.

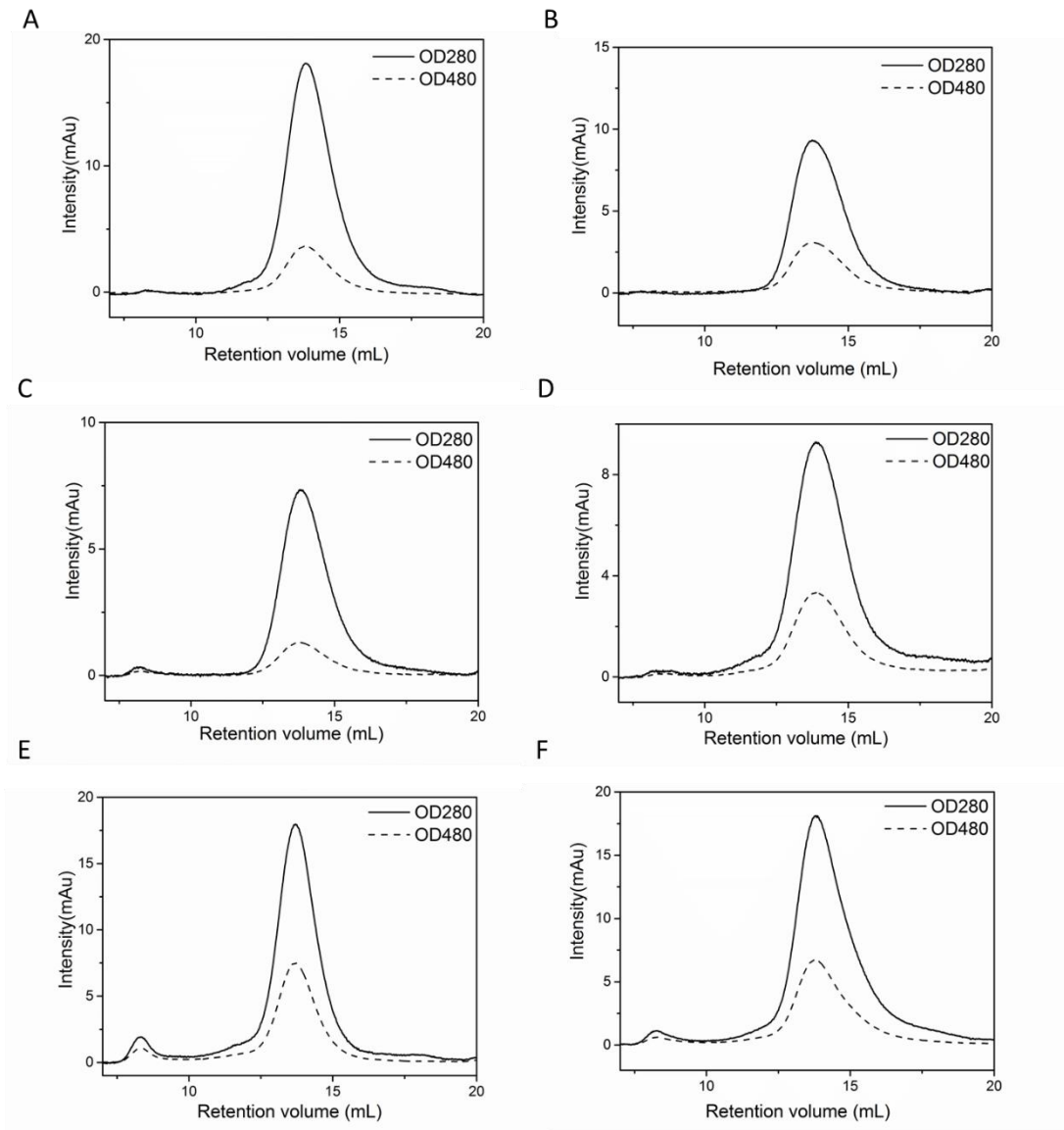


Figure B7 Size-exclusion chromatograms of 45 °C HF_n-GFLG-PAS-RGDK/DOX samples. **A**, pH 7.0, 2 h. **B**, pH 7.5 2 h. **C**, pH 7.0, 4 h. **D**, pH 7.5 4 h. **E**, pH 7.0 6 h. **F**, pH 7.5 6 h.

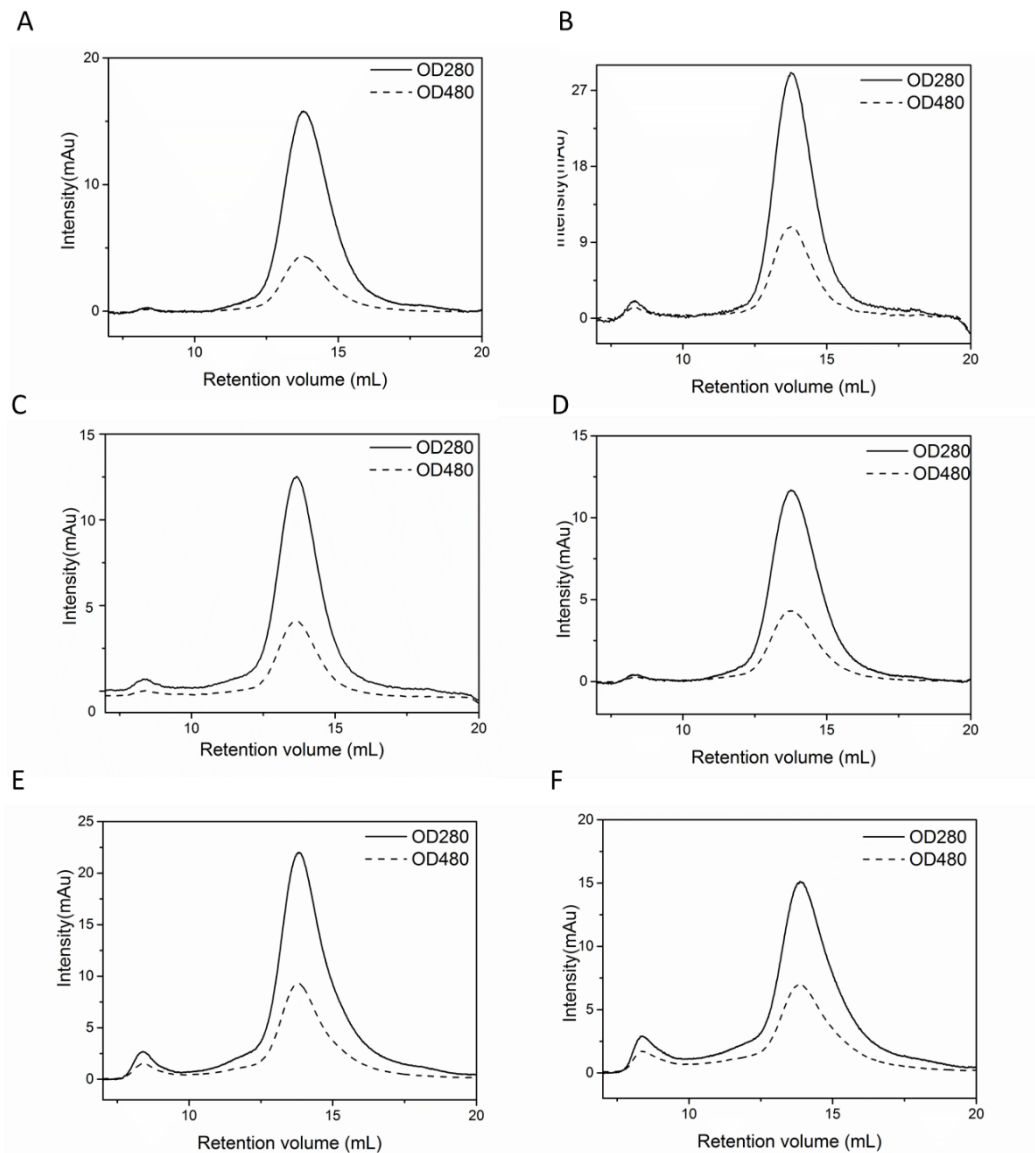


Figure B8 Size-exclusion chromatograms of 50 °C HF_n-GFLG-PAS-RGDK/DOX samples. **A**, pH 7.0, 2 h. **B**, pH 7.5 2 h. **C**, pH 7.0, 4 h. **D**, pH 7.5 4 h. **E**, pH 7.0 6 h. **F**, pH 7.5 6 h.

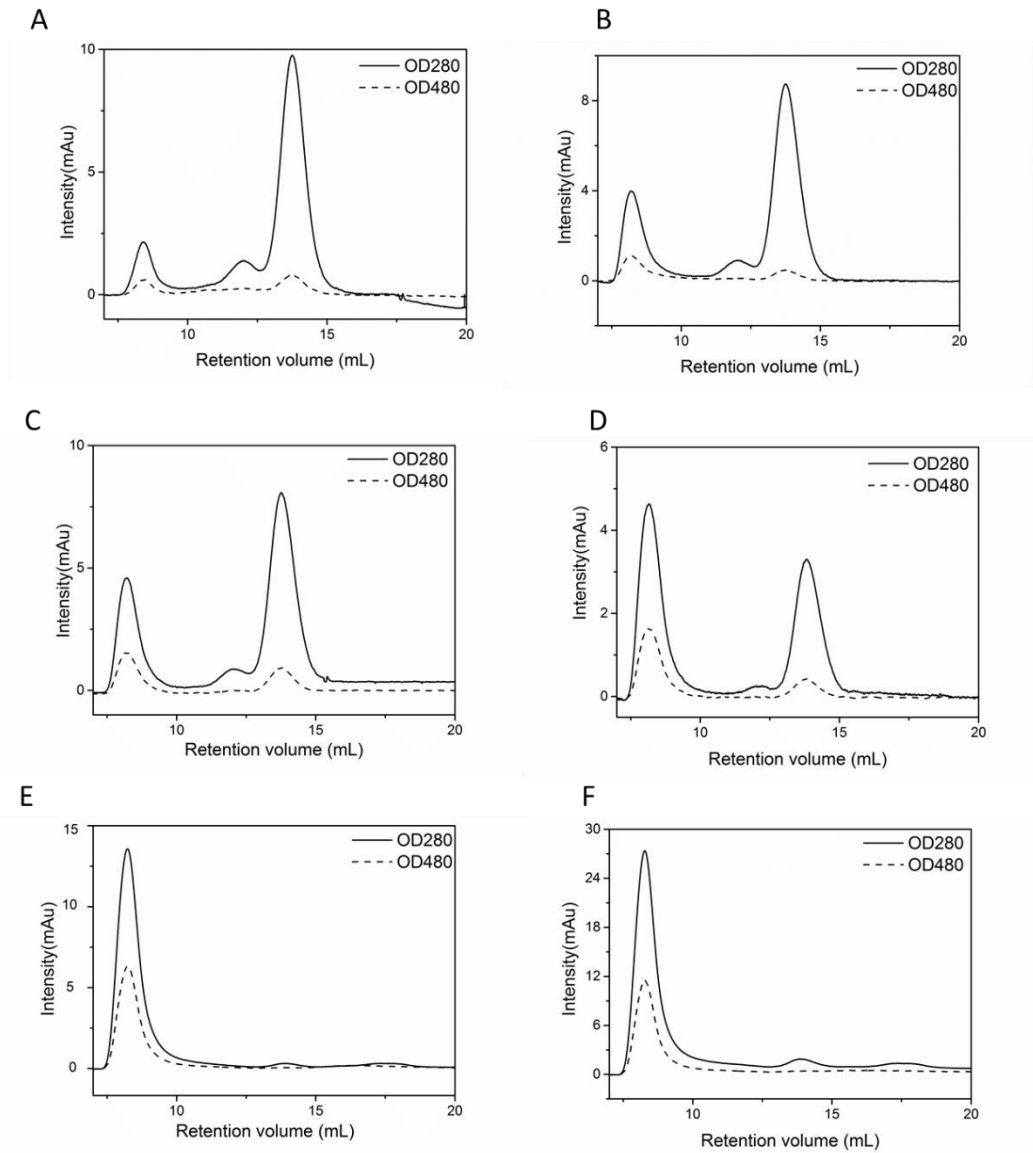
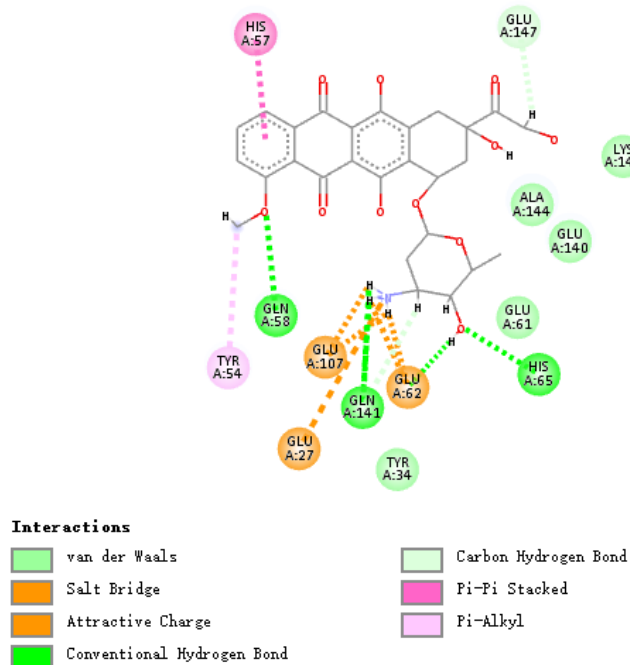


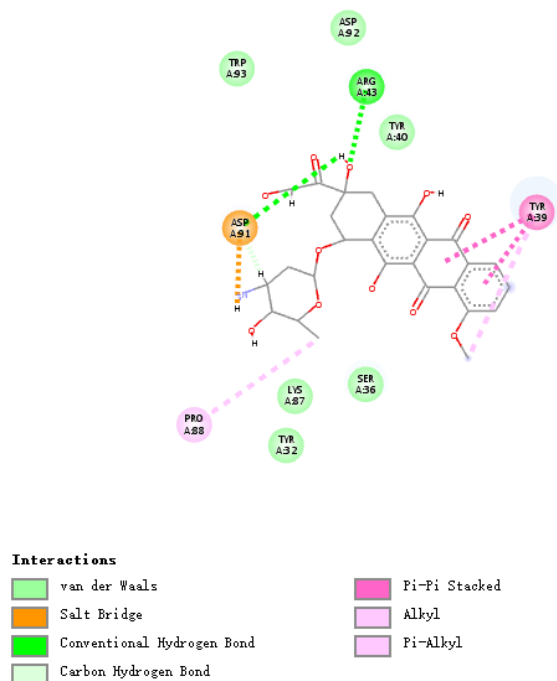
Figure B9 Size-exclusion chromatograms of 60 °C HFn-GFLG-PAS-RGDK/DOX samples. **A**, pH 7.0, 2 h. **B**, pH 7.5 2 h. **C**, pH 7.0, 4 h. **D**, pH 7.5 4 h. **E**, pH 7.0 6 h. **F**, pH 7.5 6 h.

Figure B10 is the 2D diagram of 9 complexes of HFn subunit and DOX after 10 ns MD simulation in computational analysis in **Chapter 5**. The Hydrogen bond, salt bridge and Pi effect interactions are illustrated. Complexes are in the order of stability from high to low.

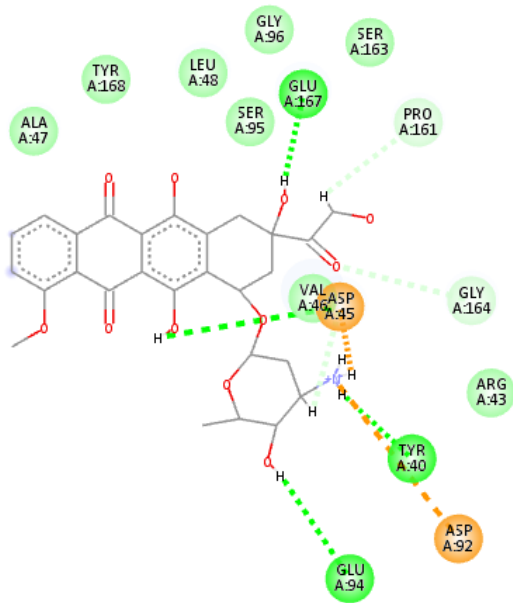
1) Complex 1



2) Complex 4



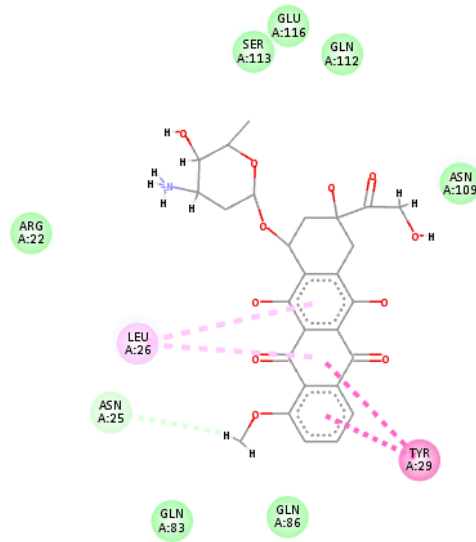
3) Complex 5



Interactions

- | | |
|-------------------|----------------------------|
| van der Waals | Conventional Hydrogen Bond |
| Salt Bridge | Carbon Hydrogen Bond |
| Attractive Charge | |

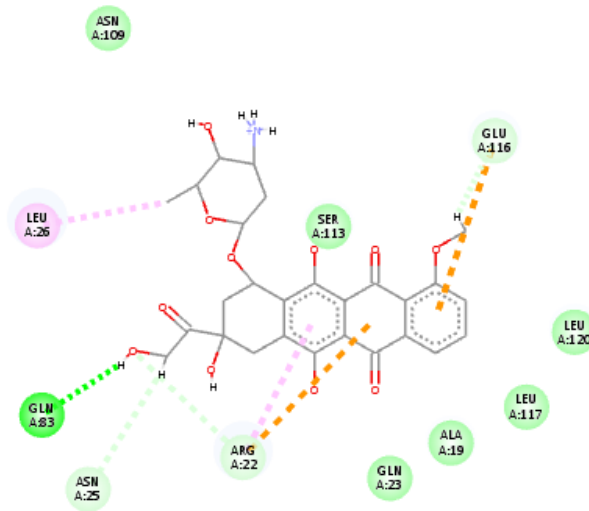
4) Complex 3



Interactions

- | | |
|----------------------|----------------|
| van der Waals | Pi-Pi T-shaped |
| Carbon Hydrogen Bond | Pi-Alkyl |

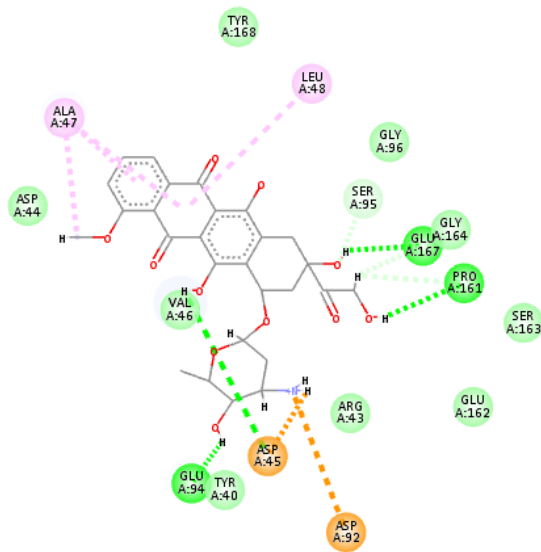
5) Complex 2



Interactions

- | | |
|----------------------------|----------|
| van der Waals | Pi-Anion |
| Conventional Hydrogen Bond | Alkyl |
| Carbon Hydrogen Bond | Pi-Alkyl |
| Pi-Cation | |

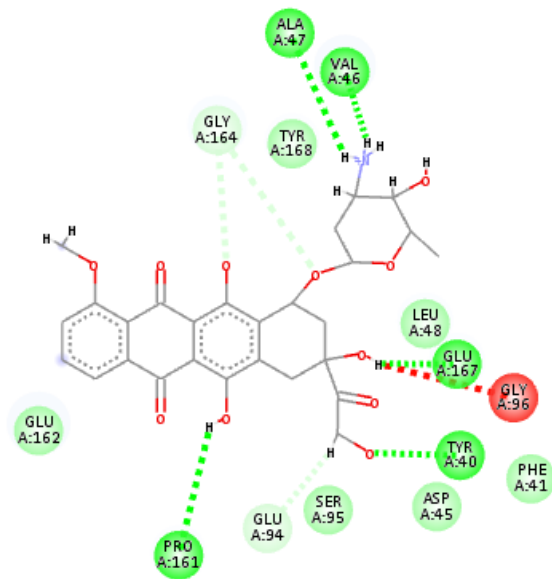
6) Complex 6



Interactions

- | | |
|----------------------------|----------------------|
| van der Waals | Carbon Hydrogen Bond |
| Salt Bridge | Alkyl |
| Attractive Charge | Pi-Alkyl |
| Conventional Hydrogen Bond | |

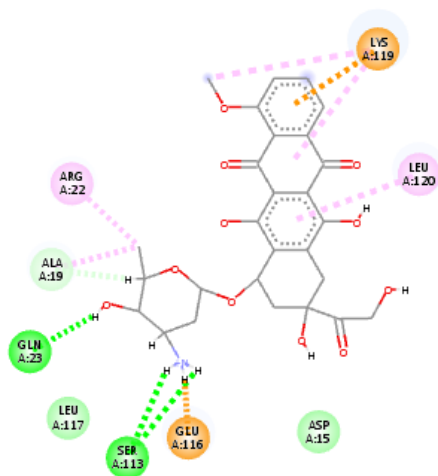
7) Complex 7










Interactions

- | | |
|---|--|
|  van der Waals |  Carbon Hydrogen Bond |
|  Conventional Hydrogen Bond |  Unfavorable Donor-Donor |

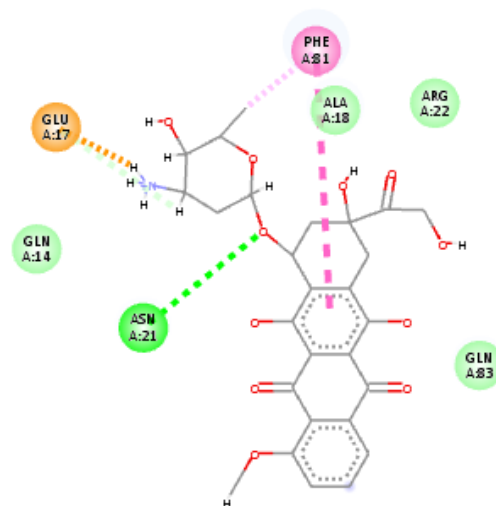
8) Complex 8



Interactions

- | | |
|--|---|
|  van der Waals |  Pi-Cation |
|  Salt Bridge |  Alkyl |
|  Conventional Hydrogen Bond |  Pi-Alkyl |
|  Carbon Hydrogen Bond | |

9) Complex 9



Interactions







	van der Waals		Carbon Hydrogen Bond
	Salt Bridge		Pi-Pi Stacked
	Conventional Hydrogen Bond		Pi-Alkyl

Figure B10 2D diagrams of 9 complexes of HFn subunit with DOX after 10 ns MD simulation.

Figure B11 is the standard curve of fluorescence intensity-for determination of DOX concentration in SD rat plasma in pharmacokinetic study in **Chapter 6**.

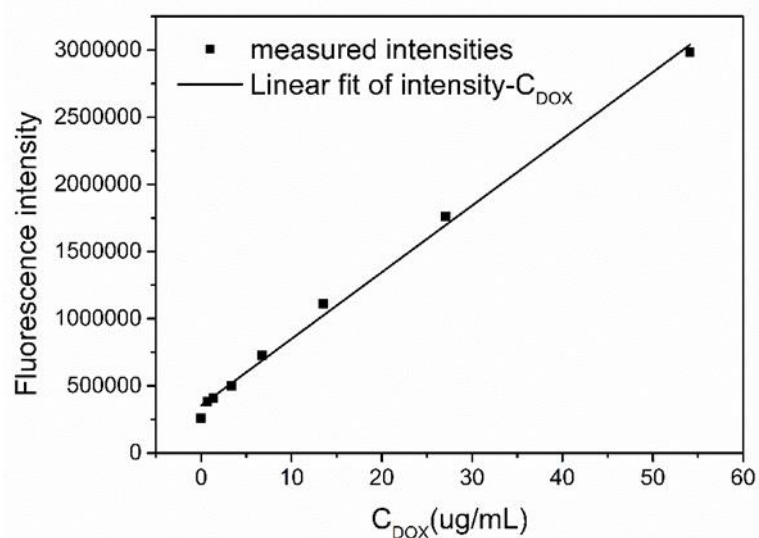


Figure B11 Standard curve of fluorescence intensity-doxorubicin concentration in SD rat plasma. Fluorescence intensity = 49656 C_{DOX} + 353005, R² = 0.996.

PUBLICATIONS DURING CANDIDATURE

Published

1. **Yin, S.**, Wang, Y., Zhang, B., Qu, Y., Liu, Y., Dai, S., Zhang, Y., Wang, Y. and Bi, J. Engineered ferritin with half-life extension and tumor targeting by PAS and RGDK peptide functionalization. *Pharmaceutics* 2021, 13(4), 521.
2. **Yin, S.**, Davey, K., Dai, S., Liu, Y. and Bi, J. A critical review of ferritin as a drug nanocarrier: structure, properties, comparative advantages and challenges. *Particuology* 2021. (Accepted).
3. **Yin, S.**, Zhang, B., Liu, Y., Su, Z., Bi, J. 2021. Development of purification process for dual-function recombinant human heavy-chain ferritin by the investigation of genetic modification impact on conformation. *Engineering in Life sciences*. (Accepted).
4. Zhang, B., **Yin, S.**, Su, Z., Bi, J. Cost-effective purification process development for chimeric Hepatitis B core (HBc) virus-like particles assisted by molecular dynamic simulation. *Engineering in Life Sciences* 2021, 21(6), 438-452.
5. Qu, Y., Wang, L., **Yin, S.**, Zhang, B., Jiao, Y., Sun, Y., Middelberg, A., Bi, J. Stability of Engineered Ferritin Nanovaccines Investigated by Combined Molecular Simulation and Experiments. *The Journal of Physical Chemistry, part B, Biophysics, Biomaterials, Liquids, Soft Matter* 2021, 125(15), 3830–3842.
6. Zhang, Y.; Liu, Y.; Zhang, B.; **Yin, S.**; Li, X.; Zhao, D.; Wang, W.; Bi, J.; Su, Z., In vitro preparation of uniform and nucleic acid free hepatitis B core particles through an optimized disassembly-purification-reassembly process, *Protein Expression and Purification* 2021, 178, 105747.
7. Zhang, Y.; **Yin, S.**; Zhang, B.; Bi, J.; Liu, Y.; Su, Z., HBc-based virus-like particle assembly from inclusion bodies using 2-methyl-2, 4-pentanediol, *Process Biochemistry* 2020, 89, 233-237.
8. Wang, C.; Zhang, C.; Li, Z.; **Yin, S.**; Wang, Q.; Guo, F.; Zhang, Y.; Yu, R.; Liu, Y.; Su, Z., Extending half life of H-ferritin nanoparticle by fusing albumin binding domain for doxorubicin encapsulation, *Biomacromolecules* 2018, 19(3), 773-781.
9. Liu, L.; Zhang, C.; Li, Z.; Wang, C.; Bi, J.; **Yin, S.**; Wang, Q.; Yu, R.; Liu, Y.; Su, Z., Albumin binding domain fusing R/KXXR/K sequence for enhancing tumor delivery of doxorubicin. *Molecular pharmaceutics* 2017, 14(11), 3739-3749.
10. **Yin, S.**; Zhang, C.; Li, Z.; Wang, Q.; Shi, H.; Yu, R.; Liu, Y.; Su, Z., Identification, characterization, and stabilization of the deamidation degradation of recombinant human tumor necrosis factor- α . *Process Biochemistry* 2017, 53, 216-223.

In progress

1. **Yin, S.**; Liu, Y.; Zhang, B.; Zhang, Y.; Qu, Y.; Choe, W.; Dai, S.; Bi, J., Thermally induced drug loading of engineered human heavy-chain ferritin and mechanism study by computational analysis.
2. **Yin, S.**; Liu, Y.; Dai, S.; Bi, J., Investigation of pro-apoptotic peptide impact on human heavy-chain ferritin assembling using computational analysis.
3. Qu, Y.; Zhang, B.; Wang, Y.; **Yin, S.**; Pederick, J. L.; Bruning, J.; Sun, Y.; Middelberg, A., Immunogenicity study of engineered ferritins with C- and N- terminus insertion of Epstein-Barr nuclear antigen 1 epitope. *Vaccine* 2021, submitted.
4. Qu, Y.; Zhang, B.; Yingli, Wang.; **Yin, S.**; Sun, Y.; Middelberg, A.; Bi, J., Immunogenicity and vaccine efficacy boosted by engineering human heavy chain ferritin and chimeric hepatitis B virus core nanoparticles. *ACS Applied Bio Materials* 2021, submitted.