THE UNIVERSITY of ADELAIDE

Stability Study of Engineered Ferritin Nano-vaccines by Combined Molecular simulation and Experiments

Yang Yang

School of Chemical Engineering

The University of Adelaide

A thesis submitted for the examination for the degree of

Master of Philosophy

November 2022

STATEMENT OF 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, Yang Yang, 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 acknowledge that the copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I permit 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.

Signature: Yang Yang

Date: 15/11/2022

EXECUTIVE SUMMARY

Ferritins are a type of protein that is present in almost all living things. They have a symmetric spherical shape with strong physical and chemical stabilities. Human heavy-chain ferritins (HFns) have an inner cavity of 12 nm and 24 identical subunits. Ferritins are known to be assembled and stabilised by hydrogen bonds, salt bridges, and hydrophobic interactions at 2-3-4 symmetry. Previous research has indicated that it is possible to create vaccines using human heavy-chain ferritin with a portion of EBNA1 inserted at the C-terminus.

Four mutant ferritins with short C-terminal amino acid mutations and the letters C1 through C4 are used in this project. C1 was neutrally charged (Gln 167 and 171), C2 was positively charged (Arg 167 and 171), and C3 and C4 were hydrophobic (Val 165 and 169, Ile 166,168, and 170) and hydrophilic (Gln 165 and 169, Asn 166,168, and 170). The changes were done in order to comprehend the relationship between the general stability of the structures and E-helices mutations. The majority of past research has focused on the mutations and stability changes of ferritin. To study the effects of the mutations, this project, however, analysed the ferritin carrier, linker, and antigen as a unified system. As a result, this study is more useful for demonstrating the safety of the developed vaccine.

In this project, experiments and MD simulations on Gromacs were used to examine the hydrophobic, thermal, and pH stability of four genetically altered human heavy-chain ferritins (C1, C2, C3, and C4). By working on this project, we can determine how the mutation will impact the HFns' stability performance. Additionally, to employ simulation in the project to examine the viability of using simulation for vaccine development research.

E. coli was used to express the four mutant HFns. On the basis of recently released data, the purification process for the 4 mutants was examined and improved. C1 was satisfactorily purified by heat-acid precipitation, followed by HIC utilising a prepacked butyl FF column to

remove the nucleic acids. C2, C3, and C4 each contained an extra IEC stage with a Sartobind Q column to remove the remaining HCPs. The final SEC flow through samples of C1 and C2 had purity levels of over 90% and were suitable for characterisation tests. Unexpectedly, C3 and C4 were unable to be purified to an extremely high purity throughout this process.

Following purification, the self-assembly of these proteins was verified using SEC-MALS, and the purity and recovery rate were measured using the Bradford assay and SDS-PAGE. A prepacked Butyl FF column was utilised to compare the hydrophobicity of the proteins, and fluorescence was employed to evaluate the changes in tertiary structure over a volume of temperature and pH. Additionally, Gromacs was used to conduct MD simulations at varied pH and temperature levels. For analysis, it was necessary to calculate the solvent accessible surface area (SASA), root mean square deviation (RMSD), radius of gyration (Rg), root mean square of fluctuation (RMSF), and the distance between subunits.

While C1 and C2 were correctly assembled into a symmetric spherical shape, according to SEC-MALS and MD simulation, they had differing hydrodynamic diameters. C1 is extremely close to the unmutated HFn (F1L3E1), while C2 is larger. The results demonstrate that these helices E mutations would not prevent ferritin from self-assembling. However, the positively charged amino acids of C2 strengthen the electrostatic interaction at the C-terminus, which may result in the larger structure of C2.

Based on experiments, C2 had a higher hydrophobicity, while C1 was similar to F1L3E1. And the computed SASA confirmed these findings, since C2 had a larger SASA than C1 and F1L3E1, and a larger SASA theoretically implied a higher hydrophobicity. The repelling power of the altered amino acids may contribute to additional hydrophobic interactions as the E-helices create the hydrophobic fourfold symmetries. The SASA of C3 and C4 was used to compare their hydrophobicity. The SASA in C3 was the smallest, while C4 had the largest.

Therefore, the hydrophobic interactions at the C-terminus contribute a greater contribution to the final structure and consequently influence the hydrophobicity.

The fluorescence results demonstrated that the modification of C1 increased its pH stability while having little effect on its thermal resistance. And C1 had demonstrated a more stable structure than others in all of the simulation results. The C2, C3, and C4 thermal and pH experiments were carried out using MD simulations. The most unstable compounds were C3, C4, and C2. Due to the mutation of C2 into positively charged amino acids, the initial electrostatic bond has been destroyed, increasing interactions in acidic environments, resulting in lower thermal and pH stability. As a result of changing the hydrophobic channels, C3 and C4 displayed greater fluctuation and instability in simulations, potentially leading to significant changes in their structural stability. However, the simulation failed to indicate the denature of all proteins at the experimentally determined temperature and pH.

This study showed that the C-terminal E-helices can significantly affect stability through mutations. Overall, C-terminus mutation research is essential for the security and effectiveness of this ferritin-based vaccine. And MD simulation is a useful method for developing studies, but more research is required to advance this method.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Associate Professor Jingxiu Bi and Dr. Lukas Gerstweiler my principal supervisors from the School of Chemical Engineering and Advanced Materials, The University of Adelaide, for their dedicated support, patience, and time. I am grateful for their constructive guidance and help throughout my candidature. It is a pleasure to work with you.

I am grateful to all the staff members in the School of Chemical Engineering and Advanced Materials for their support, both in administration and laboratory services.

My gratitude to my colleagues, Bingyang Zhang, Afshin Karami and Yiran Qu. It was great to share the journey with you. It was challenging but at the same time, the most interesting and rewarding experience.

I greatly appreciate my friends, Sijia Fu, Lei Guo, Ay Chin Hee and Yang Li, who gave me unconditional support.

Table of Contents

Table of Figures

Abbreviation

SpFN	Spike Ferritin Nanoparticle
TFR1	Transferrin receptor 1
MD	Molecule dynamic simulation
EBNA1	Epstein-Barr nuclear antigen 1
Ftn	Classic ferritin
Bfr	Bacterioferritin
DPS	Mini-ferritin
HA	Haemagglutinin
EBV	Epstein-Barr virus
SEC	Size exclusion chromatography
MALS	Multi-angle light scattering
HPLC	High-pressure liquid chromatography
DLS	Dynamic light scattering
HIC	Hydrophobic interaction chromatography
RPC	Reverse-phase chromatography
IEC	Ion exchange chromatography
GF	Gel filtration
AC	Affinity chromatography
IMAC	Immobilized metal ion affinity chromatography
DSC	Differential scanning calorimetry
ΔH	Enthalpy
CD	Circular dichroism

AAs	Aromatic amino acid
RMSD	Root mean square deviation
Rg	Radius of gyration
DCCM	Dynamic cross-correlation map
RMSF	Root mean square fluctuation
CI2	Chymotrypsin inhibitor 2
E. coli	Escherichia coli
HFn	Human heavy-chain ferritin
HCPs	Host cell proteins
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
LB	Luria-Bertani medium
IPTG	isopropyl -d-1-thiogalactopyranoside
PB	Phosphate buffer
CV	Column volume

Chapter 1: Introduction

1.1 Background Introduction

In the past ten years, biological therapeutics have become very popular, including antibody therapy, cell therapy and T-cell therapy. The market share of biopharmaceuticals is expected to increase to $399 billion by 2025 (Moorkens et al., 2017). The drug manufacturing processes can be divided into upstream and downstream operations. The upstream process is the protein design and expression to produce the target protein for further steps. While the downstream process includes protein purification and isolation, protein is exposed to different environments, including chromatography resin, freeze-thaw and storage (Li et al., 2021a). The details of the drug manufacturing process are shown in Figure 1.

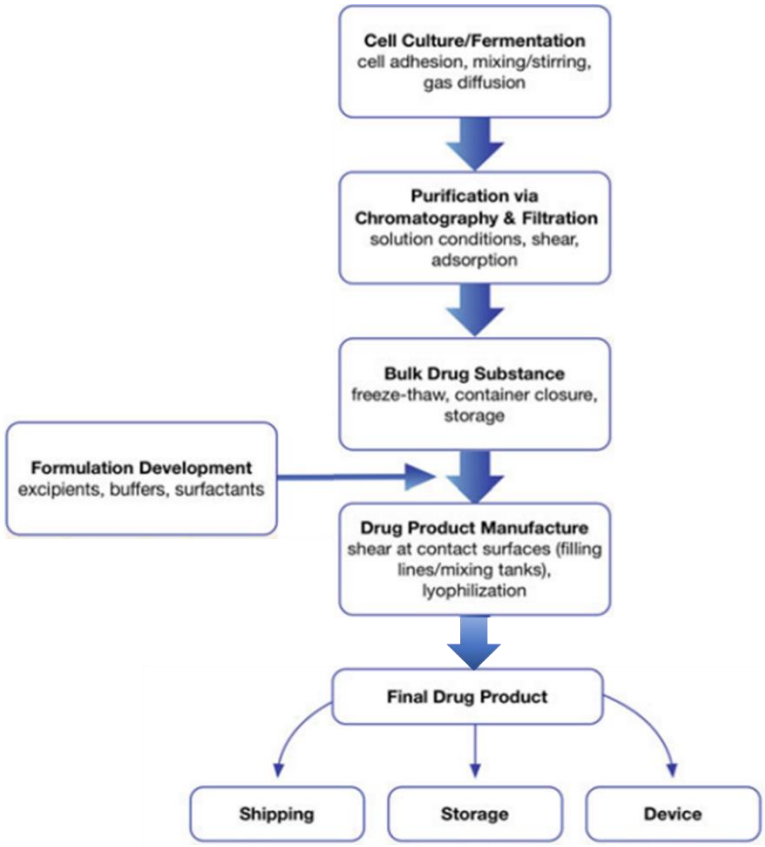

Figure 1. Flowchart of drug development and manufacturing processes (Li et al., 2021a).

The specific protein structures are related to their functions, so preventing protein unfolding is critical for protein therapeutics to maintain their designed position and safety (Wang & Roberts, 2010). The protein unfolding process needs little energy to achieve, and protein adsorption at the interface can cause the protein to denature to form aggregation or particles (Yano, 2012). In utilisation, the collection and other protein unfolding can cause severe damage to safety and efficacy (Li et al., 2021b). As a result, it's vital to maintain structural stability during drug development, manufacturing and future storage.

Ferritin is gaining attention as an excellent drug delivery vehicle in drug development as the cavity or surface can be loaded with a drug of interest (Gagliardi et al., 2022; Lee, Cho, & Kim, 2022; Tang et al., 2022) Targeted therapy is the treatment of most interest in modern clinical practice. In the treatment of common tumours, drugs can be designed to specifically select cancer-causing sites in the body to combine and cause the specific death of tumour cells without affecting the normal tissue cells surrounding the tumour. Ferritin can often be studied as a drug carrier (He, Fan, & Yan, 2019). The issue of targeted drug delivery has always been a hot issue in the field of targeted therapy (Lv, He, & Wu, 2021). Ferritin is involved in iron metabolism and plays a key role in maintaining the body's iron metabolic homeostasis (Srivastav et al., 2022) and cellular antioxidant processes, and is associated with inflammation, vascular proliferation, and tumours. Ferritin is physically stable and is not easily affected by temperature and pH (Liu & Theil, 2003; Kim et al, 2011). Structurally ferritin has internal cavities, is low cost, and can be produced on a large scale (Lawson et al., 1991). In addition, ferritin can be surface modified by simple chemical or genetic routes to achieve efficient targeting (Kanekiyo et al., 2013; Uchida et al., 2007). These characteristics make ferritin a highly promising drug delivery vehicle. In addition, genetic modification and surface modification of ferritin is the most common methods used in the development of ferritin drug carriers (An, Lee, & Kang, 2012; Yin et al., 2021). Coupling folic acid molecules to the surface of ferritin nanocages can

significantly increase the aggregation of drug molecules on the surface of tumour cells (Tesarova et al., 2019). The loading of photosensitisers on the surface of ferritin allows the combination of chemotherapy and photodynamic therapy to effectively inhibit tumour growth after light exposure with minimal impact on normal tissues, ultimately achieving good therapeutic results (Zhen et al., 2013). The development of ferritin drug carriers is a key project in the field of medicine, and the expansion of their functions is conducive to the development of precision therapy.

A study has shown that the Spike Ferritin Nanoparticle (SpFN) COVID-19 vaccine not only elicits an effective immune response but may also provide broad protection against the feared SARS-CoV-2 variant as well as other coronaviruses (Johnston et al., 2022).

The SpFN vaccine is a protein subunit nanoparticle vaccine platform, meaning it presents a fragment of the virus to the immune system to elicit a protective response. SpFN contains multiple coronavirus spike-in proteins that are attached to the surface of a multifaceted ferritin nanoparticle. In clinical practice, the targeting of drugs has always been a major concern. Ferritin has innate tumour-targeting properties (Lee et al., 2017; Yoo et al., 2020). Due to the exuberant growth of tumour cells, the demand for iron ions is much greater than that of normal cells, resulting in the overexpression of Transferrin receptor 1 (TFR1) on the surface of tumour cells (He et al., 2019). Only when TfR1 expression reaches a fixed threshold does ferritin enter the cell and release the encapsulated drug in the acidic microenvironment characteristic of tumour cells.

In this project, human heavy-chain ferritin is used because ferritins are found in almost all the living-organism, have self-assembly characteristics, and have high stability against temperature, pH and chemicals (He et al, 2015). In previous work comparing epitope insertion at N-terminus, C-terminus and DE loop region, ferritins with C-terminus insertion showed the

best stability. (Lindberg et al, 2006; Kim et al, 2001; Choe et al, 2000) There are many studies on the relationship between transformation at different positions on C-terminus and protein performance (Ingrassia et al, 2006). However, understanding the mutation at the C-terminus with linked epitope is limited. In this work, a series of hydrophobic, thermal and pH experiments were conducted to understand the effect of the modifications at the C-terminus on protein stability performance. Furthermore, molecule dynamic (MD) simulation was used to explore the underlying mechanisms of engineered ferritin stability.

1.2. Aims and Objectives of the project

1.2.1 Aims

This project aims to test the stability of three engineered ferritins with the insertion of Epstein-Barr nuclear antigen 1 (EBNA1) as the epitope model. The study involved molecular dynamic simulation and laboratory experiments.

1.2.2 Objectives

1 Set up the optimised purification methods and scale up the system for the stability study.

2 Use the molecular dynamic simulation to predict the mutant ferritin stability.

3 Study hydrophobic, thermal, pH and chemical stability through experiments.

4 Understand the denature mechanism of the engineered proteins.

1.3 Thesis Outline

Chapter 1: Introduction, a brief introduction of this project, including the background, thesis aims and objectives. The significance of studying the engineered ferritin platforms.

Chapter 2: A literature review. This chapter focus on the theory and other previous research on ferritin structures, functions, and characteristics. In addition, the theory of the characteristic experiments and computation methods.

Chapter 3: Purification of engineered ferritins. In this chapter, a series of purification methods were tested, and a purification process including heat-acid precipitation, hydrophobic interaction chromatography and ion exchange chromatography was used. Furthermore, the structure of engineered ferritins was confirmed by (SEC-MALS) and MD simulation.

Chapter 4: Characteristics performance of engineered ferritins. Two purified ferritins were used to study the thermal and pH stability of the system through the experiments including the HIC and fluorescence, and the simulations.

Chapter 5: Conclusion and future directions. Conclusions of major results in all the chapters and future research related to the use of engineered ferritins as vaccine platforms.

1.4 References

An, Lee, M. S., & Kang, J. H. (2012). Oxidative modification of ferritin induced by methylglyoxal. BMB Reports, 45(3), 147–152. https://doi.org/10.5483/BMBRep.2012.45.3.147

Choe, Li, L., Matsudaira, P. T., Wagner, G., & Shakhnovich, E. I. (2000). Differential Stabilization of Two Hydrophobic Cores in the Transition State of the Villin 14T Folding Reaction. Journal of Molecular Biology, 304(1), 99–115.

Demant. (1984). Transfer of ferritin-bound iron to adriamycin. FEBS Letters, 176(1), 97–100. https://doi.org/10.1016/0014-5793(84)80919-9

Foged, Foged, C., Rades, T., Perrie, Y., & Hook, S. (2015). Subunit Vaccine Delivery (Foged, T. Rades, Y. Perrie, & S. Hook, Eds.; 1st ed. 2015.). Springer New York. https://doi.org/10.1007/978-1-4939-1417-3

Li, Krause, M. E., & Tu, R. (2021). Overview of the Impact of Protein Interfacial Instability on the Development of Biologic Products. In Protein Instability at Interfaces During Drug Product Development (pp. 1–8). Springer International Publishing. https://doi.org/10.1007/978-3-030-57177-1_1

Gagliardi, Truffi, M., Tinelli, V., Garofalo, M., Pandini, C., Matteo Cotta Ramusino, Perini, G., Costa, A., Negri, S., Mazzucchelli, S., Bonizzi, A., Sitia, L., Busacca, M., Sevieri, M., Mocchi, M., Ricciardi, A., Prosperi, D., Corsi, F., Cereda, C., & Morasso, C. (2022). Bisdemethoxycurcumin (BDC)-Loaded H-Ferritin-Nanocages Mediate the Regulation of Inflammation in Alzheimer's Disease Patients. International Journal of Molecular Sciences, 23(16), 9237–. https://doi.org/10.3390/ijms23169237

He, & Marles-Wright, J. (2015). Ferritin family proteins and their use in bionanotechnology. New Biotechnology, 32(6), 651–657. https://doi.org/10.1016/j.nbt.2014.12.006

He, Fan, K., & Yan, X. (2019). Ferritin drug carrier (FDC) for tumor targeting therapy. Journal of Controlled Release, 311-312, 288–300. https://doi.org/10.1016/j.jconrel.2019.09.002

Ingrassia, Gerardi, G., Biasiotto, G., & Arosio, P. (2006). Mutations of Ferritin H Chain C-Terminus Produced by Nucleotide Insertions Have Altered Stability and Functional Properties. Journal of Biochemistry (Tokyo), 139(5), 881–885. https://doi.org/10.1093/jb/mvj101Johnston, S. C., Ricks, K. M., Lakhal-Naouar, I., Jay, A., Subra,

Johnston, Ricks, K. M., Lakhal-Naouar, I., Jay, A., Subra, C., Raymond, J. L., King, H. A. D., Rossi, F., Clements, T. L., Fetterer, D., Tostenson, S., Cincotta, C. M., Hack, H. R., Kuklis, C., Soman, S., King, J., Peachman, K. K., Kim, D., Chen, W.-H., … Pitt, M. L. M. (2022). A SARS-CoV-2 Spike Ferritin Nanoparticle Vaccine Is Protective and Promotes a Strong Immunological Response in the Cynomolgus Macaque Coronavirus Disease 2019 (COVID-19) Model. Vaccines (Basel), 10(5), 717–. https://doi.org/10.3390/vaccines10050717

Kim, Kim, Y.-H., & Lee, J. (2001). Thermal Stability of Human Ferritin: Concentration Dependence and Enhanced Stability of an N-Terminal Fusion Mutant. Biochemical and Biophysical Research Communications, 289(1), 125–129. https://doi.org/10.1006/bbrc.2001.5931

Kim, Rho, Y., Jin, K. S., Ahn, B., Jung, S., Kim, H., & Ree, M. (2011). pH-Dependent Structures of Ferritin and Apoferritin in Solution: Disassembly and Reassembly. Biomacromolecules, 12(5), 1629–1640. https://doi.org/10.1021/bm200026v

Lawson, Artymiuk, P. J., Yewdall, S. J., Smith, J. M. A., Livingstone, J. C., Treffry, A., Luzzago, A., Levi, S., Arosio, P., Cesareni, G., Thomas, C. D., Shaw, W. V., & Harrison, P. M. (1991). Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. Nature (London), 349(6309), 541–544. https://doi.org/10.1038/349541a0

Lee, Lee, E. J., Kim, S., Nam, G., Kih, M., Hong, Y., Jeong, C., Yang, Y., Byun, Y., & Kim, I.-S. (2017). Ferritin nanocage with intrinsically disordered proteins and affibody: A platform for tumor targeting with extended pharmacokinetics. Journal of Controlled Release, 267, 172–180. https://doi.org/10.1016/j.jconrel.2017.08.014

Li, Krause, M. E., & Tu, R. (2021). Protein instability at interfaces during drug product development : fundamental understanding, evaluation, and mitigation (Li, M. E. Krause, & R. Tu, Eds.; 1st ed. 2021.). Springer. https://doi.org/10.1007/978-3-030-57177-1

Liu, Lin, Q., Fu, Y., Huang, S., Guo, C., Li, L., Wang, L., Zhang, Z., & Zhang, L. (2020). Target delivering paclitaxel by ferritin heavy chain nanocages for glioma treatment. Journal of Controlled Release, 323, 191–202. https://doi.org/10.1016/j.jconrel.2019.12.010

Lindberg, Haglund, E., Hubner, I. A., Shakhnovich, E. I., & Oliveberg, M. (2006). Identification of the Minimal Protein-Folding Nucleus through Loop-Entropy Perturbations. Proceedings of the National Academy of Sciences-PNAS, 103(11), 4083–4088. https://doi.org/10.1073/pnas.0508863103

Liu, Jin, W., & Theil, E. C. (2003). Opening Protein Pores with Chaotropes Enhances Fe Reduction and Chelation of Fe from the Ferritin Biomineral. Proceedings of the National Academy of Sciences - PNAS, 100(7), 3653–3658. https://doi.org/10.1073/pnas.0636928100

Lv, He, W., & Wu, W. (2021). Editorial of Special Issue of Hot Topic Reviews in Drug Delivery. Acta Pharmaceutica Sinica. B, 11(8), 2094–2095. https://doi.org/10.1016/j.apsb.2021.08.005

Moorkens, Meuwissen, N., Huys, I., Declerck, P., Vulto, A., & Simoens, S. (2017). The Market of Biopharmaceutical Medicines: A Snapshot of a Diverse Industrial Landscape. Frontiers in Pharmacology, 8, 314–314. https://doi.org/10.3389/fphar.2017.00314

O'Hagan, & De Gregorio, E. (2009). The path to a successful vaccine adjuvant – "The long and winding road." Drug Discovery Today, 14(11-12), 541–551. https://doi.org/10.1016/j.drudis.2009.02.009

Srivastav, Mir, I. A., Bansal, N., Singh, P. K., Kumari, R., & Deshmukh, A. (2022). Serum Ferritin in Metabolic Syndrome—Mechanisms and Clinical Applications. Pathophysiology (Amsterdam), 29(2), 319–325. https://doi.org/10.3390/pathophysiology29020023

Tang, Li, M., Chen, L., Dai, A., Liu, Z., Wu, M., Yang, J., Hao, H., Liang, J., Zhou, X., & Qian, Z. (2022). Codelivery of SARS-CoV-2 Prefusion-Spike Protein with CBLB502 by a Dual-Chambered Ferritin Nanocarrier Potentiates Systemic and Mucosal Immunity. ACS Applied Bio Materials, 5(7), 3329–3337. https://doi.org/10.1021/acsabm.2c00328

Tesarova, Charousova, M., Dostalova, S., Bienko, A., Kopel, P., Kruszyński, R., Hynek, D., Michalek, P., Eckschlager, T., Stiborova, M., Adam, V., & Heger, Z. (2019). Folic acid-mediated re-shuttling of ferritin receptor specificity towards a selective delivery of highly cytotoxic nickel(II) coordination compounds. International Journal of Biological Macromolecules, 126, 1099–1111. https://doi.org/10.1016/j.ijbiomac.2018.12.128

Wang, & Roberts, C. J. (2010). Aggregation of Therapeutic Proteins (1. Aufl.). Wiley.

Yano. (2012). Kinetics of protein unfolding at interfaces. Journal of Physics. Condensed Matter, 24(50), 503101–503101. https://doi.org/10.1088/0953-8984/24/50/503101

Yin,, Zhang, B., Lin, J., 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, 21(10), 630–642. https://doi.org/10.1002/elsc.202000105

Yoo, Bae, S. M., Seo, J., Jeon, I. S., Vadevoo, S. M. P., Kim, S.-Y., Kim, I.-S., Lee, B., & Kim, S. (2020). Designed ferritin nanocages displaying trimeric TRAIL and tumor-targeting

peptides confer superior anti-tumor efficacy. Scientific Reports, 10(1), 19997–. https://doi.org/10.1038/s41598-020-77095-x

Zhang, Wang, R., Li, Z., Wang, L., Gao, Z., Tu, Y., & Cao, X. (2021). Anti-EGFR Single-Chain Fv Antibody Fragment Displayed on the Surface of Ferritin H-Chain Protein Nanoparticle for Asthma Therapy. ACS Applied Bio Materials, 4(9), 6690–6702. https://doi.org/10.1021/acsabm.1c00308

Zhen, Tang, W., Guo, C., Chen, H., Lin, X., Liu, G., Fei, B., Chen, X., Xu, B., & Xie, J. (2013). Ferritin Nanocages To Encapsulate and Deliver Photosensitizers for Efficient Photodynamic Therapy against Cancer. ACS Nano, 7(8), 6988–6996. https://doi.org/10.1021/nn402199g

Chapter 2: literature review

2.1 Ferritin structures and characteristics

Ferritin is a self-assembling protein found in most living organisms. that is simple to isolate and manufacture (Theil et al., 2006). Ferritins have a symmetric and well-structured hollow sphere. Due to its flexible self-assembly properties and symmetrical structure, ferritin is suitable for administration using drug delivery or antigen binding. Numerous studies (Kang et al., 2011; Klem et al., 2010) have demonstrated the role of ferritin as a drug delivery platform. In addition, the ferritin complex contains 24 identical subunits allowing it to be easily modified to meet different physical and chemical requirements. This has also encouraged more researchers to improve the medical applications of ferritin by mutating its amino acids (Kanekiyo et al., 2013; Uchida et al., 2007).

Classic ferritin (Ftn), bacterioferritin (Bfr), and mini-ferritin (DPS) are all members of the ferritin family and each has a unique structure and function (He & Marles-Wright, 2015). Homo ferritin is a typical ferritin. It consists of 24 subunits (20 kDa) of both heavy (H) and light (L) chains with molecular weights of 21 kDa and 19 kDa respectively (Zhang & Orner, 2011). Almost 55% of the amino acid sequences of ferritin are identical between the chains. Ferritin has an inner diameter of 8 nm and an outside diameter of 12 nm (Lawson et al., 1991). The internal cavity can hold up to 4,500 iron atoms in the form of an aqueous iron oxide mineral core with varying amounts of phosphate (Theil, 1987). Each monomer consists of a four-alpha-helix bundle (A, B, C, and D helices) and a short fifth helix (E helix) at the C-terminus (Figure *2*) (Zhang & Orner, 2011).

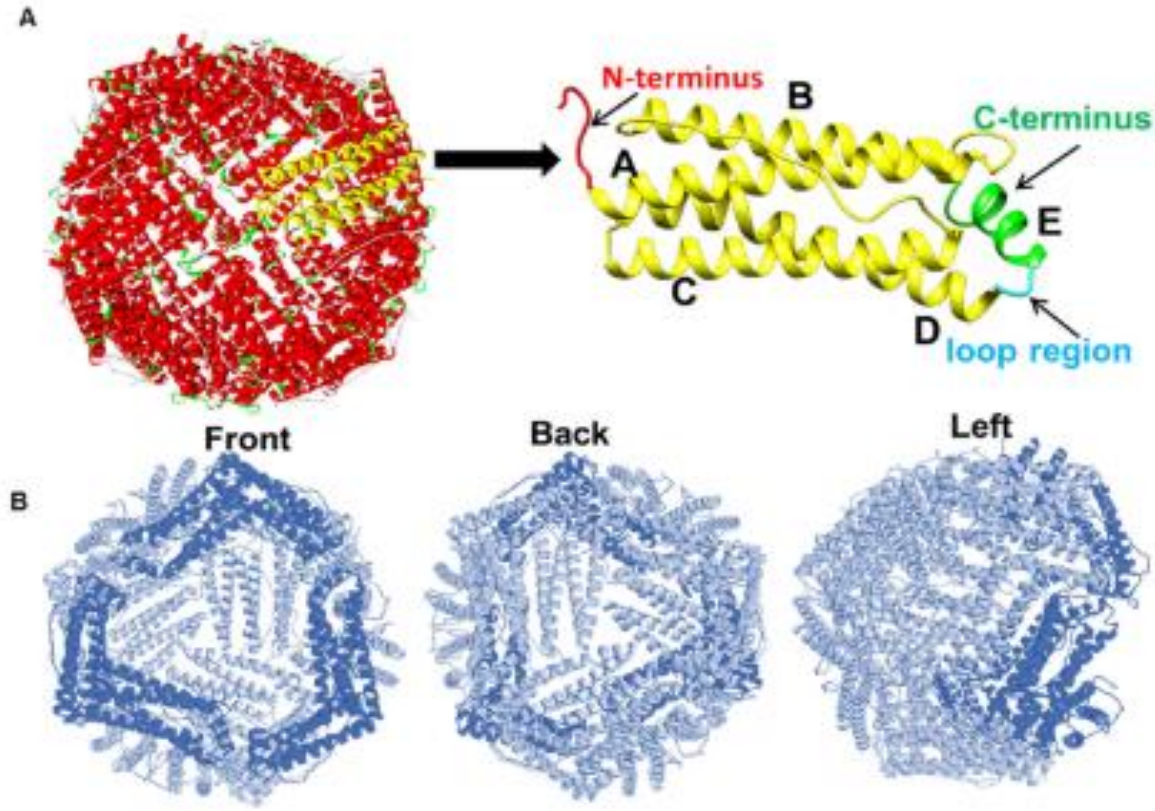

Figure 2. A. Structure of human heavy-chain ferritin, shown in ribbon, and one subunit is shown in yellow (PDB ID: 2fha). Chain A, B, C, and D are yellow, chain E is green, and N-terminus is red. B. Human heavy-chain structure shown from different directions (Qu et al., 2021).

Each subunit in the octahedral cage structure interacts with six neighbouring monomers via three types of symmetries-related interfaces. There are twelve dimerization interaction surfaces along the two-fold axis, eight trimerization interfaces along the three-fold axes, and six tetramerization interfaces along the four-fold axes (Figure 3) (Zhang & Orner, 2011).

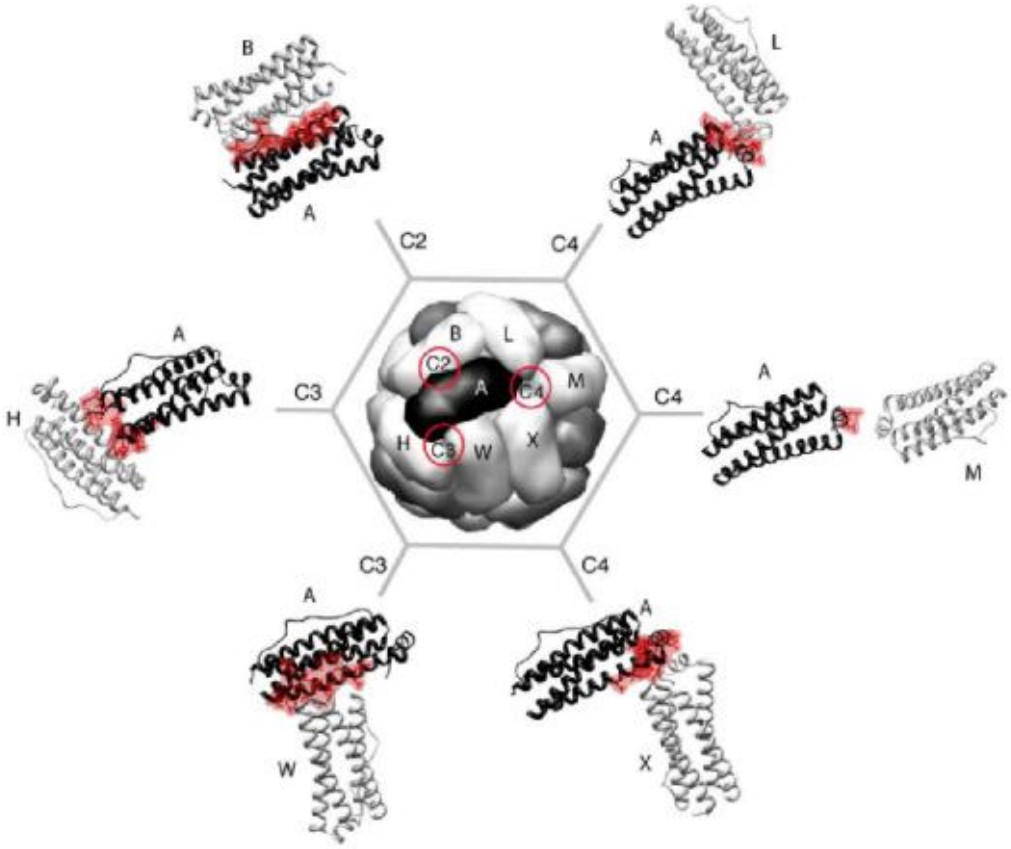

Figure 3. Ferritin is involved in six unique inter-subunit interactions at the respective symmetry-related interfaces (C2, C3, C4). The subunits (B, H, W, X, M, and L) that interact with subunit A are indicated (Zhang & Orner, 2011).

Self-assembly of ferritin into nanoparticles has demonstrated robust thermal and chemical stability. Furthermore, ferritin consists of eight units with triple axial symmetry. This high rigid and consistent structures are easier and more accurate for antigen delivery (López-Sagaseta et al., 2016).

2.2 Ferritin as vaccine delivery platform

2.2.1 Ferritin used for nanoparticle vaccine

Ferritin is widely utilized in the manufacture of nanoparticle vaccines due to its highly symmetric structure and its ability to self-assemble under varying pH and buffer conditions. The team of Kanekiyo introduced influenza haemagglutinin (HA) at the interface of the

neighbouring subunit of Helicobacter pylori-haem ferritin. The ferritin then self-assembled and featured eight trimeric viral spikes on its surface (Figure *4*) (Kanekiyo et al., 2013).

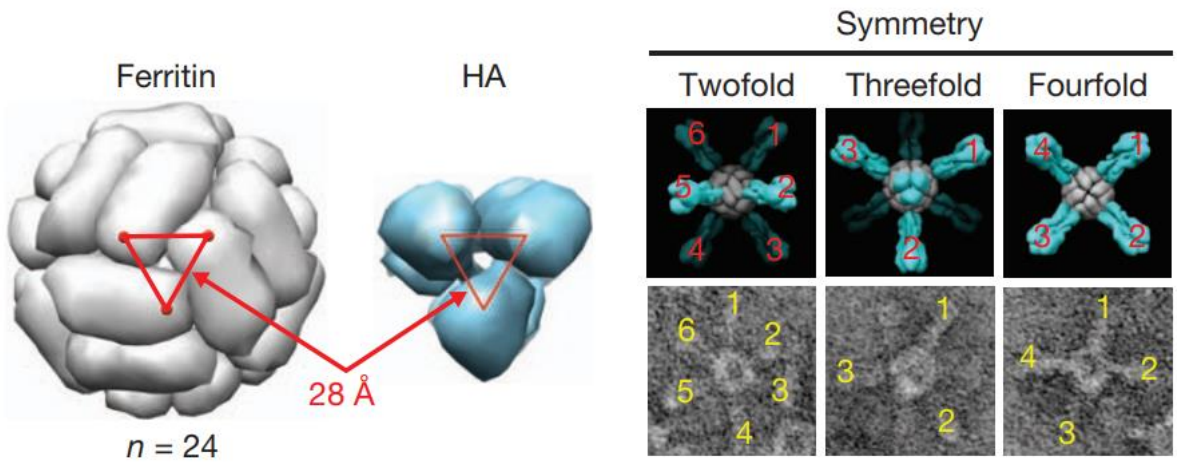

Figure 4. A. An assembled ferritin nanoparticle and an HA trimer (PDB: 3sm5). The connecting points are shown in red. B. Computational and TEM images of HA-nanoparticles with two, three, and fourfold axes are shown (right, top, and bottom panels). The visible HA spikes in the images are labelled (Kanekiyo et al., 2013).

The immunization test on mice revealed that the genetically engineered nanoparticle influenza vaccine had more than tenfold higher inhibitory antibody titters than other commercially available influenza vaccinations (Kanekiyo et al., 2013). Moreover, the new influenza vaccine offers immune protection against the two different strains of the influenza virus (Kanekiyo et al., 2013). The proposed influenza vaccine enhances the vaccination's efficacy and extends immunity to the virus (Kanekiyo et al., 2013). Further immune experiments demonstrated that ferritin neither promotes autoimmunity nor impacts iron homeostasis in mice (Kanekiyo et al., 2013). The study of the HA-Ftn nanoparticle vaccine showed how ferritin could be used to make vaccines and highlighted the possibilities of vector-induced autoimmunity (He & Marles-Wright, 2015).

2.2.2 Drug delivery via self-assembly loading font size

Ferritin is a promising option for drug delivery. Under adequate acidic buffer conditions, ferritins can disassemble, enabling medicines or nanoparticles that are smaller than the ferritin cavity size to enter. When the pH returns to about seven, ferritins can combine back to the original symmetric form. This characteristic can load metal-containing medicines because they bind to ferritin more efficiently (Domı et al., 2004; Ji et al., 2012) and tend to diffuse into ferritin through tunnels (Laghaei et al., 2013). By surface modification with peptides or labels, ferritin can be manipulated to target specific cells and tumours for more efficient therapeutics while avoiding damaging others. Figure *5* illustrates a typical nanoparticle platform configuration (Schoonen & Hest, 2014).

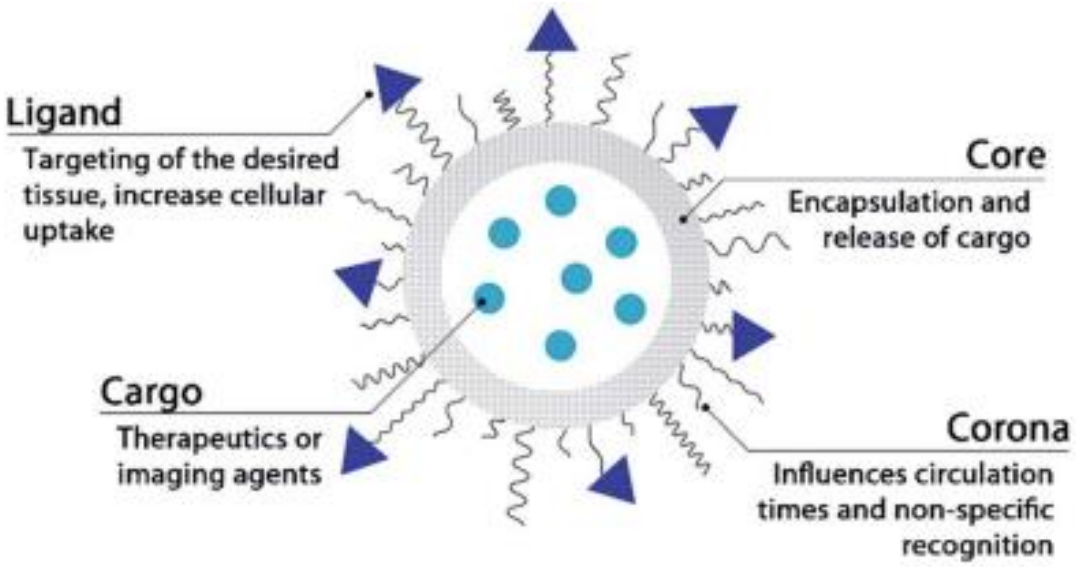

Figure 5. Structure a typical nanoparticle platform for drug delivery (Schoonen & Hest, 2014).

The ability of mutant ferritin cages to be exploited as a multifunctional drug carrier and enable cell-targeting drug delivery was proven. The researchers designed an engineered human heavy-chain ferritin and introduced a synthesized iron oxide nanoparticle into the ferritin cavity (Uchida et al., 2006). Furthermore, they bound RGD-4C, a cell-targeting peptide, on the

engineered ferritin's exterior surface. The vitro experiment results indicated that the designed ferritin system could achieve iron oxide loading and targeting treatment (Uchida et al., 2006).

2.3 Mutation of ferritins

Previous research on ferritin alterations and mutations has demonstrated that different residues and locations can impact ferritin assembly and function. Based on previous investigations, many mutations have little effect on the protein self-assembly, emphasizing their robustness and suitability for engineering modifications and applications (Zhang & Orner, 2011). In contrast, other single-point mutations may substantially affect the self-assembly of ferritin (Zhang & Orner, 2011).

The Nabel team also published an EBV vaccine based on self-assembling ferritin. Epstein-Barr virus (EBV) glycoprotein 350/220 is the primary immune target (gp350). The design resulted in self-assembled ferritins displayed in a symmetric array with gp350. In a mouse study, this modified vaccine showed a 10- to 100-fold increase in neutralisation capacity over soluble gp350 (Kanekiyo et al., 2015). The researchers insert the viral trimeric glycoprotein at 3-fold axes. They combined the amino-terminal extension of bullfrog ferritin with *H.pylori* ferritin to create hybrid ferritin with an antigen attachment site distributed equally across the surface (Figure 6) (Kanekiyo et al., 2015). These symmetrically arrayed antigens showed higher potential to activate B cells and enhance immunity through low affinity BCRs (Jardine et al., 2013).

Figure 6. Assembled hybrid H. pylori-bullfrog ferritin and engineered insertion site are red (Kanekiyo et al., 2015).

A series of C-terminus mutations were studied to determine the relationship between amino acid and ferritin self-assembly functions. The E helices along the four-fold symmetry axis are not required for human ferritin heavy-chain assembly. Various ferritin mutants were generated from the C-terminus deletions extending to different lengths, and it was found that ferritin with deletions extending Gly-159 can still assemble appropriately (Luzzago & Cesareni, 1989). However, the loss of the last four amino acids in the D helix can hinder the recovery of formed ferritin (Luzzago & Cesareni, 1989). On the other hand, the E helix of ferritin could choose to either flip inside (flip) or outside (flop) (Figure 7) (Luzzago & Cesareni, 1989).

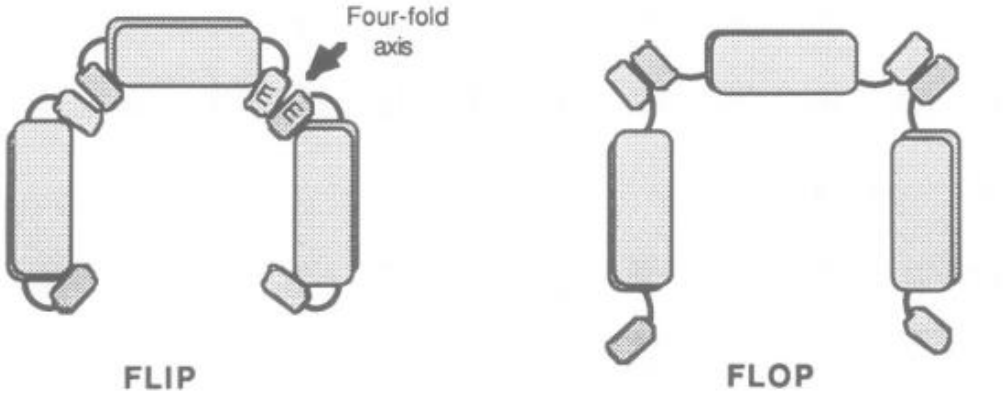

Figure 7. These two configurations have four-helix bundles, while the E helix flip model flip inside and the flop model flip outside (Luzzago & Cesareni, 1989).

Based on the previous research, deletion of 22 residues at the c-terminus of human h-chain ferritin (including the E helix) has little effect on the assembly of the mutant, but still catalyzes iron oxidation (Levi et al., 1988). Moreover, another study showed that 14-26 residue extensions could be hidden in the cavity (Ingrassia et al., 2006). However, when they deleted a further six residues at the c-terminus (28 in total), the mutant did not assemble into a Ferritin-like protein. The results showed that the six residues at the C-terminus tip of the D helix could affect the self-assembly and basic molecular function of human h-chain ferritin (Levi et al., 1989).

On the other hand, the N-terminal mutation led to a different result. After deleting the first thirteen amino acids at the N-terminus, researchers discovered that the protein could still form and function correctly (Levi et al., 1988). The effect of n-terminal modifications on ferritin stability is even more pronounced. Due to the absence of N-terminus residues, there were fewer inter- and intra-subunit hydrogen bonds. Furthermore, the out-of-bounds are responsible for the mutants' reduced strength. Another study showed the effect of mutation on the BC loop.

They removed two residues (L82 and I85), and the mutants were unable to form ferritin cages. This is a result of intermediate changes in dimer assembly (Yoshizawa et al., 2007).

Another research is related to the renaturation of human H-chain ferritin altered at the N-terminus near the 4-fold axis (Leu-169-Arg), the 3-fold axis (Asp-131-Ile + Glu-134-Phe) or the 2-fold axis (Ile-85-Cys) (Figure *8*). They give rise to monomers and dimers to assembly intermediates of various oligomers (Santambrogio et al., 1997). Thus, the results supported that the symmetric subunit dimers are the fundamental building units for ferritin self-assembly and conforms to the recombination path where dimer is added stepwise after dimer binding until the ferritin cage is completed (Santambrogio et al., 1997).

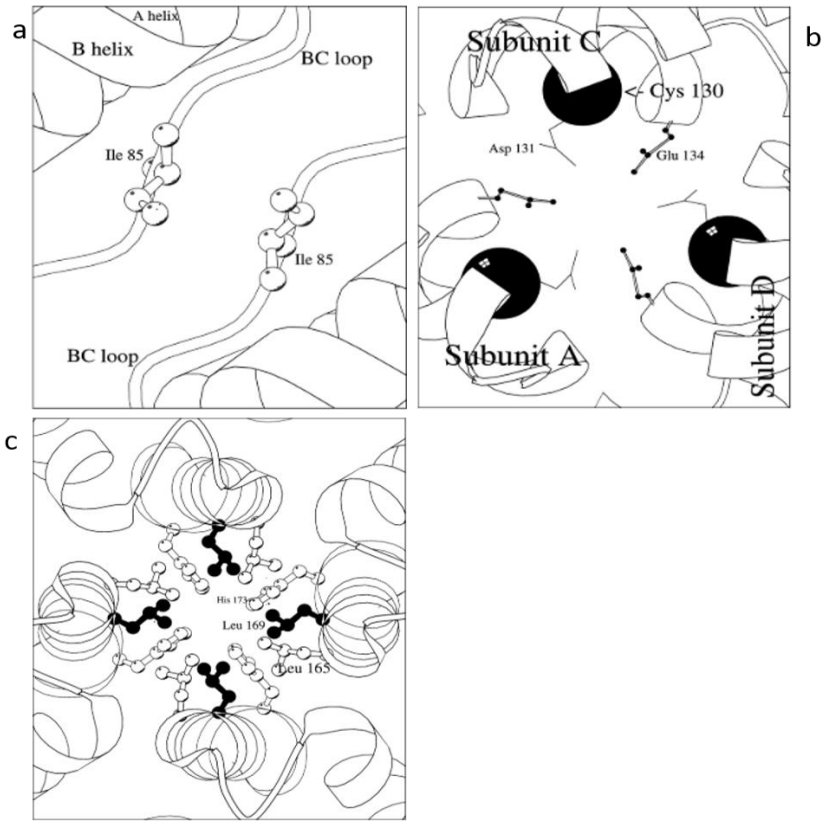

Figure 8. a. Schematic diagram of the interface between two subunits around a 2-fold axis, viewed from the molecule's centre; b. schematic diagrams of inter subunit contacts in the neighbourhood of a 3-fold axis; c. diagram of parts of four subunits adjacent to a 4-fold axis (Santambrogio et al., 1997).

2.4 Characteristics of ferritin structure and stability studies

A major obstacle to the development and production of protein products is protein aggregation. Challenges include protein aggregation affecting product quality and safety (Ross & Wolfe, 2016). Protein aggregation starts a series of processes from changes in internal structure involving a limited concentration of protein interactions, which also includes the process by which amino acids begin to interact with surrounding proteins (Fink, 1998). Firstly, reversible dimer or oligomer are formed which can be separated from or in equilibrium with the monomer. Furthermore, particles grow up until they reach the size of macroscopic particles or droplets. (Roberts, 2014). Irreversible polymerization dissociates only under extreme denaturing conditions and the secondary or tertiary structure is compromised. (Figure 9) (Roberts, 2014).

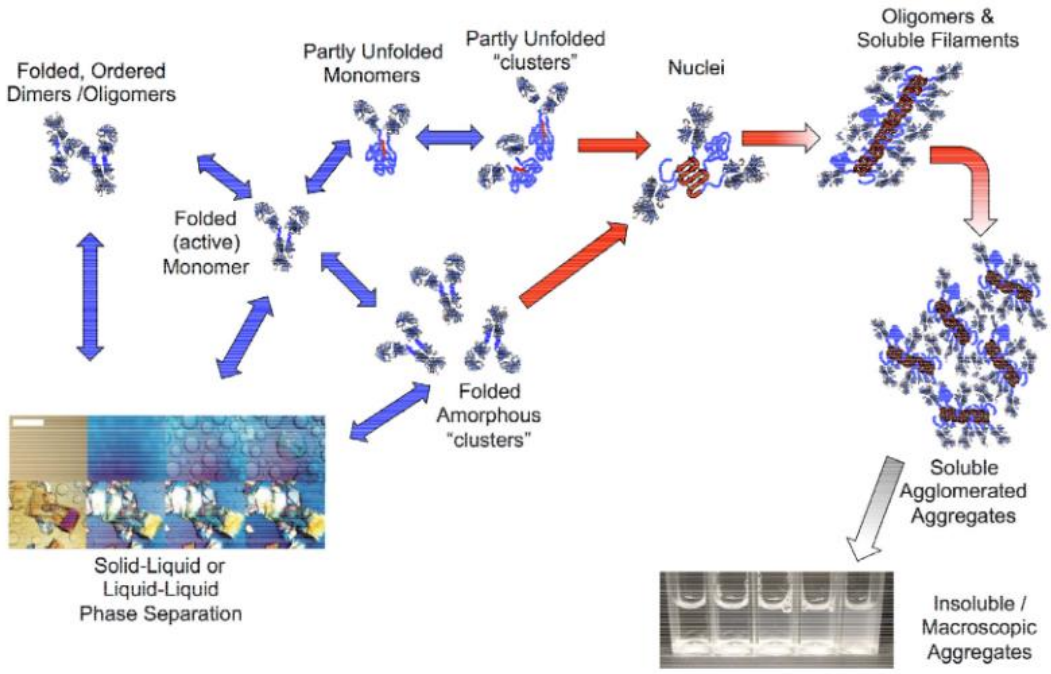

Figure 9. Illustration of different aggregated states of folded or fully/partially unfolded proteins (Roberts, 2014).

Protein aggregation begins with a molecular aggregate, and the maximum size varies based on the protein and environment (Roberts, 2014). The soluble protein aggregates can be observed by size exclusion chromatography (SEC-HPLC) (Brych et al., 2010). SEC-HPLC is a typical approach for detecting and determining protein products and studying aggregation formation (Den Engelsman et al., 2011). However, this method still has some limitations. SEC-HPLC would (1). dilute samples; (2). cause aggregates dissociation; (3). require a trial of mobile running buffer; (4). have a limited resolution and protein size range (He et al., 2010). In addition, the characteristic of the SEC column can cause high back pressure or temperature aggregation (Bobály et al., 2016). In some studies, SEC-HPLC fails to indicate the soluble protein aggregates (Barnard et al., 2011; He et al., 2010; Paul et al., 2017).

Dynamic light scattering (DLS) would be used as an alternative technique (Figure 10). DLS is a straightforward and non-destructive approach for studying the large molecular weight aggregates in their original buffer (Al-Ghobashy et al., 2017). The reason is that DLS measures the protein's mutual diffusion coefficient, which is not affected by molecular weight (Drenski et al., 2013). Moreover, the scattered intensity of DLS is autocorrelated to the mutual diffusion coefficient of the particles. Thus DLS has no requirement for sample concentration either (Goldburg, 1999).

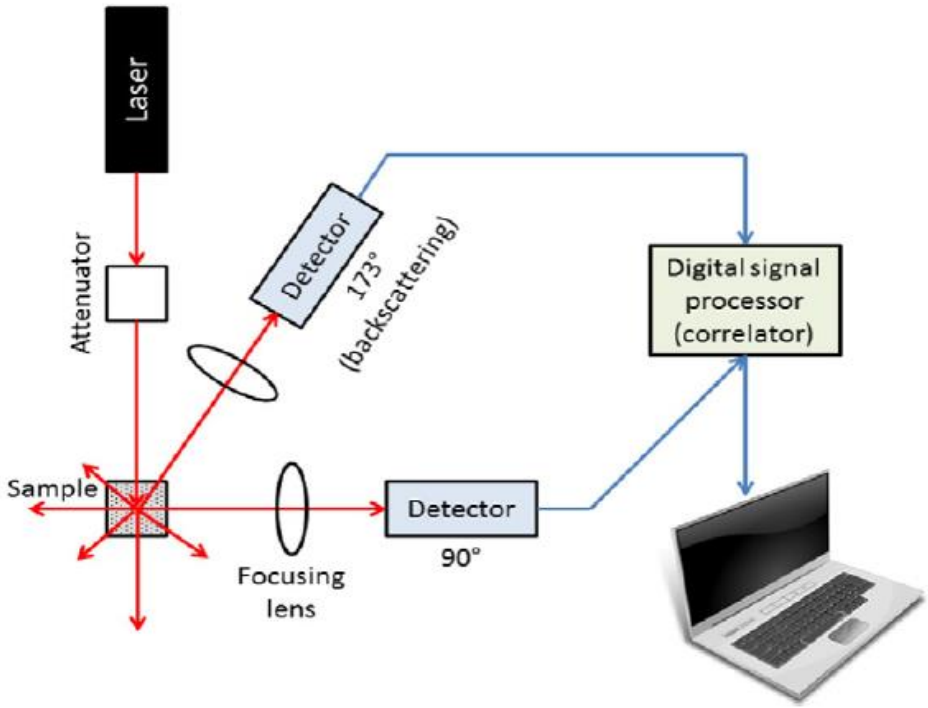

Figure 10. Instruments of DLS (Bhattacharjee, 2016).

Chromatography is the most common method of protein purification (Cytiva, 2020). These methods separate proteins according to their different affinity toward the stationary phase. Different stationary phases are available, allowing the separation based on various principles, as outlined in Table 1.

Table 1. Protein properties are used for chromatographic purification.

Protein Property	Method
Hydrophobicity	Hydrophobic interaction chromatography (HIC); Reverse-phase chromatography (RPC)
Charge	Ion exchange chromatography (IEC)
Size	Gel filtration (GF)
Specific ligand recognition	Affinity chromatography (AC)

Metal ion binding	Immobilized metal ion affinity chromatography (IMAC)
Isoelectric point	Chromatofocusing

HIC is based on surface hydrophobic properties of various proteins (Figure 11). For this approach, the protein will interact with the hydrophobic surface of the chromatography medium. A high-ionic-strength buffer enhances the interaction. Proteins bind to the HIC column at high ionic strength (typically 1 to 2 M ammonium sulphate or 3 M Sodium Chloride). The salt concentration is then decreased, and the different components are eluted differently. The more hydrophobic the protein of interest, the more strongly it binds to the column and is subsequently eluted at lower salt concentrations. The HIC was shown to be feasible in hydrophobic study by comparing the elution time of different proteins under the same condition.

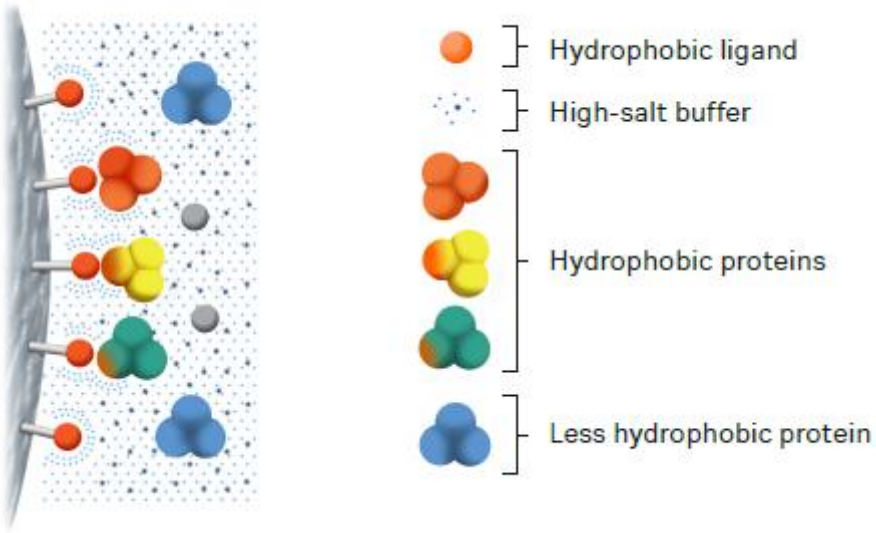

Figure 11. Schematic of HIC (Cytiva, 2020).

Fluorescence spectroscopy (Figure 12) can provide helpful information on the folding tertiary structure, stability, and ligand interaction of ferritins. In addition, fluorescence is sensitive to

changes in ferritin's secondary and tertiary structures (Kessenbrock & Groth, 2017). Consequently, fluorescence is a powerful tool for the ferritin stability study.

Proteins can emit photons when excited by ultraviolet light (Eftink, 2000). Each fluorophore has unique and observable fluorescence properties, including intensity, the position of emission wavelength, and lifetime. Each structure shows a different spectrum as a function of the measured temperature as the power and wavelength change (J. Albani, 2007). Another reason why fluorescence spectroscopy is widely used is that it only requires only low sample concentration.

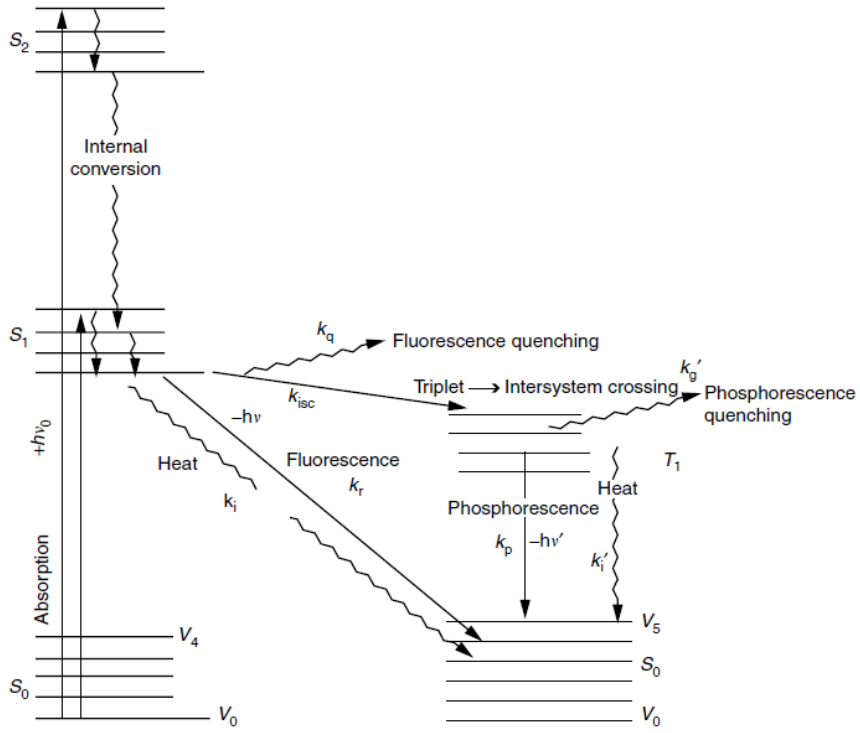

Figure 12. Jablonski (electronic transitions) diagrams (J. Albani, 2007).The electronic states (S) and the vibrational levels (V) and energy transfer constant (k) of a protein molecule.

Differential scanning calorimetry (DSC) is a susceptible thermo-analysis technique. DSC aims to investigate changes in the heat-induced conformation of proteins, including thermal stability, folding mechanism, and thermal behaviour (Johnson, 2013). The thermal stability of a protein

usually contains two components: equilibrium and kinetic. Both of them can be indicated by DSC. For proteins, thermal absorption peaks were observed in the scan. After correcting the instrument baseline, transitional baseline, and normalization to concentration, the height can be integrated to obtain a direct calorimetric measurement of Enthalpy (ΔH) and Melting temperature (Tm) (Eq. 1).

$$\Delta H = \int_{T0}^{T1} Cp.dT \qquad \textit{Eq. 1}$$

DSC does not depend on changes in spectral signal, such as CD and fluorescence-based thermal denaturation. In addition, this method directly measures process-related heat absorption and therefore addresses multiple overlapping functions. In addition, DSC can measure samples under pressure, which allows scanning temperatures up to approximately 140 °C without the need for liquid boiling (Sawano et al., 2008). Ferritin has shown versatile thermal stability, and the DSC can reach a considerably high temperature suitable for ferritin study.

2.5 Simulation for protein stability studies

Protein folding is a multi-step process, and the mechanism and force behind protein folding are complex. In addition, most current studies of protein folding, and unfolding are based on well-structured native proteins. Engineered or mutated proteins were investigated based on their amino acid sequences to understand the performance of their future utilization. Factors that promote protein folding into specific structures include (1). Hydrogen bonds; (2). Van der Waals forces; (3). Backbone angles; (4). Electrostatic interactions; (5). Hydrophobic interactions; (6). Chain Entropy (Daggett, 2006). It may take a long time to consider all these interactions and a lot of calculations are required. Molecular Dynamics (MD) simulation can calculate all details of the interaction on the atom level of protein and solve on the timescale of nanoseconds to microseconds (ns to ms) (Kumar, 2020). There are many studies on proteins

carried out at elevated temperatures, pH, or chemicals using MD simulation (Camilloni et al., 2008; Day et al., 2002; Li et al., 2022; Palanisamy et al., 2020).

For a specific protein structure, including atom coordinates, several MD simulation programs can be used to analyse the dynamics of the system (Beck & Daggett, 2004). The typical processes of GROMACS MD simulation are presented in the flow chart (Figure 13).

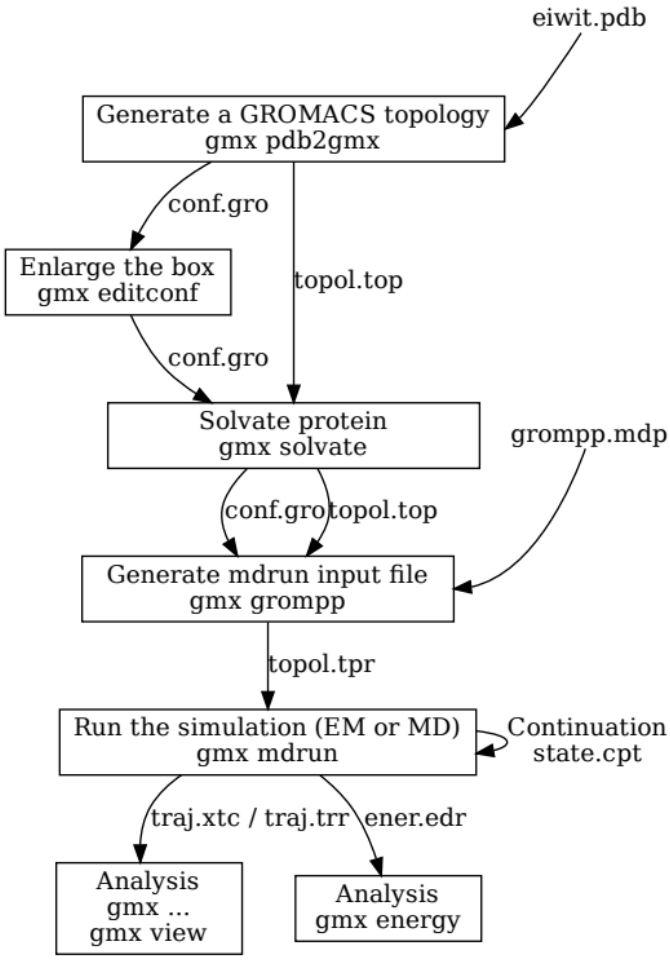

Figure 13. A flow chart of a typical GROMACS MD simulation of protein in the box with water (Spoel, 2021).

The all-atom molecular mechanics field describes the potential energy of the system analytically with the geometry of the atomic centres (Beck & Daggett, 2004). In addition, they define the functional form and parameter sets for MD simulation to measure the potential

power of atoms or coarse-grained particles in a system. Widely used biomolecular force fields include Amber (Cornell et al., 1996), CHARMM (MacKerell et al., 1998), GROMOS (Oostenbrink et al., 2004), and OPLS-AA (Jorgensen et al., 1996). These force fields can be classified according to their molecular properties: bonded and nonbonded. Furthermore, hydrogen bonding is not considered in these force fields, and hydrogen bonds are represented by the combination of the LJ and Coulomb terms (Guvench & Mackerell, 2008).

In Palanisamy's study (Palanisamy et al., 2020), MD simulations were used to predict the relationship between protein stability and imidazolium-water (ImH+Wn). They chose three aromatic amino acid (AAs) models for simulations: Phe, Trp, and Tyr. The AA model was placed in a cubic box during the simulation and the amber FF14SB force field was used (Maier et al., 2015). Water molecules were treated with SPC/E water models (Mark & Nilsson, 2001). The system was equilibrated for 250 ps at 300 K and 1 bar. After equilibration, the simulation was performed using GROMACS 2016.3 package, a simulation run for 100 ns and visualized by the PyMOL program. The interaction energies of ImH+ and water with aromatic acids were calculated by the MMPBSA tool (Kumari et al., 2014). For the denaturing ability, the root means square deviation (RMSD), the radius of gyration (Rg), and the dynamic cross-correlation map (DCCM) were used to confirm the structure changes.

Protein unfolding simulations could run under high temperatures to accelerate the simulation speed. In a previous study, the researchers run a series of unfolding simulations of chymotrypsin inhibitor 2 (CI2) under temperatures from 298 K to 498 K. The simulation length of CI2 denatures from 94 ns to 20 ns. At 298 K, CI2 kept stable for a 50 ns simulation; at 348 K, the experimentally melting point of the protein, CI2, unfolds over 25 ns and then refold over the last 35 ns. Under the extremely high temperature, the protein underwent a whole thermal denature process, starting from the unfolding transition state to the fully denature unfolding. In

conclusion, they found that a temperature much higher than the melting point was more suitable for the simulation of thermal denature study (Day et al., 2002).

By using PDB 2PQR (Dolinsky et al., 2004) to protonate protein and generate structure files for different pH, MD simulation could be used to study the protein's unfolding mechanism and interface interaction for the pH-induced unfolding (Li et al., 2022). The researchers used PDB 2PQR to protonate ferritin at pH 2 and 7, which protonated all the Asp, Glu, and His residues except Asp 174. The simulation found that the dimer is the most stable structure under pH 2, an extremely low pH. Furthermore, the monomer's rotation caused the dimer's swell, triggering the unfolding process (Figure 14). This simulation study helped to understand the fundamental mechanism (Figure 15) of ferritin pH denature.

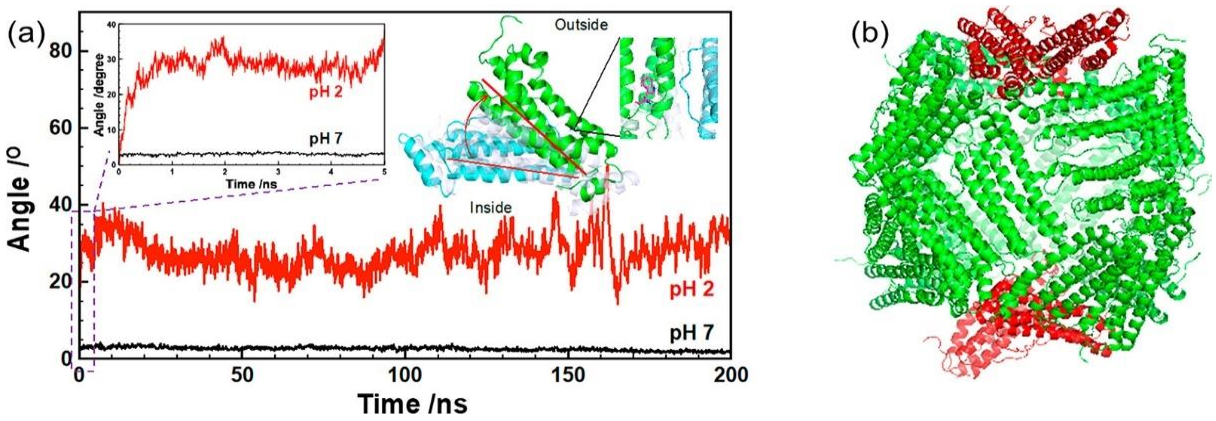

Figure 14(a). Rotation of dimer during disassembly; (b). The cycle of dimmers is highlighted in red (Li et al., 2022).

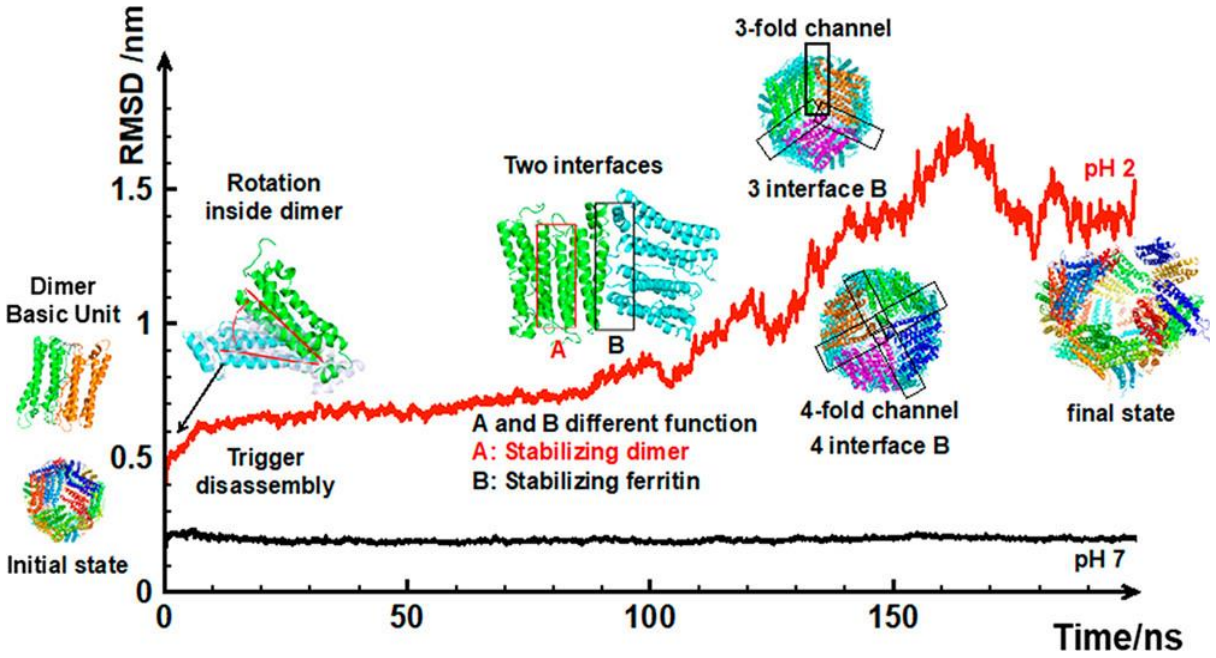

Figure 15 Ferritin unfolding mechanism under low pH (Li et al., 2022).

Adding various chemicals can also alter the stability of proteins. Some, like guanidine chloride (GdmCl) and urea destabilised proteins (Alonso et al, 1991). Given the significance of GdmCl and urea to our understanding of protein stability, it is noteworthy that a molecular understanding of how they alter protein structure remains unclear. As a result, there is still debate regarding whether denaturant chemicals influence solution characteristics or interact directly with proteins. Moreover, the hydrophobic side chains of the polypeptide and the

comparatively nonpolar protein backbone both contribute to the denature transition, and the chemical denaturants are connected to this process (England et al, 2011).

Therefore, Tiana's team simulated the chemical denaturation of protein L by placing the protein into two typical protein-denaturing chemicals, water, urea, and GdmCl two typical protein denature chemicals. The simulation was conducted under the Gromos96 force field. A box filled with denaturants and water was prepared by equilibrium at 100 bar and with charge optimizations, and then a 10-ns MD simulation was carried out. The researchers compared the structure changes in the two denaturants during simulations. They found that the simulations in GdmCl showed that the α-helix destabilized fast while the β-sheet was relatively stable. In contrast, the simulation in urea suggested that the disassembly started from the *b*-sheet (Figure 16). The simulation results suggested that the unfolding mechanism depends on the denaturants that experiments cannot reveal (Camilloni et al., 2008).

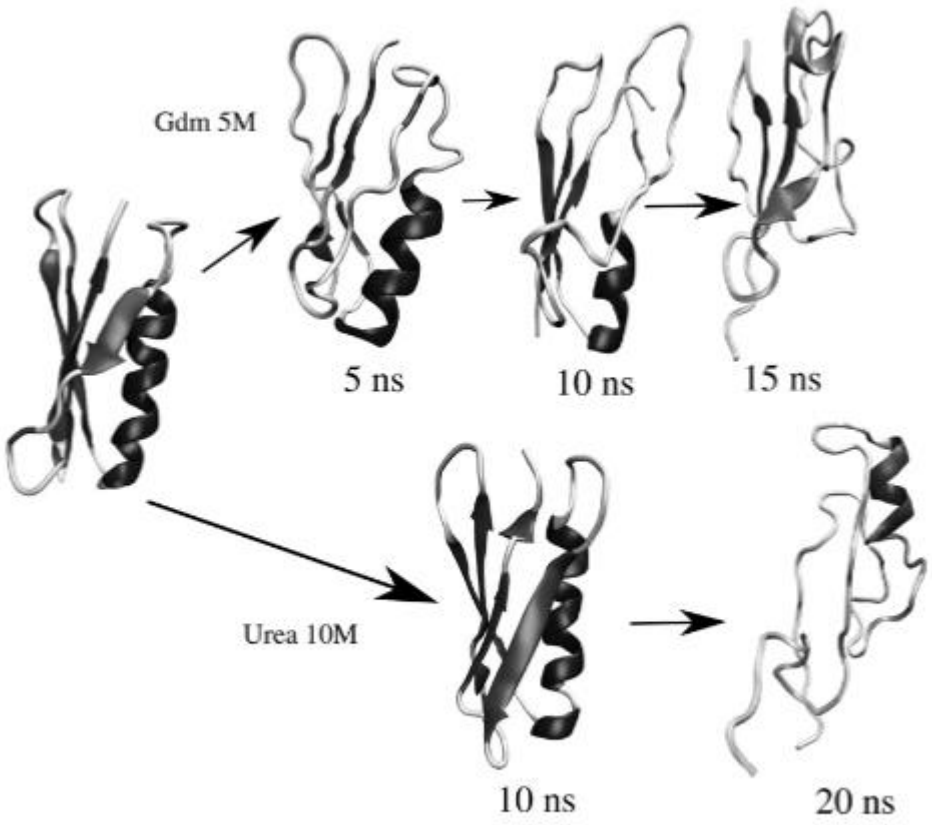

Figure 16 Protein L was unfolding in GdmCl and urea at 400 K (Camilloni et al., 2008).

2.6 References

Al-Ghobashy, M. A., Mostafa, M. M., Abed, H. S., Fathalla, F. A., & Salem, M. Y. (2017). Correlation between Dynamic Light Scattering and Size Exclusion High-Performance Liquid Chromatography for Monitoring the Effect of pH on Stability of Biopharmaceuticals. Journal of chromatography. B, Analytical technologies in the biomedical and life sciences, 1060, 1-9. https://doi.org/10.1016/j.jchromb.2017.05.029

Albani, J. R. (2007). Principles and applications of fluorescence spectroscopy. Blackwell Science.

Alonso D, Dill K. Solvent denaturation and stabilization of globular proteins. Biochemistry. 1991; 30:5974–5985. [PubMed: 2043635]. https://doi.org/10.1021/bi00238a023

Barnard, J. G., Singh, S., Randolph, T. W., & Carpenter, J. F. (2011). Subvisible Particle Counting Provides a Sensitive Method of Detecting and Quantifying Aggregation of Monoclonal Antibody Caused by Freeze-Thawing: Insights Into the Roles of Particles in the Protein Aggregation Pathway. Journal of Pharmaceutical Sciences, 100(2), 492-503. https://doi.org/10.1002/jps.22305

Beck, D. A. C., & Daggett, V. (2004). Methods for molecular dynamics simulations of protein folding/unfolding in solution. Methods (San Diego, Calif.), 34(1), 112-120. https://doi.org/10.1016/j.ymeth.2004.03.008

Bhattacharjee, S. (2016). DLS and zeta potential – What are they and what are they not? Journal of controlled release, 235, 337-351. https://doi.org/10.1016/j.jconrel.2016.06.017.

Bobály, B., Sipkó, E., & Fekete, J. (2016). Challenges in the liquid chromatographic characterisation of proteins. Journal of chromatography. B, Analytical technologies in the biomedical and life sciences, 1032, 3-22. https://doi.org/10.1016/j.jchromb.2016.04.037

Brych, S. R., Gokarn, Y. R., Hultgen, H., Stevenson, R. J., Rajan, R., & Matsumura, M. (2010). Characterisation of antibody aggregation: Role of buried, unpaired cysteines in particle formation. Journal of Pharmaceutical Sciences, 99(2), 764-781. https://doi.org/10.1002/jps.21868

Camilloni, C., Guerini Rocco, A., Eberini, I., Gianazza, E., Broglia, R. A., & Tiana, G. (2008). Urea and Guanidinium Chloride Denature Protein L in Different Ways in Molecular Dynamics Simulations. Biophysical Journal, 94(12), 4654-4661. https://doi.org/10.1529/biophysj.107.125799

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., & Kollman, P. A. (1996). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J. Am. Chem. Soc. 1995, 117, 5179−5197. Journal of the American Chemical Society, 118(9), 2309-2309. https://doi.org/10.1021/ja955032e

Cytiva. (2020). Hydrophobic Interaction Chromatography (HIC). https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=16231

Daggett, V. (2006). Protein Folding−Simulation. Chemical Reviews, 106(5), 1898-1916. https://doi.org/10.1021/cr0404242

Day, R., Bennion, B. J., Ham, S., & Daggett, V. (2002). Increasing Temperature Accelerates Protein Unfolding Without Changing the Pathway of Unfolding. Journal of molecular biology, 322(1), 189-203. https://doi.org/10.1016/S0022-2836(02)00672-1

Den Engelsman, J., Garidel, P., Smulders, R., Koll, H., Smith, B., Bassarab, S., Seidl, A., Hainzl, O., & Jiskoot, W. (2011). Strategies for the Assessment of Protein Aggregates in

Pharmaceutical Biotech Product Development. Pharmaceutical Research, 28(4), 920-933. https://doi.org/10.1007/s11095-010-0297-1

Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., & Baker, N. A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Research, 32(Web Server), W665-W667. https://doi.org/10.1093/nar/gkh381

Domı, x, & nguez-Vera, J. M. (2004). Iron(III) complexation of Desferrioxamine B encapsulated in apoferritin. Journal of inorganic biochemistry, 98(3), 469-472. https://doi.org/10.1016/j.jinorgbio.2003.12.015

Drenski, M. F., Brader, M. L., Alston, R. W., & Reed, W. F. (2013). Monitoring protein aggregation kinetics with simultaneous multiple sample light scattering. Analytical biochemistry, 437(2), 185-197. https://doi.org/10.1016/j.ab.2013.02.014

Eftink, M. R. (2000). Intrinsic Fluorescence of Proteins. In Topics in Fluorescence Spectroscopy, Volume 6: Protein Fluorescence (pp. 1-13). Springer. DOI:10.1007/0-306-47102-7_1

England, JL & Haran, G 2011, 'Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back', Annual Review of Physical Chemistry, vol. 62, no. 1, pp. 257–277.

Fink, A. L. (1998). Protein aggregation: folding aggregates, inclusion bodies and amyloid. Folding & design, 3(1), R9-R23. https://doi.org/10.1016/S1359-0278(98)00002-9

Goldburg, W. I. (1999). Dynamic light scattering. American Journal of Physics, 67(12), 1152-1160. https://doi.org/10.1119/1.19101

Guvench, O., & Mackerell, A. D. (2008). Comparison of Protein Force Fields for Molecular Dynamics Simulations. In (pp. 63-88). Humana Press. https://doi.org/10.1007/978-1-59745-177-2_4

He, D., & Marles-Wright, J. (2015). Ferritin family proteins and their use in bionanotechnology. New biotechnology, 32(6), 651-657. https://doi.org/10.1016/j.nbt.2014.12.006

He, F., Phan, D. H., Hogan, S., Bailey, R., Becker, G. W., Narhi, L. O., & Razinkov, V. I. (2010). Detection of IgG Aggregation by a High Throughput Method Based on Extrinsic Fluorescence. Journal of Pharmaceutical Sciences, 99(6), 2598-2608. https://doi.org/10.1002/jps.22036

Ingrassia, R., Gerardi, G., Biasiotto, G., & Arosio, P. (2006). Mutations of Ferritin H Chain C-Terminus Produced by Nucleotide Insertions Have Altered Stability and Functional Properties. Journal of biochemistry (Tokyo), 139(5), 881-885. https://doi.org/10.1093/jb/mvj101

Jardine, J., Julien, J.-P., Menis, S., Ota, T., Kalyuzhniy, O., McGuire, A., Sok, D., Huang, P.-S., MacPherson, S., Jones, M., Nieusma, T., Mathison, J., Baker, D., Ward, A. B., Burton, D. R., Stamatatos, L., Nemazee, D., Wilson, I. A., & Schief, W. R. (2013). Rational HIV Immunogen Design to Target Specific Germline B Cell Receptors. Science, 340(6133), 711-716. https://doi.org/doi:10.1126/science.1234150

Ji, X.-T., Huang, L., & Huang, H.-Q. (2012). Construction of nanometer cisplatin core-ferritin (NCC-F) and proteomic analysis of gastric cancer cell apoptosis induced with cisplatin released from the NCC-F. Journal of proteomics, 75(11), 3145-3157. https://doi.org/10.1016/j.jprot.2012.03.013

Johnson, C. M. (2013). Differential scanning calorimetry as a tool for protein folding and stability. Archives of biochemistry and biophysics, 531(1-2), 100-109. https://doi.org/10.1016/j.abb.2012.09.008

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. Journal of the American Chemical Society, 118(45), 11225-11236. https://doi.org/10.1021/ja9621760

Kanekiyo, M., Bu, W., Joyce, M. G., Meng, G., Whittle, James R. R., Baxa, U., Yamamoto, T., Narpala, S., Todd, J.-P., Rao, Srinivas S., McDermott, Adrian B., Koup, Richard A., Rossmann, Michael G., Mascola, John R., Graham, Barney S., Cohen, Jeffrey I., & Nabel, Gary J. (2015). Rational Design of an Epstein-Barr Virus Vaccine Targeting the Receptor-Binding Site. Cell, 162(5), 1090-1100. https://doi.org/10.1016/j.cell.2015.07.043

Kanekiyo, M., Wei, C.-J., Yassine, H. M., McTamney, P. M., Boyington, J. C., Whittle, J. R. R., Rao, S. S., Kong, W.-P., Wang, L., & Nabel, G. J. (2013). Self-assembling influenza nanoparticle vaccines elicit broadly neutralising H1N1 antibodies. Nature, 499(7456), 102-106. https://doi.org/10.1038/nature12202

Kang, Y. J., Uchida, M., Shin, H.-H., Douglas, T., & Kang, S. (2011). Biomimetic FePt nanoparticle synthesis within Pyrococcus furiosus ferritins and their layer-by-layer formation. Soft matter, 7(23), 11078-11081. https://doi.org/10.1039/c1sm06319g

Kanthe, A. D., Maldarelli, C., & Tu, R. (2021). Interfacial Behaviors of Proteins. In (pp. 51-114). Springer International Publishing. https://doi.org/10.1007/978-3-030-57177-1_3

Kessenbrock, & Groth, G. (2017). Circular Dichroism and Fluorescence Spectroscopy to Study Protein Structure and Protein–Protein Interactions in Ethylene Signaling. Ethylene Signaling, 141–159. https://doi.org/10.1007/978-1-4939-6854-1_12

Klem, M. T., Young, M., & Douglas, T. (2010). Biomimetic synthesis of photoactive a-Fe(2)O(3) templated by the hyperthermophilic ferritin from Pyrococus furiosusf. Journal of materials chemistry, 20(1), 65-67. https://doi.org/10.1039/b918620d

Kumar, A. (2020). Protein Purification, Estimation, Storage, and Effect on Structure–Function–Dynamics. In (pp. 1-22). Springer Singapore. https://doi.org/10.1007/978-981-15-5530-5_1

Kumari, R., Kumar, R., & Lynn, A. (2014). g_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations. Journal of Chemical Information and Modeling, 54(7), 1951-1962. https://doi.org/10.1021/ci500020m

Laghaei, R., Evans, D. G., & Coalson, R. D. (2013). Metal binding sites of human H-chain ferritin and iron transport mechanism to the ferroxidase sites: A molecular dynamics simulation study. Proteins, structure, function, and bioinformatics, 81(6), 1042-1050. https://doi.org/10.1002/prot.24251

Lawson, D. M., Artymiuk, P. J., Yewdall, S. J., Smith, J. M. A., Livingstone, J. C., Treffry, A., Luzzago, A., Levi, S., Arosio, P., Cesareni, G., Thomas, C. D., Shaw, W. V., & Harrison, P. M. (1991). Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts. Nature (London), 349(6309), 541-544. https://doi.org/10.1038/349541a0

Levi, S., Luzzago, A., Cesareni, G., Cozzi, A., Franceschinelli, F., Albertini, A., & Arosio, P. (1988). Mechanism of ferritin iron uptake: activity of the H-chain and deletion mapping of the ferro-oxidase site. A study of iron uptake and ferro-oxidase activity of human liver, recombinant H-chain ferritins, and of two H-chain deletion mutants. The Journal of biological chemistry, 263(34), 18086-18092. https://doi.org/10.1016/S0021-9258(19)81326-1

Levi, S., Luzzago, A., Franceschinelli, F., Santambrogio, P., Cesareni, G., & Arosio, P. (1989). Mutational analysis of the channel and loop sequences of human ferritin H-chain. Biochemical journal, 264(2), 381-388. https://doi.org/10.1042/bj2640381

Li, Z., Maity, B., Hishikawa, Y., Ueno, T., & Lu, D. (2022). Importance of the Subunit–Subunit Interface in Ferritin Disassembly: A Molecular Dynamics Study. Langmuir, 38(3), 1106-1113. https://doi.org/10.1021/acs.langmuir.1c02753

López-Sagaseta, J., Malito, E., Rappuoli, R., & Bottomley, M. J. (2016). Self-assembling protein nanoparticles in the design of vaccines. Computational and structural biotechnology journal, 14(C), 58-68. https://doi.org/10.1016/j.csbj.2015.11.001

Luzzago, A., & Cesareni, G. (1989). Isolation of point mutations that affect the folding of the H chain of human ferritin in E. coli. The EMBO journal, 8(2), 569-576.

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., & Karplus, M. (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. The journal of physical chemistry. B, 102(18), 3586-3616. https://doi.org/10.1021/jp973084f

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. Journal of chemical theory and computation, 11(8), 3696-3713. https://doi.org/10.1021/acs.jctc.5b00255

Mark, P., & Nilsson, L. (2001). Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. The Journal of Physical Chemistry A, 105(43), 9954-9960. https://doi.org/10.1021/jp003020w

Oostenbrink, C., Villa, A., Mark, A. E., & Van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. Journal of computational chemistry, 25(13), 1656-1676. https://doi.org/10.1002/jcc.20090

Palanisamy, K., Prakash, M., & Rajapandian, V. (2020). Combined DFT and MD simulation studies of protein stability on imidazolium–water (ImH+Wn) clusters with aromatic amino acids. New Journal of Chemistry, 44(41), 17912-17923. https://doi.org/10.1039/d0nj03085f

Paul, A. J., Bickel, F., Röhm, M., Hospach, L., Halder, B., Rettich, N., Handrick, R., Herold, E. M., Kiefer, H., & Hesse, F. (2017). High-throughput analysis of sub-visible mAb aggregate particles using automated fluorescence microscopy imaging. Analytical and Bioanalytical Chemistry, 409(17), 4149-4156. https://doi.org/10.1007/s00216-017-0362-2

Qu, Y., Wang, L., Yin, S., Zhang, B., Jiao, Y., Sun, Y., Middelberg, A., & Bi, J. (2021). Stability of Engineered Ferritin Nanovaccines Investigated by Combined Molecular Simulation and Experiments. The journal of physical chemistry. B, 125(15), 3830-3842. https://doi.org/10.1021/acs.jpcb.1c00276

Roberts, C. J. (2014). Protein aggregation and its impact on product quality. Current Opinion in Biotechnology, 30, 211-217. https://doi.org/10.1016/j.copbio.2014.08.001

Ross, P. L., & Wolfe, J. L. (2016). Physical and Chemical Stability of Antibody Drug Conjugates: Current Status. Journal of Pharmaceutical Sciences, 105(2), 391-397. https://doi.org/10.1016/j.xphs.2015.11.037

Santambrogio, P., Pinto, P., Levi, S., Cozzi, A., Rovida, E., Albertini, A., Artymiuk, P., Harrison, P. M., & Arosio, P. (1997). Effects of modifications near the 2-, 3- and 4-fold symmetry axes on human ferritin renaturation. Biochemical journal, 322 (Pt 2)(2), 461-468. https://doi.org/10.1042/bj3220461

Sawano, M., Yamamoto, H., Ogasahara, K., Kidokoro, S.-i., Katoh, S., Ohnuma, T., Katoh, E., Yokoyama, S., & Yutani, K. (2008). Thermodynamic Basis for the Stabilities of Three CutA1s from Pyrococcus horikoshii, Thermus thermophilus, and Oryza sativa, with Unusually High Denaturation Temperatures. Biochemistry (Easton), 47(2), 721-730. https://doi.org/10.1021/bi701761m

Schoonen, L., & Hest, J. C. M. v. (2014). Functionalisation of protein-based nanocages for drug delivery applications. Nanoscale, 6(13), 7124-7141. https://doi.org/10.1039/C4NR00915K

Spoel, L. A. H. v. d. (2021). GROMACS 2021.3 Manual. https://dx.doi.org/10.5281/zenodo.5053220

Theil, E. C. (1987). Ferritin: Structure, Gene Regulation, and Cellular Function in Animals, Plants, and Microorganisms. Annual review of biochemistry, 56(1), 289-315. https://doi.org/10.1146/annurev.bi.56.070187.001445

Theil, Matzapetakis, M., & Liu, X. (2006). Ferritins: iron/oxygen biominerals in protein nanocages. Journal of Biological Inorganic Chemistry, 11(7), 803–810. https://doi.org/10.1007/s00775-006-0125-6

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. (2006). Targeting of Cancer Cells with Ferrimagnetic Ferritin Cage Nanoparticles. Journal of the American Chemical Society, 128(51), 16626-16633. https://doi.org/10.1021/ja0655690

Uchida, M., Klem, M. T., Allen, M., Suci, P., Flenniken, M., Gillitzer, E., Varpness, Z., Liepold, L. O., Young, M., & Douglas, T. (2007). Biological Containers: Protein Cages as Multifunctional Nanoplatforms. Advanced materials (Weinheim), 19(8), 1025-1042. https://doi.org/10.1002/adma.200601168

Yoshizawa, K., Mishima, Y., Park, S.-Y., Heddle, J. G., Tame, J. R. H., Iwahori, K., Kobayashi, M., & Yamashita, I. (2007). Effect of N-terminal Residues on the Structural Stability of Recombinant Horse L-chain Apoferritin in an Acidic Environment. Journal of biochemistry (Tokyo), 142(6), 707-713. https://doi.org/10.1093/jb/mvm187

Zhang, Y., & Orner, B. P. (2011). Self-assembly in the ferritin nano-cage protein superfamily. International journal of molecular sciences, 12(8), 5406-5421. https://doi.org/10.3390/ijms12085406

Chapter 3: Fermentation and purification process development of engineered human heavy-chain ferritins

3.1 Introduction

The need for protein treatments will continuously grow in the coming decades, as mentioned in Chapter 2 .Proteins of high purity and proper shape are essential for the safety and effectiveness of protein-related products (Wang & Roberts, 2010). The development of the purification of these mutant ferritins will be thoroughly covered in this chapter.

Escherichia coli (*E. coli*) remains one of the best options for industrial manufacturing and research labs because of its usually high yields and easy of expression (Chen, 2012). *E. coli* systems are frequently employed in research on the production and structure of modified proteins (Gordon et al., 2008; Koehn & Hunt, 2009; Zou et al., 2016). *E. coli* has been thoroughly studied for a considerably long time. It is affordable, easy to grow, and suitable for large-scale protein expression in industrial bioreactors or lab fermentation (Middelberg et al., 2011). All the mutant proteins in this chapter are expressed in *E. coli*. However, one of the trickiest parts of making human heavy-chain ferritin (HFn) is the downstream isolation procedure (Wang & Roberts, 2010). The target proteins should be isolated from nucleic acids and host cell proteins of *E. coli*. The separation processes are difficult, especially for the soluble form of designed proteins (Palomares et al., 2004).

Native HFn is comparable stable, is thermally stable up to 85 °C, and has high pH stability in pH 2-10 (Liu & Theil, 2003; Kim et al, 2011). Host cell proteins (HCPs) are typically precipitated by heat-acid precipitation and separated by centrifugation (Song et al., 2021). Additionally, size-exclusion chromatography (SEC) can be used to refine them even further owing to the molecular weight of HFn, which is approximately 450 kDa. The nucleic acid of

the host bacteria is typically removed using chromatography or DNase and RNase treatment (Li et al.).

Normally, functionalized ferritin was purified with the same technique as recombinant HFn (Boumaiza et al., 2017; Weidenbacher et al., 2022). Although a few studies have been conducted on modifications to the purification process for functionalized HFns, the underlying mechanisms remain unclear. It has been demonstrated in previous research that 15 linkers and EBNA1 peptide inserted at C-terminus do not impact HFn assembly and purification in *E. coli* (Qu et al., 2021). Generally, there is a close connection between genetic modifications and changes in purification techniques since they have an impact on the HFn's biochemical properties and purification efficiency.

This chapter describes the development of the purification strategies for the four mutant HFns, comprising chromatography and heat-acid precipitation. The initial stage involved the removal of the HCPs under various acid and temperature conditions, and the HIC was then utilized to remove the nucleic acid. However, an additional IEC step was adopted when the purification performance was subpar. Proteins were quantified, and their structures were verified using Bradford, sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-Page), and size-exclusion chromatography-multi-angle light scattering (SEC-MALS).

Table 2 List of general purification methods of human heavy chain ferritins.

Target protein	**HCPs removal**	**Nucleic acid removal**	**Purity of the target protein**	**The recovery rate of the target protein**
Recombinant human ferritin heavy chain (Song et al., 2021)	0.5-1 M NaCl pH 4.4 70 or 75 ℃ for 10 min	Hydrophobic interaction chromatography (HIC): Hiscreen Capto Butyl pre-packed column	Above 98 %	~ 66 %
ferritin-EBNA1 (Qu et al., 2021)	1M NaCl pH 5 60 ℃	HIC: HiTrap Butyl FF	Not reported	Not reported
HFn (Yin, 2021)	pH 4.5 60 ℃	HIC: HiTrap Butyl FF	98.52 %	98.18 %

3.2 Materials and Methods

3.2.1 Materials

The mutant ferritins came from earlier work by Yiran (Qu et al., 2022). To obtain the control protein ($F_1L_3E_1$) for this experiment, she first inserted the EBNA1 (PDB ID: 2fz3) epitope after adding three linkers (GGS) to the C-terminus of human heavy chain ferritin (PDB ID: 2fha). The E-helices of F1L3E1 were modified to introduce different amino acids for creating the engineered ferritin (C1-C4) in this project. The order of the amino acids is listed in Table 1.

Table 3. The amino acid sequence of control HFn and four mutated HFns.

Mutant No.	**Amino sequence from position "164" to "173."**
$F_1L_3E_1$ (Native ferritin)	GL AEYLF DKH
C1 (No charge)	GL AQYLF QKH
C2 (Same charge)	GL ARYLF RKH
C3 (Most hydrophobic)	GV IEIVI DKH
C4 (Most hydrophilic)	GQ NEQNQ DKH

12% $^{w}/_{v}$ separation gel and 5% $^{w}/_{v}$ stacking layer were used for SDS-PAGE. A 5 × loading buffer containing 6% $^{w}/_{v}$ 1 M Tris-HCL at pH 6.8, 25% $^{w}/_{v}$ glycerol, 1 mg/ml bromophenol blue, 2% $^{w}/_{v}$ SDS, and 5% $^{w}/_{v}$ -mercaptoethanol was employed to prepare the samples. After being heated at 100 °C for 10 min, the samples were loaded into the gel wells. The concentration and the separation were performed for 20 minutes at 90 V and 30 minutes at 200 V, respectively. The background was removed by washing the gel with de-stain buffer (10% $^{v}/_{v}$ ethanol and 10% $^{v}/_{v}$ acetic acid) overnight after the gel had been stained with 0.25% $^{w}/_{v}$ Coomassie R-250 for 60 minutes.

The chromatography columns used in this chapter were all purchased from GE Healthcare (USA). All analytical-grade reagents were obtained by Chem-Supply (Australia). Milli Q water from a Millipore purification system (Merck, USA) was used throughout the procedure.

3.2.2 Fermentation of different E. coli strains

The expression of engineered ferritins was accomplished using *E. coli* BL21 (DE3). A 100 µl *E. coli* strain was inoculated in a 50 ml Luria-Bertani (LB) medium with 50 µl ampicillin and incubated overnight at 37 °C in a Ratek incubator, shaken at 150 rpm. Then, 10 ml *E. coli* was transferred to 500 ml LB medium with 500 µl ampicillin and shacked at 200 rpm until OD600 reached 0.8. Afterwards, 1 mM isopropyl -d-1-thiogalactopyranoside (IPTG) was added for protein induction and shaken for another 4 h at 200 rpm. Subsequently, Bacteria were harvested by centrifugation for 30 minutes at 4 °C at 3700 rpm.

The bacteria pallets were re-suspended in 20 mM phosphate buffer (PB, 0.2 M Na_2HPO_4, 0.2 M NaH_2PO_4), 2 mM EDTA, and pH 7 lysis buffer at a mass/volume ratio of approximately 1:20 (mass of bacteria (g): volume of lysis buffer (ml). Next, the pallets were broken by an ultrasonicate for 20 minutes using 300 W power, 4-s on, 6-s off (Ningbo Keyi Technology, China).

3.2.3 Optimisation of host cell proteins removal by heat-acid precipitation

The buffer pH, salt concentration, and heating temperature were all tested during the heat-acid precipitation of mutated HFns to determine the best conditions for high protein purity and recovery. First, 0.5 ml *E. coli* lysate supernatant (9 mg/ml) was mixed with 0.3 ml acid buffer (2 M acetic acid - sodium acetate at pH 4.5, 5.0, and 5.5) and 1 M sodium chloride (NaCl). Then, the mixture was adjusted to a total volume of 1 ml with Milli Q water. In a water bath, the samples were heated to 50, 55, or 60 °C for 10 minutes. After heating, the samples were centrifuged for 3 minutes at 12,000 rpm and 4 °C to remove any precipitated proteins. The

protein concentrations in the supernatant were determined using the Bradford assay after centrifugation. The protein compositions in the supernatants were examined using 12% reducing SDS-PAGE (Bio-Rad, USA). The purity of the target protein was determined by a densitometry scan performed with the Image J software to calculate the target protein's recovery rate after heat-acid precipitation through Equation 2.

$$Recovery\ rate\ (\%) = 100\% \times \frac{Amount\ of\ target\ protein\ in\ step\ n}{amount\ of\ target\ protein\ in\ supernant\ of\ bacteria\ lysate} \qquad Eq.\ 2$$

3.2.4 Nucleic acid removal by hydrophobic interaction chromatography (HIC)

For all mutated HFns, HIC was used to remove the host cell nucleic acid. With the Hitrap Butyl FF column (1 ml, GE Healthcare, USA), absorbance 260 and absorbance 280 were measured during the HIC. The equilibrium buffer was 100 mM phosphate buffer (PB), 1.2 M ammonia sulphate (AS) at pH 6.5, and the elution buffer was 20 mM PB with pH 6.5. The heat-acid precipitation samples were diluted 5 times with the equilibrium buffer (100 mM PB, 1.2 M AS, pH 6.5) to adjust to the same conductivity as the buffer. The HIC column was equilibrated with equilibrium buffer (100 mM PB, 1.2 M AS, pH 6.5) for 15 CVs before loading a 100 ml sample. After the flowthrough peak was completed, the column was washed with a gradient of 0 to 100% elution buffer (20 mM PB, pH 6.5) with 10 CV. Flow-through samples and peaks from the HIC were collected, the protein concentration was calculated using the Bradford assay, and the purity of the samples was determined using a 12% reducing SDS-PAGE.

3.2.5 Further purification using ion exchange chromatography (IEC)

Considering that C2, C3, and C4 were unable to achieve the desired purity, an additional ion exchange chromatography (IEC) step was added to improve the purity even further. Prior to IEC, 10 mL of each sample was desalted into an equilibrium buffer (20 mM PB, pH 7.5) with a HiPrep 26/10 desalting column (GE Healthcare, USA) at a flow rate of 10 ml/min. IEC was performed in a 3 mL Sartobind Q (Sartorius AG, Germany) column with an equilibrium buffer

of 20 mM PB at pH 7.5 and an elution buffer with an extra 2 M NaCl. The column was first equilibrated at 3 ml/min with 15 CV equilibrium buffer (20 mM PB, pH 7.5). Subsequently, the desalted sample was loaded and washed with 15 CV equilibrium buffer (20 mM PB, pH 7.5). Afterwards, the sample was eluted in 10 CV with 0-100% elution buffer (20 mM PB, 2 M NaCl, pH 7.5). The flow-through and peak samples were collected for Bradford assay and 12% SDS-PAGE analysis.

3.2.6 Engineered ferritin's structure confirmation

The hydrodynamic radius and molecular weight of ferritins were measured using size-exclusion chromatography - multi-angle light scattering (SEC-MALS). Superose 6 increase 10/300 GL (GE Healthcare, USA) was connected to a high-performance liquid chromatography system (Shimadzu, Japan) with a Wyatt Optilab refractive index and a Wyatt DAWN MALLS detector (Wyatt, USA). The running buffer was 20 mM PB, pH 7.0, the protein concentration was 13.45 mg/ml, and the loading volume was 50 µl. ASTRA 6 and GraphPad Prism 9 were used to analyse the SEC-MALS data.

3.2.7 Computational analysis of the structure and properties of engineered ferritins

Gromacs used the Phoenix high-performance computer to run a series of molecule dynamic (MD) simulations (the University of Adelaide). The structures of human heavy-chain ferritin and EBNA1 were obtained from the RCSB Protein Data Bank (PDB ID: 2FHA and 2FZ3).

The final structure of mutated proteins was established using Discovery Studio. The OPLS - AA/L all-atom force field and SPC/E water mode were adopted during the simulation. The protein was placed in a cubic box with periodic boundary conditions in all directions, and the box's diameter was 13.7 nm, with a minimum distance of 0.1 nm between the protein and the box's wall. The protein system was first neutralized by adding Na^+ and Cl^-, and then the system was energy minimized. Next, the system ran a 100 ps MD simulation for equilibrium at

constant 300 K and 1 bar. The position restraint was released to perform a 100 ns simulation. The MD trajectory was saved every 100 ps for future analysis. Following the simulation, the root mean square deviation (RMSD) and radius gyration (Rg) were calculated to confirm the protein structures.

3.3 Results and discussions

3.3.1 Optimisation of HCPs removal by heat-acid precipitation

Densiometric analysis using ImageJ estimated that the bacterial lysis supernatant consists of about 20-30% expressed target proteins. Figure 17, Table 4 and Table 5 exhibit the heat-acid precipitation results. HFns have previously demonstrated excellent thermal and pH stability. In this heat-acid precipitation process, nonetheless, the mutated ferritins demonstrated significantly lower resistance to high temperature and low pH. The purity and yield of HFns under various conditions are rendered in *Figure 17*, and detailed results are presented in Table 2. At this stage, the most crucial perspective was purity, while an acceptable recovery rate was also considered.

The results revealed that C1 was more stable than C2, C3, and C4, and all four proteins were unstable at pH 4.5. C1 had the highest purity of 65.0% at pH 5 and 60 °C with a recovery rate of 15.5%, while C1 had a higher recovery rate of 36.5% at pH 5.5 and 55 °C with a purity of 40.9%. Under the consideration of both purity and recovery rate, 60 °C and pH 5.5 were the most desirable conditions, with 47.5% purity and 34.4% recovery rate. Nonetheless, C2 and C3 were only stable at pH 5.5 and 50 °C. C4 was stable at pH 5.5, 50 °C and 55 °C, with a low purity of 33.2% and 35.2%, respectively. Furthermore, they have a similar recovery rate of 38.0% and 37.2%. As a result, the optimal conditions for C4 were determined to be pH 5.5 and

55 °C. All of the heat-acid precipitation conditions that have been explored are listed in Table 3.

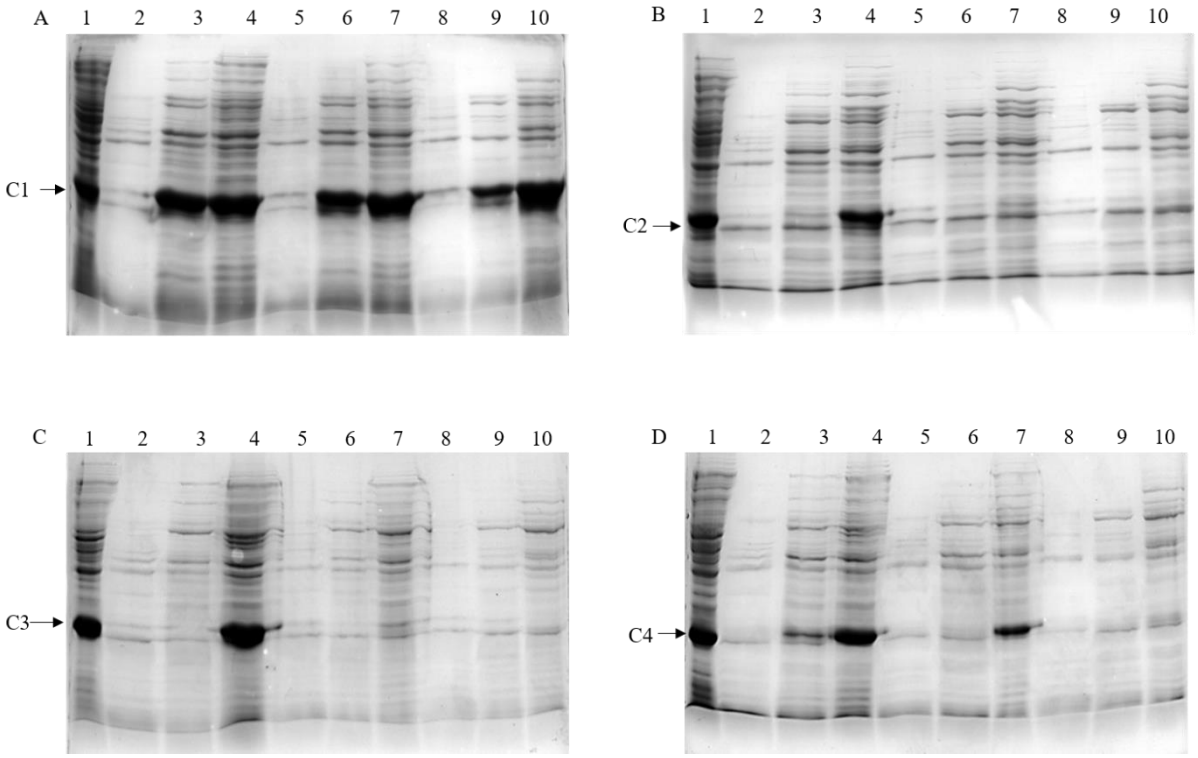

Figure 17 Heat-acid precipitation of mutated HFns. A. C1; B. C2; C. C3; D. C4; L1. Bacteria lysate supernatant; Lane2. 50 ℃, pH 4.5; Lane3. 50 ℃, pH 5; Lane4. 50 ℃, pH 5.5; Lane5. 55 ℃, pH 4.5; Lane6. 55 ℃, pH 5; Lane7. 55 ℃, pH 5.5; Lane8. 60 ℃, pH 4.5; Lane9. 60 ℃, pH 5; Lane10. 60 ℃, pH 5.5.

Table 4. Purity and protein recovery rate after heat-acid precipitation.

Protein	Conditions	50 ℃			55 ℃			60℃		
		pH 4.5	pH 5	pH 5.5	pH 4.5	pH 5	pH 5.5	pH 4.5	pH 5	pH 5.5
C1	Purity (%)	-	42.5	31.0	-	43.1	40.9	-	65.0	47.5
	Recovery rate (%)	-	27.5	36.2	-	21.0	36.5	-	15.5	34.4
C2	Purity (%)	-	19.0	38.8	-	36.3	23.1	-	30.8	21.7
	Recovery rate (%)	-	4.9	32.0	-	7.0	8.7	-	3.5	5.9
C3	Purity (%)	-	-	32.6	-	-	-	-	-	-
	Recovery rate (%)	-	-	38.3	-	-	-	-	-	-
C4	Purity (%)	-	26.9	33.2	-	-	35.2	-	-	-
	Recovery rate (%)	-	14.5	38.0	-	-	37.2	-	-	-

Table 5. Conclusion of heat-acid precipitation optimisation conditions.

Protein	C1	C2	C3	C4
Heat-acid precipitation methods	pH 5.5, 60 ℃ 10 min, 1 M NaCl	pH 5.5, 50 ℃, 10 min, 1 M NaCl	pH 5.5, 50 ℃, 10 min, 1 M NaCl	pH 5.5, 55 ℃, 10 min, 1 M NaCl

As suggested in the results of heat-acid precipitation, all of the mutated FHns were much less stable than native HFns, especially in an acidic environment. Previous experiments implied that insertions at the C-terminus curtailed HFn stability, and mutations at the C-terminus further impacted stability performance. C1 was the most stable of the four mutated HFns, reflecting that removing charge at the C-terminus exerted less impact on protein stability. C4 also surpassed C2 and C3 in thermal performance during this step. However, more experiments are needed to investigate the specific stability performance and mechanism. Due to the gentle purification parameters, a second purification step is required to obtain high-purity samples.

3.3.2 Nucleic acid removal by HIC

HFns can bind to hydrophobic columns, while nucleic acids are extremely hydrophilic and cannot bind to them (Song et al., 2021). Figure 18 illustrates the chromatography graphs in HIC. Figure 19 illuminates the purity of the mutated HFns following HIC. In *Figure 18*, OD 260 was more than twice as high as OD 280 during the flow-through step, and OD 280 was twice as high as OD 260 in the eluted sample. In other words, nucleic acid was removed during the flow through.

According to the SDS-PAGE results, HIC may aid in the removal of some HCPs. Nevertheless, there was no significant change in purity following the HIC step. The same volume of heat-acid precipitation supernatants and HIC samples were loaded into SDS-PAGE. C1 still possessed the highest purity and recovery rate as a result in Figure 19.

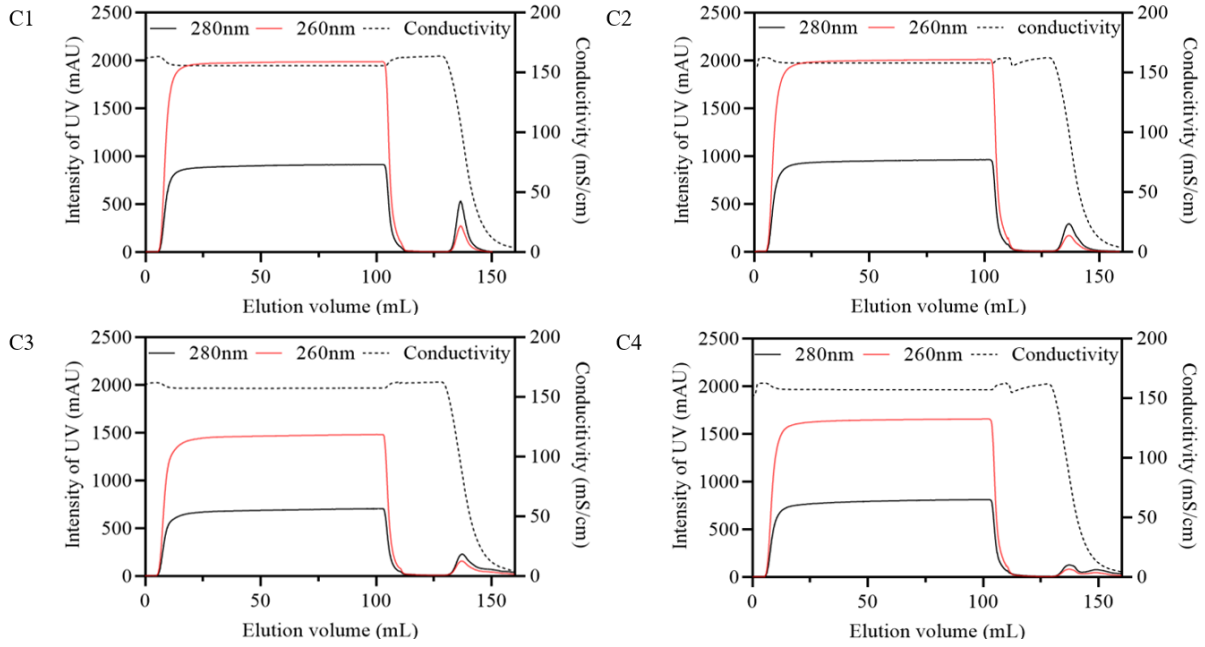

Figure 18. HIC chromatography of C1, C2, C3 and C4.

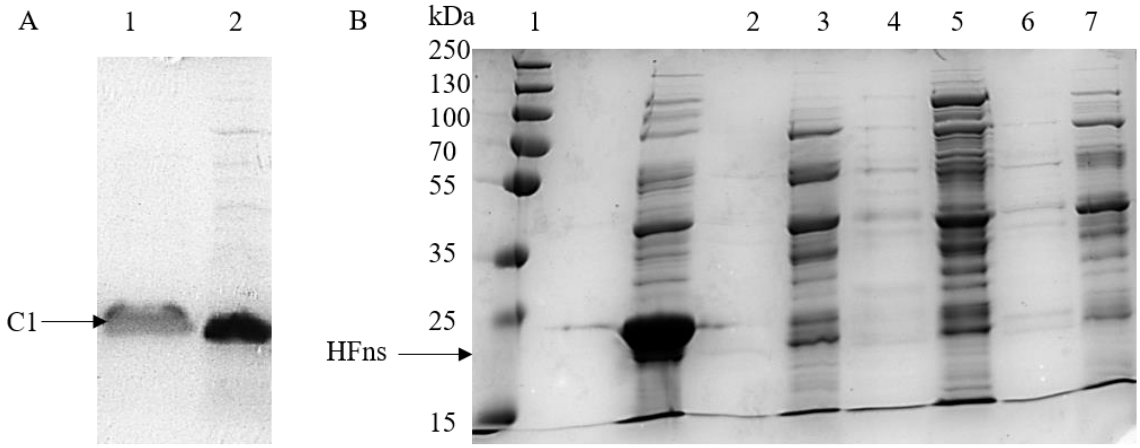

Figure 19. SDS-PAGE of HIC samples. A. HIC samples of C1. Lane1, flow-through, Lane2, HIC peak; B. Lane1, marker; Lane2, C2 flow-through sample; Lane3, C2 HIC peak sample; Lane4, C3 flow-through sample; Lane5, C3 HIC peak sample; Lane6, C4 flow-through sample; Lane7, C4 HIC peak sample.

In conclusion, only C1 had a purity of 76.5 % but was still lower than 90% after two steps of purification, including heat-acid precipitation and HIC with butyl resin. All the other HFns had

purity lower than 30 %. Since this two-step purification process could purify only C1, C2, C3, and C4 would require a further purification step.

3.3.3 Further purification by IEC

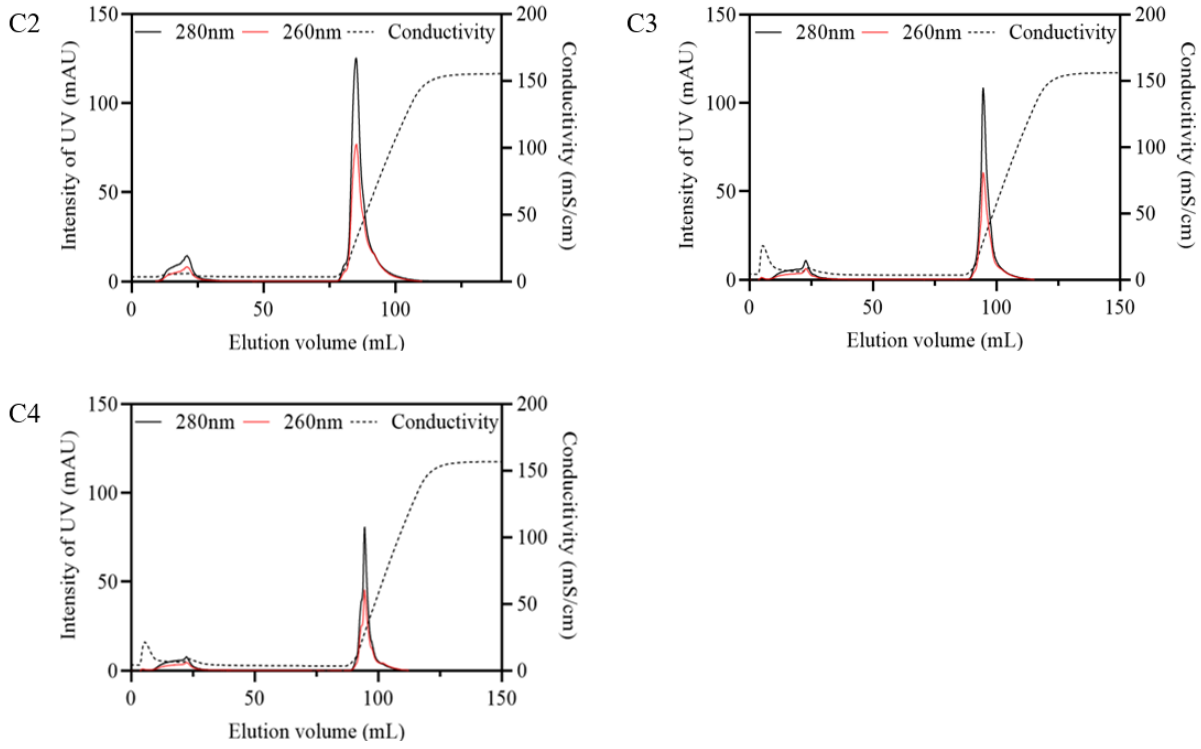

Figure 20. IEC Chromatography of C2, C3 and C4.

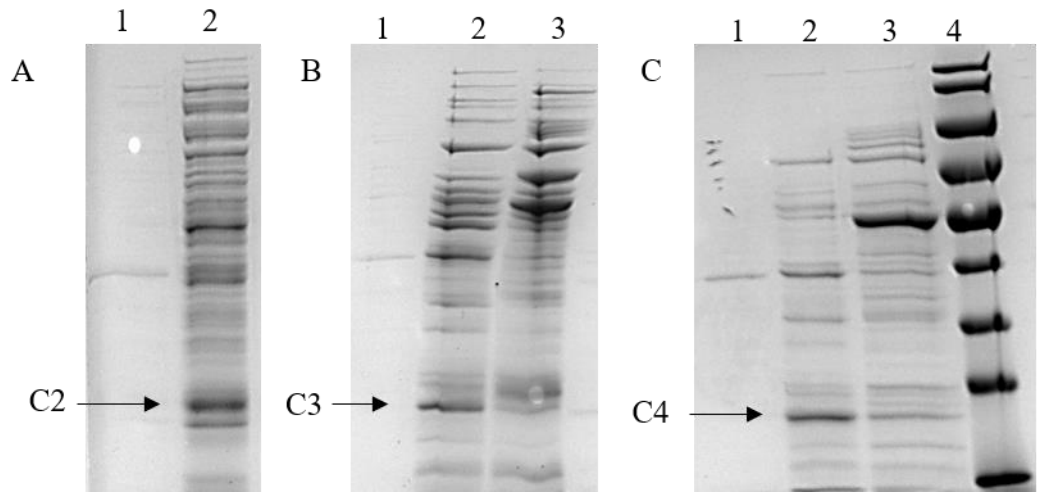

Figure 21 SDS-PAGE of IEC samples. A. SDS-PAGE of C2 collected from IEC, Lane1, flow-through sample, Lane2, peak sample; B. SDS-PAGE of C3 after IEC, Lane1, flow-through, Lane2, Lane3, two peak samples; C. SDS-PAGE of C4 from IEC, Lane1, Flow through, Lane2, Peak 1, Lane3, peak 2, Lane4, marker.

The SDS-PAGE image in Figure 21 unveiled that C2, C3, and C4 still had low purity, and the recovery rate was low. An SEC column polishing process was added, and Superose 6 Increase 10/300 GL (GE Healthcare, US) was adopted. Figure 22 and Table 6 detail the final purity of the mutated HFns. C1 and C2 had a purity of 94.4% and 92.1%, respectively. Nonetheless, the purity of C3 and C4 was 69.4% and 73.7%, respectively, falling short of the required purity of over 90% for the following stability studies. The main reason is their low resistance to high temperatures and low pH, indicating that most HCPs could not be removed during the first heat-acid precipitation step.

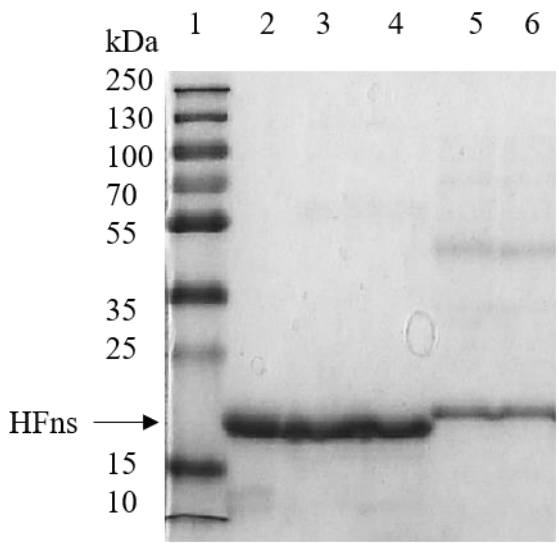

Figure 22. SDS-PAGE of mutated HFns. Lane1, marker; Lane3, C1; Lane4, C2; Lane5, C3; Lane6, C4.

Table 6. Purity of purified proteins estimated by densitometric analysis.

Protein	Purity (%)
C1	94.4
C2	92.1
C3	69.4
C4	73.7

3.3.4 Engineered ferritin's structure confirmation

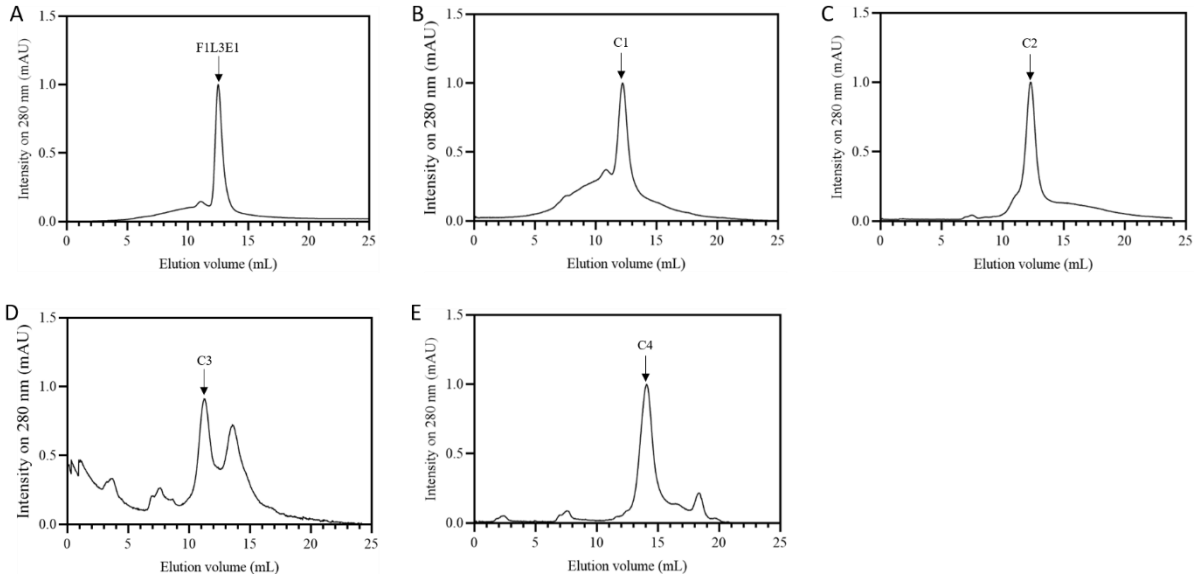

Figure 23. Size-exclusion (SEC, Superose 6) chromatogram for mutated HFns.

Table 7. Structure-related data from SEC-MALS.

Molecule	Hydrodynamic diameter (nm)	SEC elution volume (ml)
$F_1L_3E_1$	13.3	12.5
C1	13.7	12.2
C2	14.6	12.3
C3	12.2	11.3/13.5
C4	-	14.1

F1L3E1 was used as the control protein in the SEC experiment; its hydrodynamic diameter is 13.3 nm, and the elution volume in SEC is 12.5 ml. The mutated C1 and C2 have hydrodynamic diameter of 13.7 nm and 14.6 nm, and elution volumes of 12.2 ml and 12.3 ml, respectively. These findings imply that C1 and C2 self-assembled into a slightly larger global structure, whereas C3 and C4 may fail to assemble.

3.4 Simulation results for structures

Figure 24 exhibits the simulation results. As revealed by examining the final structure of four mutated HFns, they all assembled into the same symmetric hollow spherical structure as native ferritin. However, they were much more unstable than native ferritin. The RMSD of mutated HFns was similar between 0.7 and 0.8 nm, but approximately four times higher than the RMSD of native HFn, which was around 0.15 nm. Higher RMSD normally indicated lower structure stability (Palanisamy et al., 2020). The change margin of C3 and C4 was 0.1 nm greater than C1 and C2, the same as the experiment results. Thus, C3 and C4 are more unstable. Meanwhile, the stable RMSD of the four proteins demonstrated no structure change during the simulation time scale. The Rg results provided information about protein compactness (Lobanov et al., 2008). The increased Rg at 10 ns indicated that the mutated HFns did not fold properly at the start. Additionally, the large fluctuation during the simulation suggested that the structures of mutated HFns were unstable and had a lower rigidness than the control structure. However, the simulation conducted under room temperature and pH 7 did not show a clear stability difference of all mutated HFns.

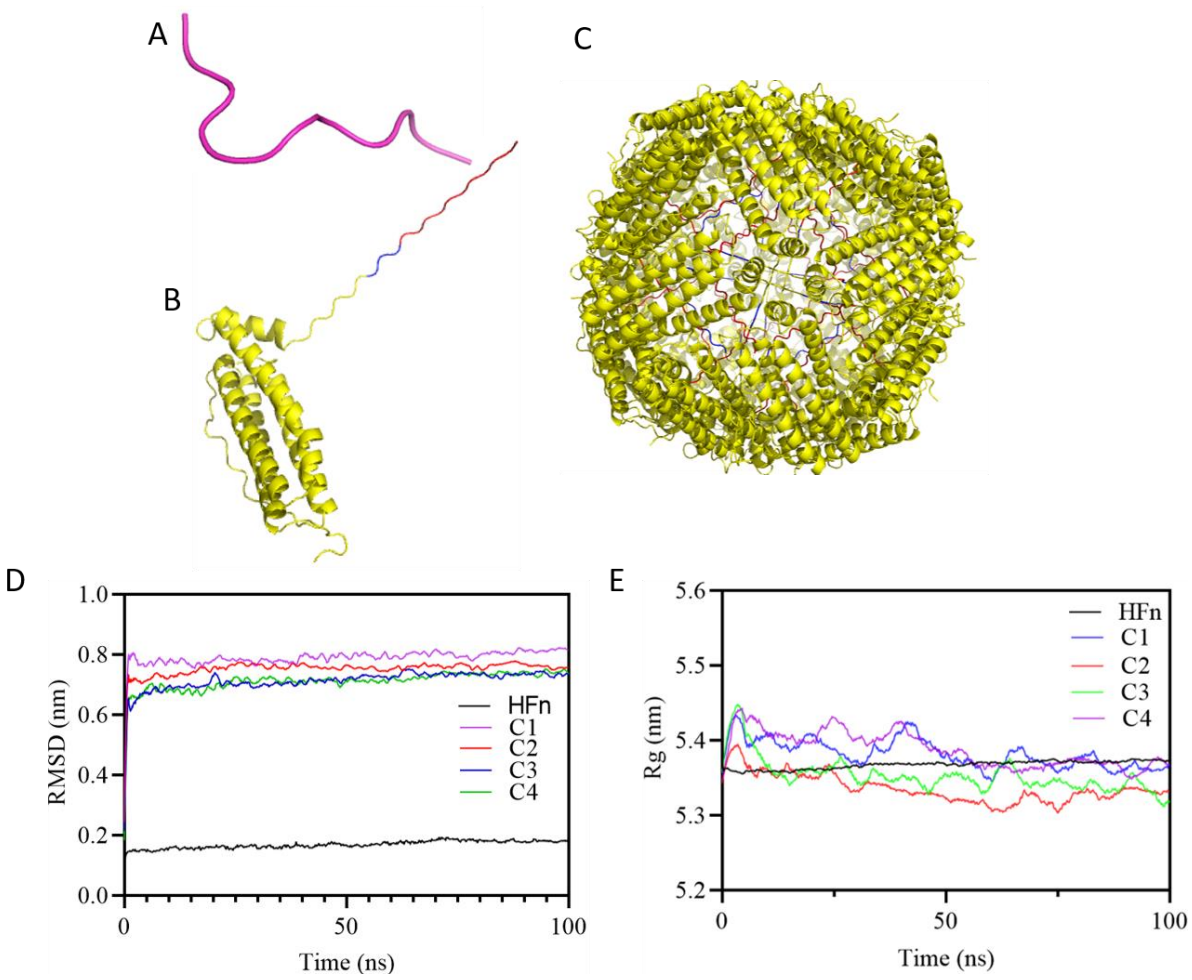

Figure 24. Structure information from 100 ns molecule dynamic simulation delivered by Gromacs. A. Structure of EBNA1 inserted section; B. Assembled ferritin subunit with 3 linkers (blue) and EBNA1 (red); C. Mutated ferritin structure after MD simulation; D. root mean square deviation results of HFns; E. radius of gyration results of HFns.

3.5 Conclusions

According to the purification and simulation results:

1. Purification process including heat-acid precipitation, hydrophobic interaction chromatography was used for C1 and an extra ion exchange chromatography was used for C2. The first step is for HCP removal, and the following chromatography steps are for nucleic acid removal, and further purity improvement.

2. The purification of C3 and C4 failed to reach purity above 90 %. Due to a decrease in the thermal and pH stability of these two proteins during the heat-acid precipitation step compared to other HFns. In HIC and IEC comparison, there was no significant difference in purity.
3. In SEC-MALS results, C1 and C2 formed the same hollow structure as control F1L3E1, which is slightly larger than the native ferritins because of the insertion of the antigens. In addition, C2 had the largest size, due to the repel interaction between positively charged residues on E-helices. However, C34 and C4 failed in assembling. This also explain why C3 and C4 failed to be separated in the designed purification methods.
4. The neutral charged residues of C1 did not have an obvious effect on protein's thermal and pH stability and did not change its structure either. On the contrast, positively charged residues of C2 decrease its thermal stability and made C2 have a looser structure.
5. From simulation results, all the four HFns could self-assemble into the same hollow structure under ambient conditions. This finding is different from the experiment results. The MD simulation can predict the movement of atoms in the time scale based on the model of physical interactions (Karplus and McCammon, 2002). In this study, MD simulation failed to calculate the introduced hydrophobic interaction by the mutations. As the simulations were conducted in a water environment which has different hydrophobic effect than buffers due to the low salt concentration.

3.6 References

Boumaiza, Carmona, F., Poli, M., Asperti, M., Gianoncelli, A., Bertuzzi, M., Ruzzenenti, P., Arosio, P., & Marzouki, M. N. (2017). Production and characterization of functional recombinant hybrid heteropolymers of camel hepcidin and human ferritin H and L chains. Protein Engineering, Design and Selection, 30(2), 77–84. https://doi.org/10.1093/protein/gzw066

Chen. (2012). Bacterial expression systems for recombinant protein production: E. coli and beyond. Biotechnology Advances, 30(5), 1102–1107. https://doi.org/10.1016/j.biotechadv.2011.09.013

Gordon, Horsefield, R., Swarts, H. G. P., de Pont, J. J. H. H. M., Neutze, R., & Snijder, A. (2008). Effective high-throughput overproduction of membrane proteins in Escherichia coli. Protein Expression and Purification, 62(1), 1–8. https://doi.org/10.1016/j.pep.2008.07.005

Karplus, & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. Nature Structural & Molecular Biology, 9(9), 646–652. https://doi.org/10.1038/nsb0902-646

Kim, Rho, Y., Jin, K. S., Ahn, B., Jung, S., Kim, H., & Ree, M. (2011). pH-Dependent Structures of Ferritin and Apoferritin in Solution: Disassembly and Reassembly. Biomacromolecules, 12(5), 1629–1640. https://doi.org/10.1021/bm200026v

Koehn, & Hunt, I. (n.d.). High-Throughput Protein Production (HTPP): A Review of Enabling Technologies to Expedite Protein Production. High Throughput Protein Expression and Purification, 1–18. https://doi.org/10.1007/978-1-59745-196-3_1

Liu, Jin, W., & Theil, E. C. (2003). Opening Protein Pores with Chaotropes Enhances Fe Reduction and Chelation of Fe from the Ferritin Biomineral. Proceedings of the National Academy of Sciences - PNAS, 100(7), 3653–3658. https://doi.org/10.1073/pnas.0636928100

Lobanov, Bogatyreva, N. S., & Galzitskaya, O. V. (2008). Radius of gyration as an indicator of protein structure compactness. Molecular Biology (New York), 42(4), 623–628. https://doi.org/10.1134/S0026893308040195

Middelberg, Rivera-Hernandez, T., Wibowo, N., Lua, L. H. ., Fan, Y., Magor, G., Chang, C., Chuan, Y. P., Good, M. F., & Batzloff, M. R. (2011). A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. Vaccine, 29(41), 7154–7162. https://doi.org/10.1016/j.vaccine.2011.05.075

Palanisamy, Prakash, M., & Rajapandian, V. (2020). Combined DFT and MD simulation studies of protein stability on imidazolium-water (ImHW) clusters with aromatic amino acids. New Journal of Chemistry, 44(41), 17912–17923. https://doi.org/10.1039/d0nj03085f

Palomares, L. A., Estrada-Mondaca, S., & Ramírez, O. T. (2004). Production of recombinant proteins: challenges and solutions. Methods in molecular biology (Clifton, N.J.), 267, 15-52.

Qu, Y., Wang, L., Yin, S., Zhang, B., Jiao, Y., Sun, Y., Middelberg, A., & Bi, J. (2021). Stability of Engineered Ferritin Nanovaccines Investigated by Combined Molecular Simulation and Experiments. The journal of physical chemistry. B, 125(15), 3830-3842. https://doi.org/10.1021/acs.jpcb.1c00276

Qu, Davey, K., Sun, Y., Middelberg, A., & Bi, J. (2022). Engineered Design of the E-Helix Structure on Ferritin Nanoparticles. ACS Applied Bio Materials, 5(7), 3167–3179. https://doi.org/10.1021/acsabm.2c00154

Song, X., Zheng, Y., Zhu, L., Zhang, L., Meng, H., Yu, R., & Zhang, C. (2021). Development of robust and facile purification process for the production of recombinant human ferritin heavy chain nanoparticle from Escherichia coli. Process biochemistry (1991), 104, 1-9. https://doi.org/10.1016/j.procbio.2021.02.014

Wang, W., & Roberts, C. J. (2010). Aggregation of Therapeutic Proteins (1. Aufl. ed.). Wiley.

Yano, Y. F. (2012). Kinetics of protein unfolding at interfaces. Journal of physics. Condensed matter, 24(50), 503101-503101. https://doi.org/10.1088/0953-8984/24/50/503101

Weidenbacher, P., Musunuri, S., Powell, A. E., Tang, S., Do, J., Sanyal, M., & Kim, P. S. (2022). Simplified Purification of Glycoprotein-Modified Ferritin Nanoparticles for Vaccine Development. Biochemistry. https://doi.org/10.1021/acs.biochem.2c00241

Yin. (2021). Engineered human heavy-chain ferritin to improve anti-tumour drug delivery performance.

Zou, W., Liu, X., Zhao, X., Wang, J., Chen, D., Li, J., Ji, L., & Hua, Z. (2016). Expression, purification, and characterization of recombinant human L-chain ferritin. Protein expression and purification, 119, 63-68. https://doi.org/10.1016/j.pep.2015.11.018

Chapter 4: Stability study of engineered ferritins by experiments and simulations

4.1 Introduction

Human heavy-chain ferritins have highly symmetric hollow structures with high thermal and pH stability (Liu & Theil, 2003; Kim et al, 2011). The structural evolution of ferritin-related drug functions is investigated by various teams based on these ferritin specificities. Protein products with higher stability, efficacy, and safety are always preferred. As a result, ferritin gene and surface modification are frequently adopted in drug development to enhance stability. The ability to reveal the impact of various mutations on ferritin performance is critical to the development of new drugs for specific needs.

This project stresses the C-terminus mutation because C-terminus insertion has demonstrated greater stability than N-terminus insertion (Yiran, Qu, 2021). In Chapter 3, the four genetically mutated ferritins demonstrated distinct properties during the purification process. Nonetheless, how can this be explained?

Subsequently, a series of experiments were performed to determine the thermal and pH properties of the ferritins, and MD simulation was utilized to compare the stability at the atomic level. SASA was calculated for the MD simulation to estimate the potential hydrophobicity. The root mean square deviation (RMSD), the radius of gyration (Rg), root mean square fluctuation (RMSF), and H-bond were employed to reveal structural changes on the simulation time scale.

The SDS-PAGE results in Chapter 3 implied that C3 and C4 did not achieve purity of more than 90% using any of the available purification methods. Owing to their low purity, the experiment results were not validated. Therefore, they were excluded from the subsequent studies.

4.2 Materials and Methods

4.2.1 Hydrophobic studies of engineered ferritins by hydrophobic chromatography and simulations

A HiTrap Butyl FF column was adopted to conduct a hydrophobic analysis on C1 and C2 proteins (GE Healthcare, USA). The binding strength was indicated by the elution volume at room temperature. The larger the elution volume, the stronger the binding between ferritin and column resin, and the higher the hydrophobicity of ferritin.

The elution buffer was 20 mM phosphate buffer (PB, with Na_2HPO_4 and NaH_2PO_4) at pH 6.5, and the equilibration buffer was 100 mM PB, 1 M ammonium sulphate, and pH 6.5. C1 and C2 were diluted to 1 mg/ml in 2 ml of elution buffer each. Then, diluted samples were mixed with 8 ml of equilibrium buffer. After the same treatment, an F1L3E1 sample was used as a control. Besides, 1 M sodium hydroxide was used to clean the column and prepare for the next sample. A prepackaged 5 ml HiTrap Butyl FF column was acquired from GE Healthcare, and the data were analyzed with GraphPad Prism 9.

All of the experiments were performed on a GE Healthcare AKTA Go. First, the column was brought to equilibrium using 3-column volumes (CV) of equilibration buffer. Then, a 10 ml control sample (F1L3E1) was loaded into the equilibrated column at 1 ml/min for the investigation. Afterwards, the sample was linearly eluted by elution buffer for 10 CV.

For the following samples, the column with 1 M sodium hydroxide was cleaned, the column was subjected to re-equilibrium, and the previous column operation was repeated.

4.2.2 Determination of thermal and pH stability of engineered ferritins by fluorescence spectroscopy

The ultraviolet absorption wavelength of ferritin was 280 nm, which was typically used to identify ferritin, and the emission wavelength of ferritin was primarily between 300 and 350

nm (Zhang et al., 2021). Since a low-concentration phosphate buffer (Na_2HPO_4 and NaH_2PO_4) was recommended to reduce the buffer's influence (Kessenbrock & Groth, 2017), a 20 mM phosphate buffer was selected.

Engineered ferritin C1 was desalted in a 20 mM pH 7 phosphate buffer (PB buffer, 0.2 M Na_2HPO_4 and 0.2 M NaH_2PO_4). Concerning the following experiment, protein concentration of the samples were adjusted to 0.1 mg/ml with a sample volume of 3 ml using the Bradford assay with BSA as standard. They were exposed to various temperatures for determining the ferritins' thermal stability. Each sample was heated to 40, 50, 60, 70, 80, 90, and 100 °C for 10 minutes, followed by 30 minutes at room temperature. Each sample was scanned three times. Besides, a series of 20 mM PB buffers (Na_2HPO_4 and NaH_2PO_4) were prepared with pH values of 4, 4.5, 5, 5.5, 6, 6.5, and 7. The protein samples were diluted to 0.1 mg/ml using the different pH PB buffers to a total volume of 3 ml, and each sample was scanned three times after 30 minutes.

The fluorescence spectrophotometer (RF-5301PC SHIMADZU, Japan)) was used in the experiment. The path length of the cuvette was 1 cm, the excitation wavelength was 280 nm, the emission wavelength was 300-600 nm, the slit width of excitation and emission was 5 nm, and the scan speed was set as a high speed.

4.2.3 Stability performance and dissemble mechanism by MD simulation

Four mutated ferritins were simulated at pH 7 and pH 4 at room temperature for pH comparisons. Three temperatures (70, 100, and 200 °C) were used for thermal studies. The Gromacs 2018 software package was utilized for all simulations. AMBER99SB force field and SPC/E water mode were applied. The protein complex was placed in a cubic unit cell with a periodic distance of 1 nm in all directions, and the resulting box size was 15.7 nm. Subsequently, the system was neutralized by adding Na^+ and Cl^-. The system was equilibrated

for 100 ps at a temperature of 300 K and a pressure of 1 bar. The system was prepared to release position restraints and execute a 200ns MD simulation for data collection (Lemkul, 2019). Next, the proteins were visualized with PyMOL and VMD, and Gromacs was then employed to analyze the data.

4.3 Results and Discussion

4.3.1 Hydrophobic studies by using HIC

The alternation position was determined based on the simulation's "hot spot" findings, and significant changes in ferritin stability were expected (Reichmann et al., 2007). However, the results in Figure 25 implied that the hydrophobicity of the two engineered ferritins (C1 and C2) was similar to that of the control (F1L3E1). The elution volumes of all three proteins were listed in Table 8. According to the elution volume, C2 possessed the highest hydrophobicity and eluted at 36.0 ml, while C1 was eluted at 35.3 ml, which was close to the control (F1L3E1) volume of 34.0 ml.

As calculated for the solvent accessible surface area (SASA), C3 had the smallest size of 2018.7 nm^2, and C4 had the highest area of 2065.3 nm^2. The SASA of F1L3E1 was smaller than C1 and C2.

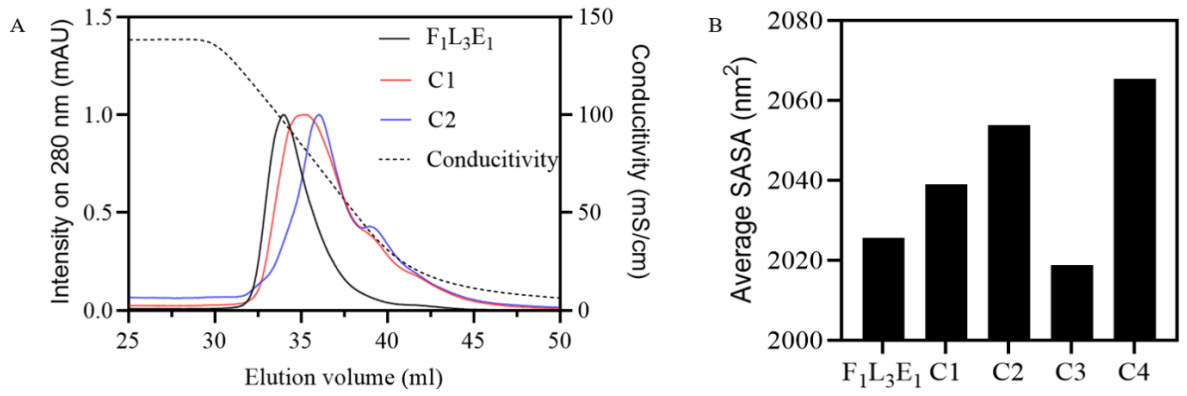

Figure 25. A. hydrophobic interaction chromatography of three HFns; B. solvent accessible surface area (SASA) of all five HFns calculated in Gromacs.

Table 8. Detailed results of hydrophobicity studies.

Molecule	**HIC elution Volume (ml)**	**Calculated SASA (nm^2)**
$F_1L_3E_1$	34.0	2025.5
C1	35.3	2038.9
C2	36.0	2053.8
C3		2018.7
C4	-	2065.3

The mutation on the protein's solution-accessible outer surface exerted a greater effect on hydrophobicity than the inner hydrophobic core (Calhoun, et al., 2003). After the self-assembly of HFns, the E-helices would fold inward. As a result, all mutated amino acids were entrapped within the globe and had a negligible effect on the hydrophobicity of the entire protein. The previous study also suggested that the C-terminus residues of ferritin had a negligible effect on hydrophobicity compared to the N-terminus residues (Qu et al., 2021).

The calculated theory of solvent accessible surface area (SASA) confirmed this phenomenon from a different perspective. As C2 had the greatest SASA of the three HFns (F1L3E1, C1, and C2), it demonstrated the greatest hydrophobicity in experiments. Mutated Arg 167 (Glu) and Arg 171 (Asp) are positively charged at pH 6.5, leading to a larger diameter of C2 than F1L3E1 and C1 due to the electrostatic interactions between these amino acids (Dill & Maccallum, 2012). The looser structure may contribute to the larger size of SASA. To prove that C2 has a larger structure, additional research is required.

Hydrophobic interaction is one of the most imperative forces that assemble ferritins into their global structure (Camilloni et al., 2016; Reichmann et al., 2007). Val and Ile are the most hydrophobic amino acids found in C3. Consequently, C3 tends to be more compactable in the

water system. In contrast, Gln and Asn are highly hydrophilic, resulting in a looser structure for C4.

4.3.2 Thermal and pH Stability study by fluorescence results and MD simulations

C1 has a structure change over 60 °C and is denatured at 100 °C, which is lower than F1L3E1's denature temperature of 63.6 °C and higher than the total denature temperature of 84.6 °C. C1 starts to lose its tertiary structure at pH 4.5, which is lower than $F_1L_3E_1$ at pH 5. (Qu et al., 2021).

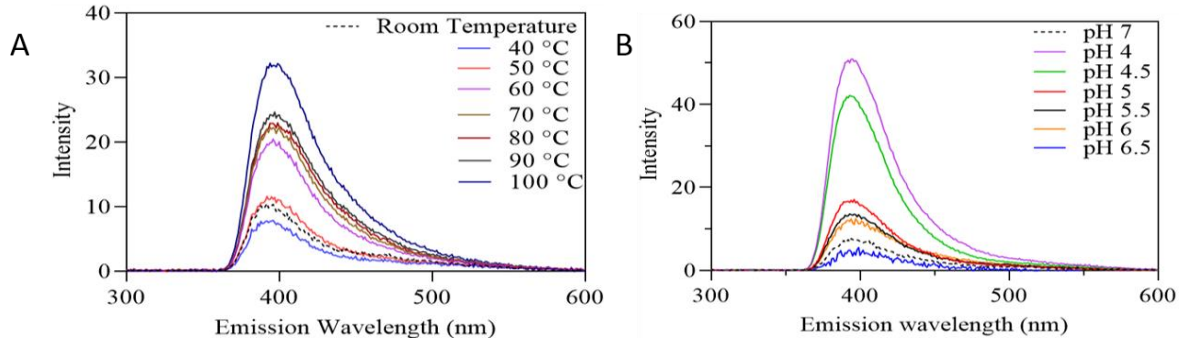

Figure 26. Fluorescence results of thermal (A) and pH (B) stability of C1.

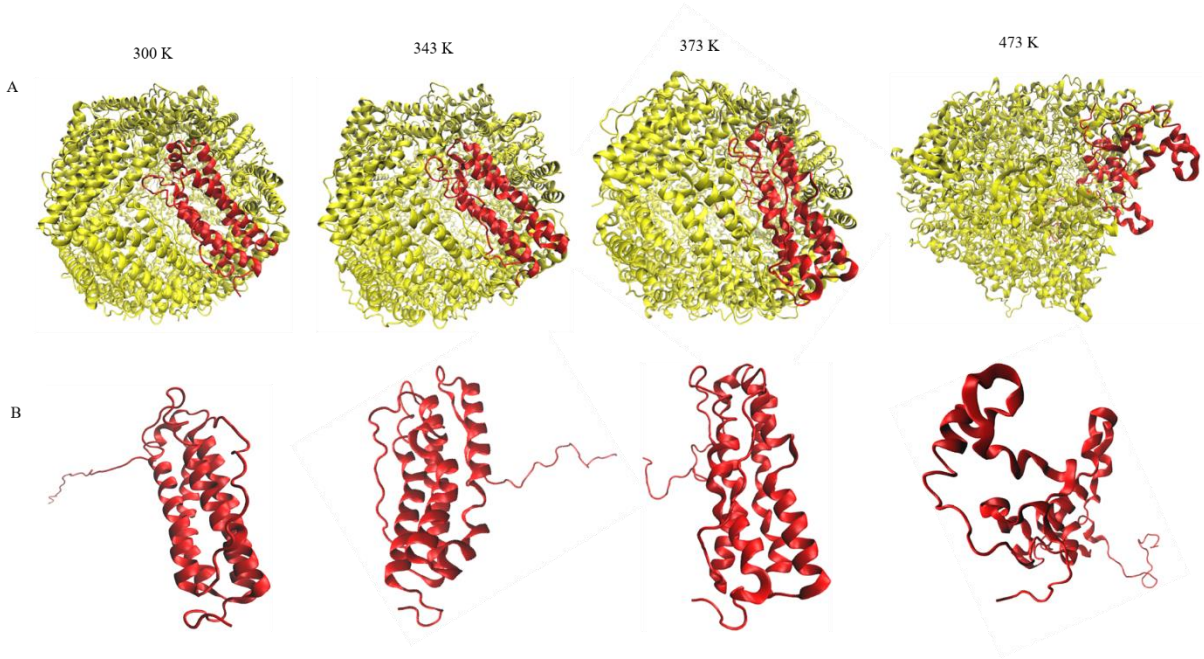

Figure 27. Structure of C1 at different temperatures. A. The whole protein structure is in yellow, and the subunit is studied in the following calculations in red; B. Structure of one single subunit at different temperatures. The structures were visualised and generated by VMD.

Figure 27 demonstrated that the tertiary structure of the protein changed after 100 ns of simulation at different temperatures, and the secondary structure was lost at 473 K. Between 300 K and 473 K, the subunit's structure altered, and the number of α-helices gradually decreased. The following calculations illustrate the quantitative outcomes.

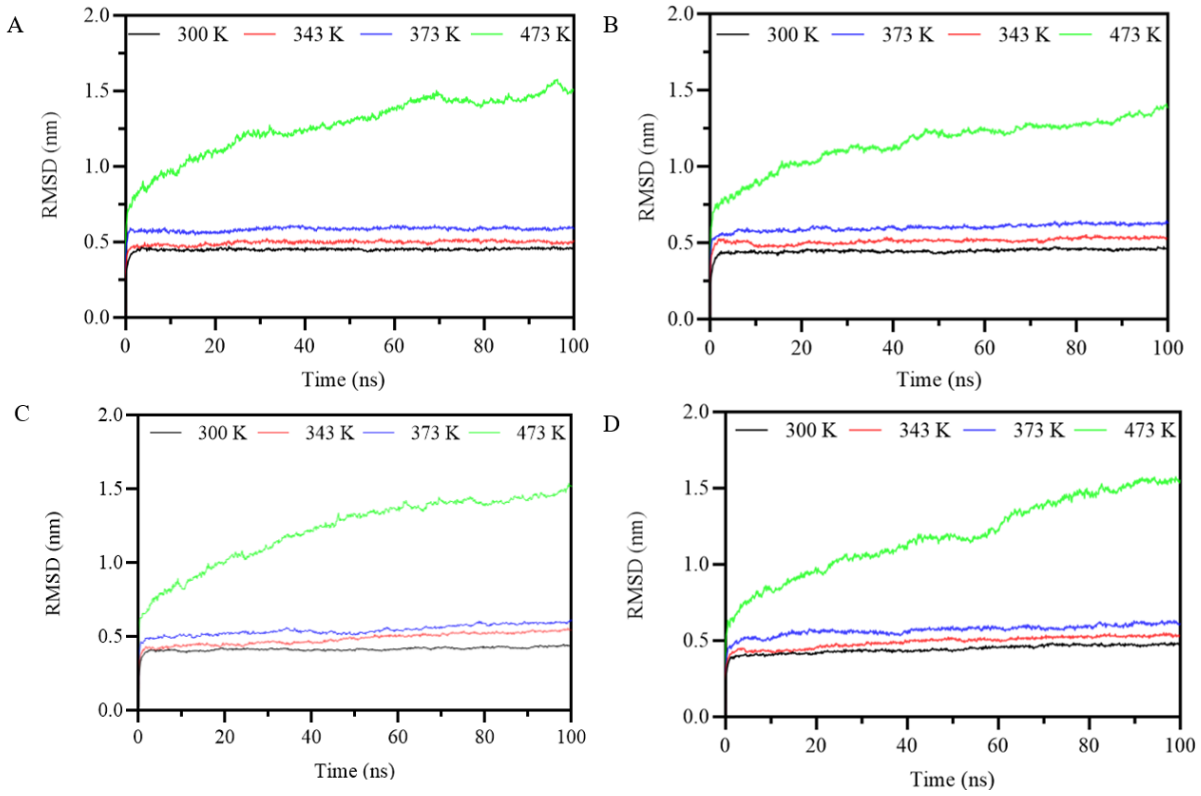

Figure 28. RMSD of C1 (A), C2 (B), C3 (C) and C4 (D) over the temperature range from 300 K to 473 K.

The root mean square deviation (RMSD) of the alpha carbon atoms of the entire protein (Figure 28) was measured to determine the protein's overall structural stability. RMSD of all HFns began at approximately 0.45 nm at 300 K and increased to approximately 0.5 nm at 343 K and 0.6 nm at 373 K, verifying that the structures were more unstable at higher temperatures. A significant increase was observed at 473 K, at which the system was destroyed.

However, there was no discernible change in RMSD between 343 K and 373 K. Consequently, the destruction of the structure cannot be concluded. These RMSD results may have been affected by the simulation's time limitations. As stated in Chapter 2, a higher temperature accelerated the simulation speed. Hence, a longer simulation would be preferable to allow this complex system to reach its evaluation threshold.

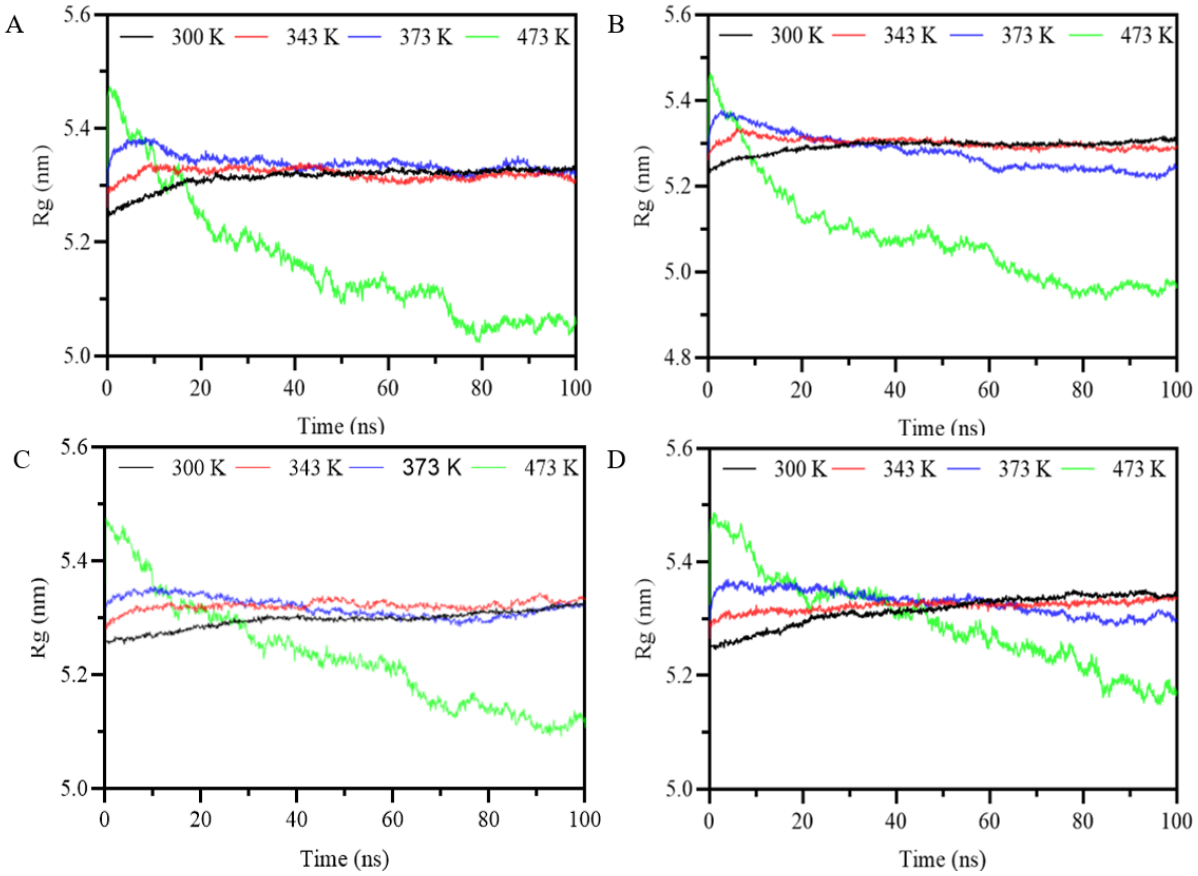

Figure 29. Rg of C1 (A), C2 (B), C3 (C) and C4 (D) over the temperature range from 300 K to 473 K.

Since the radius of gyration (Rg) (Figure 29) indicates the compactness of the protein, a smaller variation in Rg during the simulation may demonstrate a more stable and rigid structure. At 473 K, the Rg of all HFns decreased from 5.5 nm to approximately 5 nm, confirming the claim. At 373 K, only C2, C3, and C4 fluctuated by approximately 0.1 nm during the 100 ns simulation.

The variation in Rg for C2, C3, and C4 implied that they were less stable than C1. The same conclusion can be drawn from the purification processes. Thus, the positive charge and hydrophobicity mutation at E-helices exerted a substantial effect on the thermal stability of HFNs.

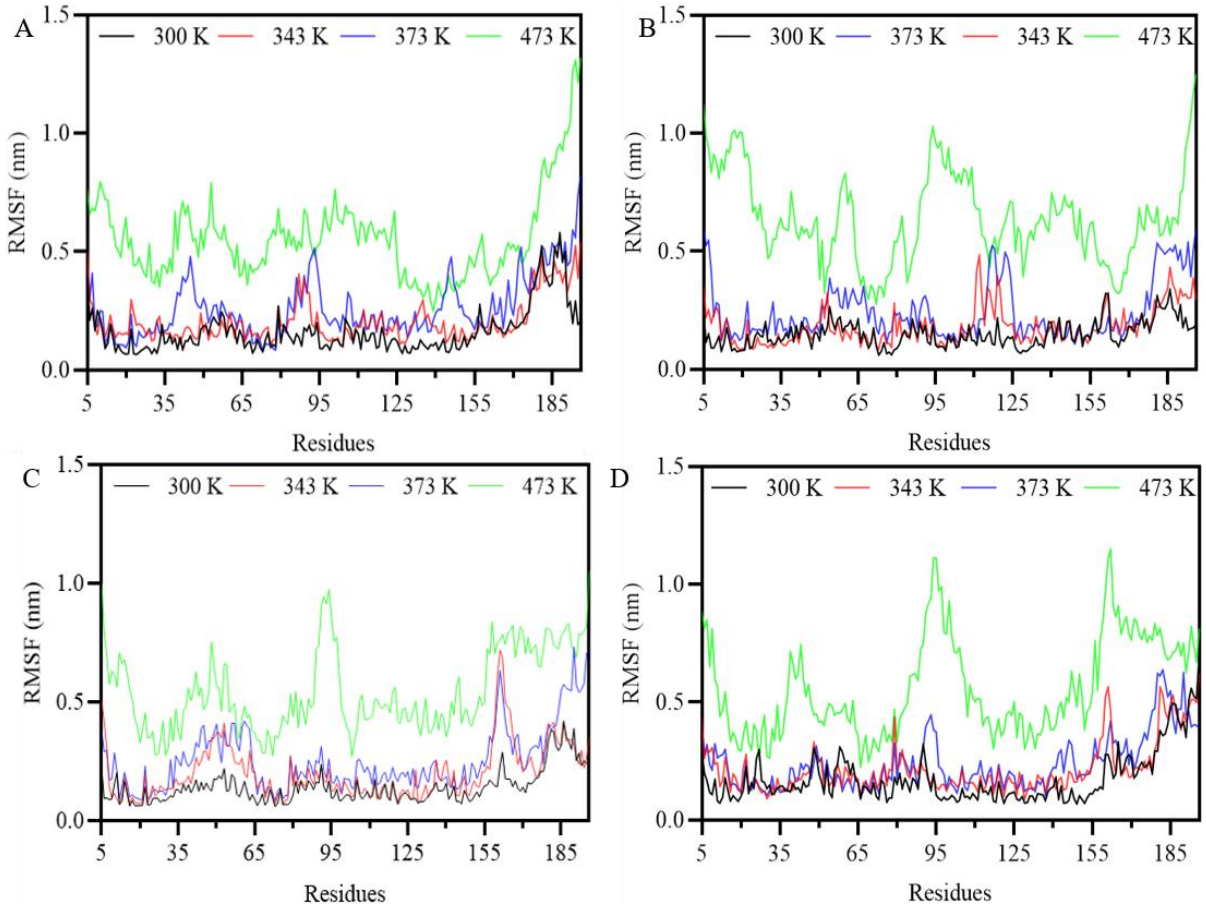

Figure 30. RMSF of all residues of monomers of C1 (A), C2 (B), C3 (C) and C4 (D) over the temperature range from 300 K to 473 K.

The root mean square fluctuation (RMSF) (Figure 30) reflected the change in the position of each residue. The larger the value of RMSF, the more significant the structure fluctuation. This principle can be used to determine the fluctuating region of each subunit (Islam et al., 2020). 473 K was a very high temperature, at which proteins had the greatest RMSF. Otherwise, the RMSF of all mutated HFns increased at 343 K and 373 K, with the higher temperature exhibiting a higher RMSF. Although the subunit contained flexible residues in all sections, the final 160-196 residues were more flexible than other sections, including mutated E-Helices and inserted antigens. Additionally, C3 and C4 had a greater variety of approximately 0.5 nm between residues 155 and 170, consisting of the mutated amino acid.

This result establishes because the hydrophobic (Val 165, 169 and Ile 166, 168, 170) and hydrophilic (Gln 165, 168, 170 and Asn 166, 169) amino acids interact more hydrophobically with the surrounding water environment.

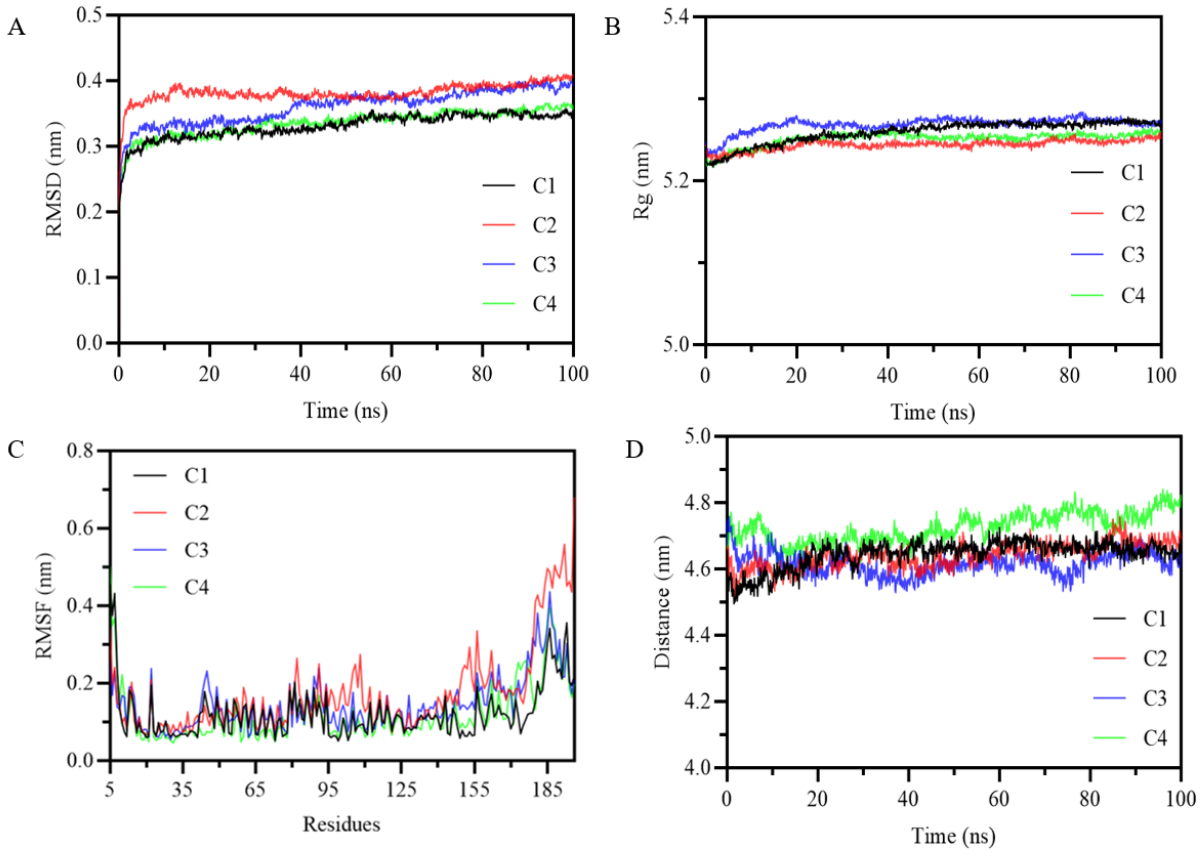

Figure 31. RMSD (A), Rg (B), RMSF (C) and the distance between the geometric centre of a monomer and the geometric centre of the protein (D) of C1, C2, C3 and C4 at pH 4.

The pH experiments were conducted at pH 4 (Figure 31), which was the lowest dissemble pH determined by the experiments. Nonetheless, this seems not low enough for the simulation analysis, especially given the short duration of the simulation (Li et al., 2022).

After 40 nanoseconds, the RMSD of C3 increased from 0.3 nm to 0.4 nm, indicating that C3 was more unstable in an acidic environment than other molecules. Based on the Rg results, C3 fluctuated between 0 and 30 ns. RMSF results uncovered that the mutated E-helices of C2 were

more flexible at lower pH, with an RMSF of 0.3 nm. One reason could be that Arg 167 and 171 were protonated and induced the segment to aggregate at low pH (Wyrzykowski et al., 2020).

4.4 Conclusion

According to the results of the stability study experiments and simulation:

1. All mutations at E-helices on the C-terminus affect the stability of ferritin. C1, which changed to amino acids with no charge, had a minimal effect and exhibited the highest hydrophobic, thermal, and pH stability among all the mutated HFns, followed by C2, which contained an amino acid with a positive charge. The introduced positively charged amino acid boosted the electrostatic interaction as the E-helices tended to repel each other and changed the compactness of assembled structure.
2. In contrast, C3 and C4, which were designed by substituting hydrophobic or hydrophilic amino acids, demonstrated the lowest stability in all of the experiments and simulations. The increase and decrease in hydrophobicity of E-helices significantly affected the self-assembly structure of C3 and C4 to form the smallest and largest structures accordingly. This structure change influenced the forming of the H-bind and van der Waals interaction. As a result, these two proteins failed to assemble into the correct structures.
3. The fluorescence results revealed that C1 was supposed to start denaturing over 60 °C and pH 4.5. However, the simulation results failed to support these findings. No structure disassembly was observed from the simulations due to the limitation of simulation results.

4.5 References

Camilloni, C., Bonetti, D., Morrone, A., Giri, R., Dobson, C. M., Brunori, M., Gianni, S., & Vendruscolo, M. (2016). Towards a structural biology of the hydrophobic effect in protein folding. Scientific reports, 6(1), 28285-28285. https://doi.org/10.1038/srep28285

Islam, R., Parves, M. R., Paul, A. S., Uddin, N., Rahman, M. S., Mamun, A. A., Hossain, M. N., Ali, M. A., & Halim, M. A. (2020). A molecular modeling approach to identify effective antiviral phytochemicals against the main protease of SARS-CoV-2. Journal of Biomolecular Structure and Dynamics, 1-12. https://doi.org/10.1080/07391102.2020.1761883

DILL, KA & MACCALLUM, JL 2012, 'The Protein-Folding Problem, 50 Years On', Science (American Association for the Advancement of Science), vol. 338, no. 6110, pp. 1042–1046.

Kessenbrock, M., & Groth, G. (2017). Circular Dichroism and Fluorescence Spectroscopy to Study Protein Structure and Protein–Protein Interactions in Ethylene Signaling. In Ethylene Signaling (Vol. Vol.1573, pp. 141–159).

Kim, M, Rho, Y, Jin, KS, Ahn, B, Jung, S, Kim, H & Ree, M 2011, 'pH-Dependent Structures of Ferritin and Apoferritin in Solution: Disassembly and Reassembly', Biomacromolecules, vol. 12, no. 5, pp. 1629–1640.

Lemkul, J. (2019). From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package [Article v1.0]. Living Journal of Computational Molecular Science, 1(1). https://doi.org/10.33011/livecoms.1.1.5068

Li, Z., Maity, B., Hishikawa, Y., Ueno, T., & Lu, D. (2022). Importance of the Subunit–Subunit Interface in Ferritin Disassembly: A Molecular Dynamics Study. Langmuir, 38(3), 1106-1113. https://doi.org/10.1021/acs.langmuir.1c02753

Liu, X, Jin, W & Theil, EC 2003, 'Opening Protein Pores with Chaotropes Enhances Fe Reduction and Chelation of Fe from the Ferritin Biomineral', Proceedings of the National Academy of Sciences - PNAS, vol. 100, no. 7, pp. 3653–3658.

Qu, Y., Wang, L., Yin, S., Zhang, B., Jiao, Y., Sun, Y., Middelberg, A., & Bi, J. (2021). Stability of Engineered Ferritin Nanovaccines Investigated by Combined Molecular Simulation and Experiments. The journal of physical chemistry. B, 125(15), 3830-3842. https://doi.org/10.1021/acs.jpcb.1c00276

Reichmann, D., Rahat, O., Cohen, M., Neuvirth, H., & Schreiber, G. (2007). The molecular architecture of protein–protein binding sites. Current opinion in structural biology, 17(1), 67-76. https://doi.org/10.1016/j.sbi.2007.01.004

Swift, J., Wehbi, W. A., Kelly, B. D., Stowell, X. F., Saven, J. G., & Dmochowski, I. J. (2006). Design of Functional Ferritin-Like Proteins with Hydrophobic Cavities. Journal of the American Chemical Society, 128(20), 6611-6619. https://doi.org/10.1021/ja057069x

Wyrzykowski, D., Pilarski, B., Chmurzyński, L., & Makowska, J. (2020). Acidic-basic properties of arginine-rich peptide fragments derived from the human Pin1 protein. Journal of molecular liquids, 312, 113379. https://doi.org/10.1016/j.molliq.2020.113379

Zhang, S., Li, Y., Bao, Z., Sun, N., & Lin, S. (2021). Internal cavity amplification of shell-like ferritin regulated with the change of the secondary and tertiary structure induced by PEF technology. International Journal of Biological Macromolecules, Vol.182, 849-857.

Chapter 5: Conclusions and future development

5.1 Conclusions

The main aims of this project are to use the experimental methods to study the stability of engineered ferritins which have mutations on E-helices at the C-terminus and to introduce MD simulation to this study to examine the potential of using MD simulation for protein-related research. And explore the underlying mechanism of the performance of these mutated ferritins. In general, this project tried to understand the relationship between various E-helices mutations and proteins' stability.

The key findings of this project are:

1. Protein C1, which changed to amino acids with no charge and C2, which contained amino acids with positive charge. could be purified with the developed methods, C3 and C4 could however not be purified to high purity levels. C1 could achieve a very high purity after heat-acid precipitation and HIC, while C2 need an extra IEC step to achieve the required purity. At the end of the purification processes, an SEC step was included for all four proteins to further polish the samples. The final purity of C1 was 94.4 %, C2 was 92.1 %, C3 was 69.4 % and C4 was 73.3 %.
2. The fluorescence results showed that the mutation of C1 did not affect its thermal resistance significantly and increased its pH stability. And in all the simulation results, C1 had shown a more stable structure than others. Both the HIC experiment and calculated SASA for simulations indicated that C1 is higher hydrophobic than F1L3E1, but the difference is not significant. These suggested that the neutrally charged amino acids at E-helices changed the electrostatic interactions however, they didn't introduce new forces, thus, had less effect on ferritins assemble and stability.

3. The thermal and pH studies of C2, C3 and C4 were conducted by MD simulations. C3 was the most unstable, followed by C4 and C2.
4. C2 had a larger structure, and lower thermal and pH stability than C1. Because the introduced positive charged amino acids not only destroy the original electrostatic bindings, but also increased the repel force.
5. C3 and C4 altered the hydrophobic fourfold symmetry, so they showed more fluctuations in the simulations and suggested the lowest stability. This was the same as the findings of purification processes. The SEC results suggested that C3 and C4 failed in assembling, however, the final structure of simulation showed same structures as C! and C2.
6. C4 had the largest SASA in simulation, which suggested a looser structure. In the contrast, the SEC-MALS indicated that C4 is the smallest protein even smaller than the control. This could be because C4 failed to assemble properly. Hydrophobic interaction between E-helices is the main force to push ferritin self-assembly, as a result, the change of hydrophobicity effect or the assemble and the stability.

However, more experiments are required to provide detailed results of the thermal and pH stability, and a longer simulation time is needed to allow the whole simulation to achieve the critical structures for the following analysing. In general, simulation is a very useful tool to study proteins, which potentially shortens the research period and reduces the experiment expense.

5.2 Future development of this project

1. The recovery rate of C2 was very low due to the introduction of IEC processes, other purification methods should be considered for C2, C3 and C4 such as IMAC. Because in current purification methods, most of the HCPs are expected to be removed by heat-

acid precipitation. But C2, C3 and C4 could not resist high temperature and low pH, as a result, the HCPs were not isolated effectively.

2. A series of pH conditions should be tested in simulations. The current simulation was only carried out at pH 7 and pH 4, therefore, pH 5 a middle number and pH 2 an extreme environment should be conducted to compare the structure changes.

Appendix A. Amino acid sequence of engineered ferritins

Protein	Full amino acid sequence
C1	TTASTSQVRQNYHQDSEAAINRQI NLELYASYVYLSMSYYFDRDDVA LKNFAKYFLHQSHEEREHAEKLM KLQNQRGGRIFLQDI KKPDCDDWESGLNAMECALHLEK NVNQSLLELHKLATDKNDPHLCD FIETHYLNEQVKAIKELGDHVTNL RKMGAPESGLAQYLFQKHTLGDS DNESGGSHPVGEADYFEY
C2	TTASTSQVRQNYHQDSEAAINRQI NLELYASYVYLSMSYYFDRDDVA LKNFAKYFLHQSHEEREHAEKLM KLQNQRGGRIFLQDI KKPDCDDWESGLNAMECALHLEK NVNQSLLELHKLATDKNDPHLCD FIETHYLNEQVKAIKELGDHVTNL

	RKMGAPESGLARYLFRKHTLGDS DNESGGSHPVGEADYFEY
C3	TTASTSQVRQNYHQDSEAAINRQI NLELYASYVYLSMSYYFDRDDVA LKNFAKYFLHQSHEEREHAEKLM KLQNQRGGRIFLQDI KKPDCDDWESGLNAMECALHLEK NVNQSLLELHKLATDKNDPHLCD FIETHYLNEQVKAIKELGDHVTNL RKMGAPESGVIEIVIDKHTLGDSD NESGGSHPVGEADYFEY
C4	TTASTSQVRQNYHQDSEAAINRQI NLELYASYVYLSMSYYFDRDDVA LKNFAKYFLHQSHEEREHAEKLM KLQNQRGGRIFLQDI KKPDCDDWESGLNAMECALHLEK NVNQSLLELHKLATDKNDPHLCD FIETHYLNEQVKAIKELGDHVTNL

	RKMGAPESGQNEQNQDKHTLGDS DNESGGSHPVGEADYFEY

Appendix B. Standard curves for Bradford Assay

Figure 32 and Figure 33 present the two standard calibration curves for Bradford Assay that was used to measure protein concentration on two UV-VIS equipment in this project. According to the instruction of Bradford Assay, bovine serum albumin (BSA) with the concentration of 0.2 to 0.9 mg/ml were prepared. A linear trend was generated, in which y is the absorbance at 595 nm from UV (AU) and x is the concentration of the protein (mg/ml).

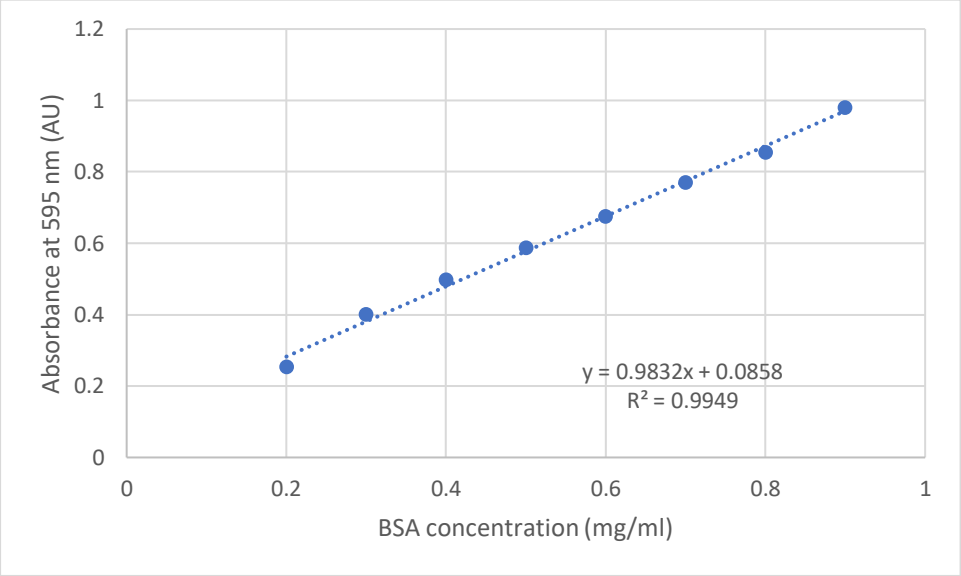

Figure 32. The calibration curve for Bradford Assay by using BSA.

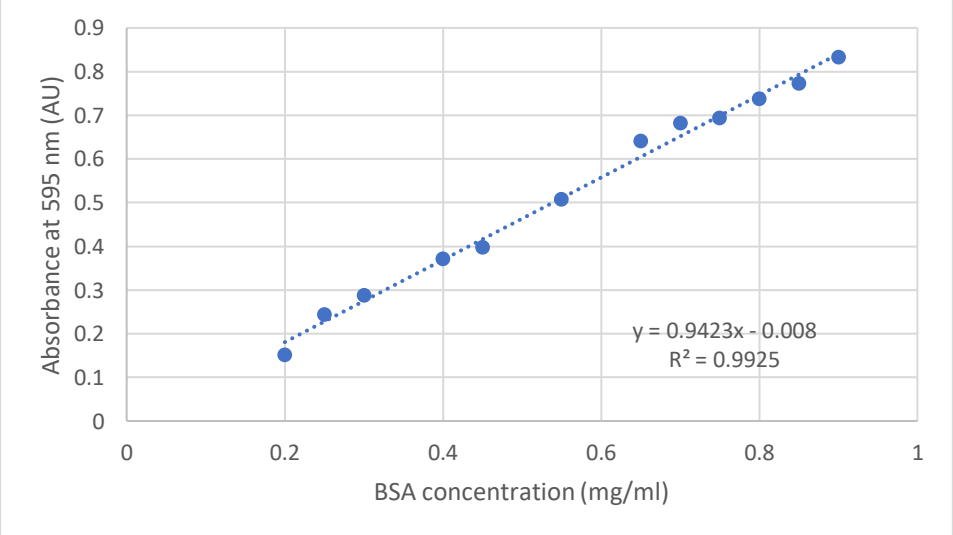

Figure 33. The calibration curve for Bradford Assay by using BSA.