University of Nottingham
UK | CHINA | MALAYSIA

Characterisation of the deubiquitinating enzyme USP4

Peni Indrayudha Rasulan

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

September 2017

Abstract

Ubiquitin Specific Protease (USP) 4 is a deubiquitinating enzyme (DUB) which is an important regulator of different cellular pathways, such as Wnt signalling, and A2A-adenosine receptor signalling. USP4 can remove ubiquitin from RIP1, PDK-1, and Ro52, interacts with SART3 at the spliceosome and regulates TNFα and IL-1β in cancer. The full-length structure of USP4 remains to be investigated. The structure of USP4 consists of an N-terminal DUSP (domain in USPs), two Ubl (Ubiquitin-like) domains, and two subdomains that form one catalytic domain. Only the structure of the DUSP-Ubl and catalytic core lacking the Ubl2 domain has so far been determined. Six constructs were cloned based on the USP4 domain architecture to investigate the structure and function of individual USP4 domains. The constructs were tested by enzymatic activity assays. Site directed mutagenesis of the catalytic site, namely mutants H881N and C311S were generated to form complexes with Ubiquitin, Ubiquitin-GGG, diubiquitin and diubiquitin-L73X.

The 6 constructs were: USP4FL (full-length), USP4htt (head-to-tail), USP4httΔUbl2, USP4C1C2 (catalytic core), USP4ΔDU (lacking the N-terminal DUSP-Ubl domains), and USP4DU (N-terminal domains). Expression and purification was generally done by the Ni column and gel filtration method. Some further purification was performed by anion exchange chromatography. Enzymatic activity assays were conducted using Ub-AMC (Ubiquitin-7-amido-4-methylcoumarin) as the fluorogenic substrate. The mutagenesis was done by exchange of Histidine (H) to Asparagine (N) and Cysteine (C) to Serine (S) in the catalytic triad and protein complexes were created by combining the mutants with Ubiquitin variants. These complexes were set up for crystallization trials. The protein complexes of active site mutant USP4 with Ubiquitin variants were analysed using ESI-MS and ITC.

All the protein constructs were successfully cloned, expressed and purified. USP4httΔUbl2 was well expressed with high yields and USP4htt needed optimisation to increase the yield. Enzymatic activity assays showed that the highest specific activity was obtained for USP4C1C2 whereas USP4htt was the lowest. Some protein crystals were obtained from complexes of active site mutant USP4 with Ubiquitin variants. Only the complex USPFL-H881N with diubiquitin produced crystals that diffracted to 3 Å resolution although the structure has not yet been solved. Binding interaction between various active site mutant constructs of USP4-C311S with ubiquitin variants showed that Ubiquitin and Ub-GGG have the highest affinity for USP4C1C2-C311S. Diubiquitin had the highest affinity for USP4FL-C311S. Together these data provide novel insights into USP4 structure and ubiquitin recognition.

Acknowledgements

I would like to thank my supervisor Dr. Ingrid Dreveny for great supervision during my Ph.D. I am very grateful for invaluable teaching, helping and supporting everything in the USP4 research. I also thank Prof Jonas Emsley who have been supported and put much guidance throughout my Ph.D. I would like to thank Dr. Cristina de Matteis for a very kind guidance and encouragement through my Ph.D.

I would also like to thank all the members of C50 who have always helped and make the laboratory very nice to study and learn: Dr. Simon, Dr. Selva, Dr. Susan, Dr. Hayley, Dr. Tasos, Dr. Affif, Emma, Fajar, Halawa, Steph, Buba, Dr. Chan, Dr. Saleem, and Dr. Monika. It has been a pleasure and unforgettable experience working with all of you.

Many thank for Dikti, Indonesian Ministry of Research, Technology and Higher Education for the funding my Ph.D. I also would like to thank Gemma and Dr. Neil Oldham for ESI mass spectrometry analysis and great discussion. Thanks to PyMOL version 1.3 for great structure picture software.

I also want to say thank to my new family and home in the Nottingham: the big family of PeDLN, special to Dr. Iswandaru & fam, Dr. Maryati & fam, Dr. Wisnu & fam, Dr. Irwan & fam, and Dr. Wahyu. Big thank the Indonesian Ph.D. forum: Mr. Nuklir family and friends. Together with you keep me on the right path, strong, and make me able to finish this Ph.D. Thank Dr. Denny and Ms. Eka for a detail proofreading. The last thank for Mr. Dewanto and family for very hospitality home. I will miss you all.

My grateful thank also to my parents: who always pray and did all the best for me: Bunda, Ibu-Bapak, and Mami-Papi. My pray for all of you always, now and forever.
The deep love and big thank my little family: my wife: Tri Yulianti, my daughters: Yumna, Izzah, and Sofia. You are my everything. Because of you, my Ph.D. life was very colorful and wonderful. As we said, this was a journey, not the destination. May Allah always bless and put us in His Jannah.

Table of Contents

Abstract.....i
Acknowledgement iii
Table of Contents 1
List of Figures 5
List of Tables..... 9
Abbreviations 11
CHAPTER 1..... 12
Introduction 12
1.1. The Ubiquitin system 13
1.2. Polyubiquitin chains..... 15
1.3. Deubiquitinating enzymes (DUBs) 17
1.4. Ubiquitin Specific Protease (USP) 4, USP 11 and USP 15..... 19
1.5. SART3 as a USP4 Binding Partner 23
1.6. Ubiquitin Specific Protease 4 (USP4) and Ubiquitin interactions 24
1.7. USP4 structure 27
1.8. Research design 29
1.9. Aims 35
1.10. Specific Objectives 35
CHAPTER 2..... 37
Introduction to General Methods..... 37
2.1. Affinity Chromatography 38
2.2. Size exclusion chromatography 39
2.3. Protein Crystallisation and Structure Determination 42
2.3.1. Growing crystals..... 43
2.3.2. Cryocooling, X-ray diffraction and structural determination..... 48
2.3.3. Anomalous scattering 49
2.4. Electron spray ionization - mass spectrometry 50
2.4.1. Introduction to mass spectrometry 50
2.4.2. Electrospray ionisation mass spectrometry (ESI-MS) 51
2.4.3. Ion Mobility Mass Spectrometry..... 54

2.5. Isothermal titration calorimetry (ITC) 55
2.6. Ubiquitin cleavage Enzymatic Assay 57
CHAPTER 3 59
Design constructs and Research Outline 59
3.1. Constructs used 60
3.2. Research Outline 63
CHAPTER 4 65
Material and Methods 65
4.1. Construct design and cloning 66
4.1.1. PCR 66
4.1.2. Ligation 68
4.1.3. Transformation 69
4.1.4. Cracking Digestion Test 69
4.1.5. Site-directed mutagenesis 70
4.1.6. Co-expression with Ubiquitin variant 72
4.2. The Protein Expression and Purification 73
4.2.1. Lysate Preparation 73
4.2.2. Affinity chromatography purification 74
4.2.3. Anion exchange 76
4.2.4. Size exclusion chromatography 76
4.3. Ubiquitin and ubiquitin variant expression and purification 77
4.4. SDS-PAGE 78
4.5. Protein concentration calculation 79
4.6. Enzymatic Activity Assay using Ubiquitin-AMC (Ub-AMC) 79
4.7. Crystallisation trials 81
4.8. Native Mass Spectrometry 81
4.9. Isothermal Titration Calorimetry (ITC) 82
Chapter 5 84
Result and Discussion: Cloning, Protein Expression and Purification of Active Proteins (USP4FL, USP4htt, USP4C1C2, USP4ΔDU, USP4httΔUbl2, and USP4DU) 84
5.1. Cloning, Protein Expression, and Purification of USP4FL 85
5.2. Cloning, Protein Expression, and Purification of USP4htt 91
5.3. Cloning, Protein Expression, and Purification of USP4httΔUbl2 94

5.4. Cloning, Protein Expression, and Purification of USP4C1C2 97
5.5. Protein Expression and Purification of USP4ΔDU 100
5.6. Protein Expression and Purification of USP4DU 104
5.7. Discussion 106
CHAPTER 6 111
Result and Discussion: Characterization of active site mutant of USP4 and Crystallisation trials (USP4FL, USP4htt, USP4httΔUbl2, USP4C1C2, and USP4ΔDU) 111
6.1. Protein expression and purification of USP4FL H881N and USP4FL C311S with Ubiquitin Variants 112
6.1.1. Expression and Purification of active site mutant USP4FL-H881N with Ubiquitin variants 112
6.1.2. Expression and Purification of active site mutant USP4FL-C311S in complex with Ubiquitin Variants 118
6.1.3. Expression and purification of active site mutant USP4FL-C311S 122
6.2. Protein expression and purification of USP4htt-H106N and C311S active site mutants 125
6.2.1. Protein Expression and Purification of USP4htt-H106N active site mutant with Ubiquitin Variants 125
6.2.2. Expression and Purification of Active Site Mutant USP4htt-C311S with Ubiquitin Variants 130
6.3. Protein expression and purification of USP4httΔUbl2-H106N active site mutant 135
6.4. Protein expression and purification active site mutant of USP4C1C2 H881N and C311S 137
6.4.1. Protein expression and purification active site mutant of USP4C1C2-H881N with Ubiquitin Variant 138
6.4.2. Protein expression and purification of active site mutant USP4C1C2-C311S with Ubiquitin Variant 144
6.5. Protein expression and purification active site mutant of USP4ΔDU H881N and C311S 154
6.5.1. Protein expression and purification of active site mutant USP4ΔDU-H881N with Ubiquitin variant 154
6.5.2. Protein expression and purification of active site mutant USP4ΔDU-C311S..... 156
6.6. Crystallisation Trials 158
6.6.1. Crystallisation trials of protein complexes of USP4FL 159
6.6.2. Crystallisation trial of protein complex USP4htt 164
6.6.3. Crystallisation trials of protein complexes of USP4C1C2 mutants 165

6.6.4. Crystallisation trials of protein complexes USP4C1C2-mitoxantrone 166
6.6.5. Crystallisation trials of protein complexes with USP4ΔDU 167
6.6.6. Crystallisation trial of protein complex USP4httΔUbl2 167
6.6.7. Crystallisation trials of protein complexes of USP4DU with the SART3 167
6.7. Discussion 171
CHAPTER 7 174
Result and Discussion: Binding Interaction between USP4FL, USP4C1C2, and USP4ΔDU with Ubiquitin, diUbiquitin-L73X, and Ub-GGG 174
7.1. USP4FL interaction with Ubiquitin variant 176
7.1.1. Mass spectrometry analysis 176
7.1.2. Isothermal Titration Calorimetry Analysis 179
7.2. USP4C1C2 interaction with Ubiquitin variant 183
7.2.1. Mass Spectrometry analysis 183
7.2.2. Isothermal Titration Calorimetry analysis 186
7.3. USP4ΔDU interaction with Ubiquitin variant 188
7.3.1. Mass spectrometry analysis 188
7.3.2. Isothermal Titration Calorimetry Analysis 191
7.4. Comparison binding affinity of various USP4-C311S with Ubiquitin variant 193
7.5. The binding interaction between USP4DU and SART3 194
7.6. Discussion 201
Chapter 8 203
Discussion and Future Outlook 203
8.1. Expression and Purification of USP4 204
8.2. Coexpression of Active Site Mutants of USP4 with Ubiquitin Variants and Crystallisation trials 204
8.3. Characterisation Binding Interactions of USP4 with Ubiquitin variants 205
8.4. Future Outlook 208
References 209
APPENDIX 218

List of Figures

Page

1.1. Structure of ubiquitin (PDB code 1UBQ) by PyMOL 1.3…… 13

1.2. Ubiquitin pathway…… 14

1.3. Ubiquitin modification…… 16

1.4. Process of ubiquitination and deubiquitination…… 18

1.5. Sequence alignment between USP4, USP11, and USP15…… 20

1.6. Structure of SART3…… 23

1.7. Schematic representation of the USP4 domain…… 28

1.8. Cartoon representative of circular permutation construct: USP4htt…… 30

1.9. Mutation of the active site from Histidine (H) to Asparagine (N) and Cysteine (C) to Serine (S) of USP4 catalytic site…… 32

1.10. Structure of Linear diubiquitin…… 33

1.11. Structure of Ubiquitin, Ubiquitin-GGG, Linear diubiquitin and diubiquitin-L73X 34

2.1. Illustration of a protein crystallisation phase diagram…… 44

2.2. Sitting drop vapour diffusion and hanging drop vapour diffusion method in crystallisation…… 46

2.3. Schematic ionisation process on mass spectrometry…… 52

2.4. Diagram of a typical power compensation ITC…… 56

3.1. Schematic representation of USP4 constructs cloned…… 60

3.2.Schematic representation of active site mutant H881N of USP4 constructs cloned 62

3.3. Schematic representation of active site mutant C311S of USP4 constructs cloned 62

3.4. Outline of research for elucidating structure of USP4 …… 64

4.1. Schematic overlap PCR for USP4C1C2 …… 68

4.2. Reaction of Ub-AMC with USP4 enzymes …… 79

5.1. Affinity chromatography and SDS-PAGE gel of USP4FL…… 86

5.2. Chromatogram of USP4FL Gel Filtration purification…… 87

5.3. Enzymatic activity assay graph of USP4FL with Ub-AMC…… 88

5.4. Linear range of progress curve to obtain initial rate of USP4FL…… 89

5.5. Sample of raw data from the kinetic assay with USP4FL at 50 nM and the Ub-AMC…… 90

5.6. Graph of the rate of reaction against substrate concentration for USP4FL…… 90

5.7. Affinity chromatography and gel filtration purification of USP4htt…… 92

5.8. Enzymatic activity assay graph of USP4htt with Ub-AMC…….. 94
5.9. Linear range of progress curve to obtain the initial rate of USP4htt…….. 94
5.10. Affinity chromatography and gel filtration purification of USP4httΔUbl2…….. 96
5.11. Protein expression and purification of USP4C1C2…… 98
5.12. Graph of fluorescent of active protein USP4C1C2 with Ub-AMC…….. 99
5.13. Enzymatic activity of active protein USP4C1C2 with Ub-AMC …….. 100
5.14. Chromatogram of Ni column purification of USP4ΔDU…… 101
5.15. Chromatogram of anion exchanges purification of USP4ΔDU…….. 102
5.16. Chromatogram of gel filtration purification of USP4ΔDU…….. 103
5.17. Enzymatic activity of active protein USP4ΔDU with Ub-AMC…….. 104
5.18. Chromatogram of affinity chromatography of USP4DU…… 105
5.19. Chromatogram of gel filtration purification of USP4DU…….. 106
6.1. Nickel column purification of USP4FL-H881N with ubiquitin…… 113
6.2. Size exclusion purification of USP4FL-H881N with ubiquitin…….. 114
6.3. Nickel column purification of USP4FL-H881N with linear diubiquitin…….. 115
6.4. Size exclusion purification of USP4FL-H881N with linear diubiquitin…… 116
6.5. Gel filtration chromatogram of USP4FL-H881N with diUb-L73X…… 117
6.6. Gel filtration chromatogram of the USP4FL-H881N with diUbL73X …….. 118
6.7. Nickel column chromatogram of the USP4FL-C311S-Y928X and diubiquitin…… 119
6.8. Gel filtration chromatogram of USP4FL-C311S-Y928X with diubiquitin…….. 120
6.9. Nickel column chromatogram of USP4FL-C311S-Y928X with Ub-GGG…….. 121
6.10. Gel filtration chromatogram of USP4FL-C311S-Y928X with Ub-GGG …….. 122
6.11. Nickel column chromatogram of USP4FL-C311S…… 123
6.12. Gel filtration chromatogram of USP4FL-C311S…….. 124
6.13. SDS-PAGE gel and chromatogram after Nickle column of USP4htt-H106N with linear diubiquitin…… 126
6.14. SDS-PAGE gel and chromatogram after gel filtration from USP4htt-H106N coexpression and purification with linear diubiquitin…….. 127
6.15. Nickel column chromatogram and SDS PAGE gel of active site mutant USP4htt-H106N with Ub-GGG…… 128
6.16. Gel filtration chromatography of USPhtt-H106N with Ub-GGG…… 129
6.17. Chromatogram of Ni column of USP4htt-H106N with diUbL73X…….. 130
6.18. Ni column chromatogram of USP4htt-C311S with diubiquitin…….. 132

6.19. Chromatogram of gel filtration of USP4htt-C311S with diubiquitin…… 132
6.20. Chromatogram of Ni column of USP4htt-C311S with UbGGG…….. 133
6.21. Chromatogram of gel filtration of USP4htt-C311S with UbGGG…….. 134
6.22. Ni column protein purification of USP4httΔUbl2-H106N with linear diubiquitin 136
6.23. Gel filtration purification of USP4httΔUbl2-H106N with diubiquitin…… 137
6.24. Ni column USP4C1C2-H881N with linear diubiquitin…….. 138
6.25. Chromatogram gel filtration of USP4C1C2-H881N with linear diubiquitin…… 139
6.26. Affinity chromatography of USP4C1C2-H881N with Ub-GGG…….. 140
6.27. Gel filtration chromatography of USP4C1C2-H881N with Ub-GGG…… 141
6.28. Chromatogram of Ni column of USP4C1C2-H881N-diUbL73X…….. 142
6.29. Gel filtration chromatogram of active site mutant USP4C1C2-H881N-diUbL73X 143
6.30. Chromatogram of Ni column of USP4C1C2-C311S with diubiquitin…… 144
6.31. Chromatogram of gel filtration of USP4C1C2-C311S with diubiquitin…….. 145
6.32. Chromatogram of Ni column USP4C1C2-C311S with UbGGG…….. 146
6.33. Chromatogram of gel filtration of USP4C1C2-C311S with UbGGG…… 147
6.34. Chromatogram of Ni column of USP4C1C2-C311S-Y928X with diubiquitin…… 148
6.35. Chromatogram of gel filtration of USP4C1C2-C311S-Y928X with diubiquitin….. 149
6.36. Chromatogram of gel filtration of USP4C1C2-C311S-Y928X with UbGGG…….. 150
6.37. Chromatogram of Ni column USP4C1C2-C311S…… 151
6.38. Chromatogram of gel filtration of USP4C1C2-C311S …….. 152
6.39. Chromatogram of Ni column of USP4ΔDU-H881N with Ubiquitin…….. 154
6.40. Chromatogram of gel filtration of USP4ΔDU-H881N with Ubiquitin…….. 155
6.41. Chromatogram of Ni column USP4ΔDU-C311S…….. 156
6.42. Chromatogram of gel filtration of USP4ΔDU-C311S…….. 157
6.43. Crystals with various solutions from a sample of complex USP4FL-H881N with linear diubiquitin…….. 159
6.44. Crystal diffraction of sample USP4FL-H881N with diubiquitin in resolution 3 Å 160
6.45. Matthews coefficient of crystal USP4FL-H881N and USP4FL-H881N-diubiquitin 161
6.46. Comparison of the SDS-PAGE gel of USP4FL-H881N with linear diubiquitin……. 163
6.47. Crystals from preparation complex USP4htt-H106N with UbGGG…… 164
6.48. Crystal from protein complex active site mutant USP4C1C2-H881N combine with Ub-GGG…… 165
6.49. Crystal from protein complex active site USP4C1C2-mitoxantrone…….. 166

7.1. Spectrum of complex USP4FL with Ubiquitin…… 177

7.2. Spectrum of complex USP4FL-H881N with diubiquitin…… 178

7.3. Spectrum of complex USP4FL-H881N with diubiquitin-L73X…….. 179

7.4. ITC graph of Ubiquitin 100 µM titrated into USP4FL-C311S 10 µM (left), and Ubiquitin-GGG 200 µM into USP4FL-C311S 20 µM (center), and diUbiquitin 200 µM titrated into USP4FL-C311S 20 µM (right)…….. 182

7.5. Spectrum of complex USP4C1C2-H881N with diubiquitin-L73X…….. 183

7.6. Spectrum of complex USP4C1C2-H881N with ubiquitin-GGG…….. 184

7.7. ITC graph of Ubiquitin 250 µM titrated into USP4C1C2-C311S 30 µM (left) and ubiquitin-GGG 300 µM titrated into USP4C1C2-C311S 30 µM (center), diubiquitin 500 µM titrated into USP4C1C2-C311S 50 µM (right)…….. 187

7.8. Spectrum of the active site USP4ΔDU…… 188

7.9. Mass spectrometry result of protein active site USP4ΔDU-H881N and Ubiquitin 189

7.10. Graph of binding interaction using ITC of USP4ΔDU-C311S with diubiquitin (left) and USP4ΔDU-C311S with Ub-GGG (center), and USP4ΔDU-C311S with diubiquitin (right)…….. 192

7.11. Mass spectrometry result of active site USP4DU-SART3 (1st sample)…….. 195

7.12. Mass spectrometry result of active site USP4DU-SART3 (2nd sample)…… 196

7.13. Mass spectrometry result of active site USP4DU-SART3 (3rd sample-1)…… 197

7.14. Mass spectrometry result of active site USP4DU-SART3 (3rd sample-2)…… 198

7.15. Mass spectrometry result of active site USP4DU-SART3 (3rd sample-3)…… 198

7.16. Mass spectrometry result of active site USP4DU-SART3 (3rd sample-4)…… 199

8.1. Schematic diagram showing the interaction between various USP4 with Ubiquitin variant…… 207

List of Tables

3.1. Summary of active site mutant USP4 constructs design used…….. 63

4.1. Summary of USP4 constructs design used…… 67

4.2. Summary the primer for USP4 constructs used…… 67

4.3. Oligonucleotide primers used for the USP4C1C2 construct…….. 67

4.4. Primers for mutagenesis…… 71

4.5. Co-expression for crystallisation…….. 72

4.6. Expressions conditions…….. 73

4.7. Types of buffer for protein purification of all constructs…….. 74

4.8. Gel filtration columns…… 77

4.9. Composition of SDS-PAGE gels…… 78

4.10. Samples for ESI- mass spectrometry…… 82

4.11. Details of ITC experiment between USP4 and Ubiquitin variant…… 83

5.1. Kinetic parameters for USP4FL…… 91

5.2. Comparison of yield of USP4 modified construct after gel filtration purification using Superdex 200 HiLoad 16/60…… 107

5.3. Comparison of elution volume of USP4 modified construct based on gel filtration purification Superdex 200 HiLoad 16/60…… 108

5.4. Specific enzymatic activity of USP4 construct…… 109

6.1. Elution volume and molecular weight prediction of protein complex USP4FL-H881N and C311S-Y928X with Ubiquitin variant…… 124

6.2. Comparison of yield, Elution volume and molecular weight prediction of protein complex USP4htt-H881N and C311S with Ubiquitin variant…….. 135

6.3. Yield, elution volume and molecular weight prediction of protein complex USP4httΔUbl2-H106N with diubiquitin…….. 137

6.4. Comparison of yield, elution volume and molecular weight prediction of protein complex USP4C1C2-H881N, C311S and C311S with Ubiquitin variant… 153

6.5. Comparison of yield, elution volume and molecular weight prediction of protein complex USP4ΔDU-H881N and C311S with Ubiquitin…….. 158

6.6. Cell parameter of sample protein complex USP4FL-H881N with diubiquitin…… 160

6.7. Calculation of Matthews probabilities and solvent content of USP4FL-H881N and USP4FL-H881N with diubiquitin…… 162

6.8. Crystallisation trials of USP4 constructs with various conditions…… 178

7.1. Spectrum of complex USP4FL-H881N with Ubiquitin variant by ESI-MS analysis 180

7.2. Spectrum of complex USP4FL-H881N with diubiquitin result by ESI-MS analysis…….. 185

7.3. Spectrum of complex of USP4ΔDU-H881N with ubiquitin result by ESI-MS analysis …… 190

7.4. Comparison binding affinity between various USP4-C311S with Ubiquitin…….. 193

7.5. Comparison binding affinity between various USP4-C311S with Ub-GGG…… 194

7.6. Comparison binding affinity between various USP4-C311S with diubiquitin…….. 194

7.7. Summary of mass spectrometry analysis of complex USP4DU (1-244) –SART3 (96-574)…….. 200

Abbreviations

BLAST	Basic Local Alignment Search tool
DTT	Dithiothreitol
DUBs	Deubiquitinating enzyme
DUSP	Domain present in USPs
E1	Ubiquitin activating enzyme
E2	Ubiquitin conjugating enzyme
E3	Ubiquitin ligase
ERAD	Endoplasmic reticulum associated degradation
ESI	Electrospray ionisation
HAT	Half-a tetratricopeptide
HECT	Homologous to E6AP C-terminus
IL	Interleukin
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
ITC	Isothermal titration calorimeter
IRF8	Interferon regulatory factor 8
K_D	Dissociation Constant
MALDI	Matrix assisted laser desorption/ionisation
MS	Mass spectrometry
NLS	Nuclear localisation signal
OTU	Otubain Protease
PDK1	Processing factor 3 Phospo-inositide-dependent kinase1
PRPF3	Pre-mRNA Processing Factor 3
QTOF	Quadrupole-time of flight
RIP-1	Receptor interacting
RING	Really interesting New Gene
RRM	RNA recognition motif
SART3	Squamous cell carcinoma antigen recognized by Tcell 3
TGF-β	Transforming growth factor-β
TGF	Transforming growth factor
TNF-α	Tumour necrosis factor-α
TOF	Time of flight
Ubl	Ubiquitin like
Ub-AMC	Ubiquitin-7-Amido-4-methylcoumarine
UCH	Ubiquitin C-terminal hydrolase
UNP	Ubiquitous nuclear protein
USP	Ubiquitin Specific Protease
USP4FL	USP4 full length
USP4C1C2	USP4 catalytic core
USP4htt	USP4 head to tail
USP4httΔUbl2	USP4 htt without Ubl2

CHAPTER 1

Introduction

1.1. The Ubiquitin system

Protein degradation is an equally important system as protein synthesis for the maintenance of protein homeostasis (Nath and Shadan, 2009). Research into protein degradation system has been rapidly expanding after the discovery of the ubiquitin system by Avram Hersko, Aaron Ciechanover, and Irwin Rose in 2004 (Ciechanover, 2005; Goldberg, 2005).

Ubiquitin is a 76-amino acid residue protein which is 8.5 kDa and highly conserved in all eukaryotes (Sippl, 2011) (Figure 1.1). Ubiquitin has functions in protein degradation, in signaling pathway regulation, DNA repair, transcription, and endocytosis (Hersko and Ciechanover, 1998; Sun and Chen, 2004; Welchman et al., 2005). Ubiquitin is a small protein which is activated by E1, E2, and E3 enzymes.

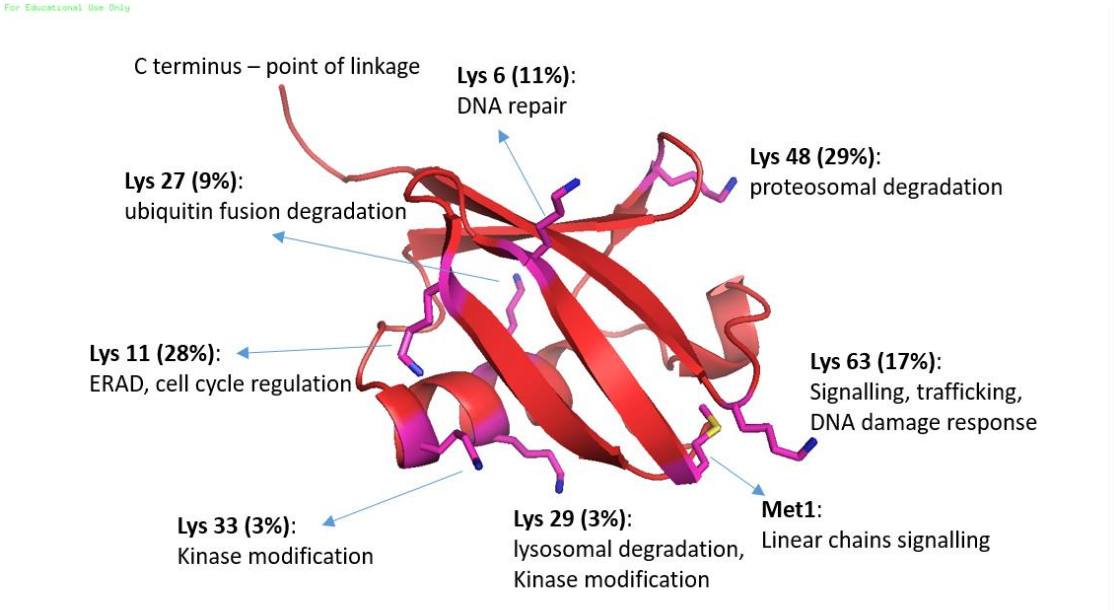


Figure 1.1. The structure of ubiquitin (PDB code 1UBQ) by PyMOL 1.3.
Ubiquitin has 7 lysine residues on the different surface of the molecule which has specific functions (Adaptation from Komander, 2009).

Ubiquitin will covalently conjugated to lysine residues of substrate proteins by E1, E2, and E3 enzymes. E1 (ubiquitin activating enzyme) will activate ubiquitin through an ATP-dependent mechanism by forming a thiol ester bond between its active cysteine and the C-terminus of Ubiquitin (Figure 1.2). Activated ubiquitin is then delivered to a ubiquitin-conjugating enzyme (E2) and finally by ubiquitin ligase (E3) transferred to a lysine residue of a target protein (Wang et al., 2012). The E3 enzyme has two major families: The homologs to E6-AP carboxyl terminus (HECT) and the really interesting new gene (RING) like ligases. The enzyme E3 defines the specificity of the target protein to be ubiquitinated. The ubiquitinated substrate is then sent to the 26S proteasome for degradation (Varshavsky, 2012) or to another function of ubiquitination.

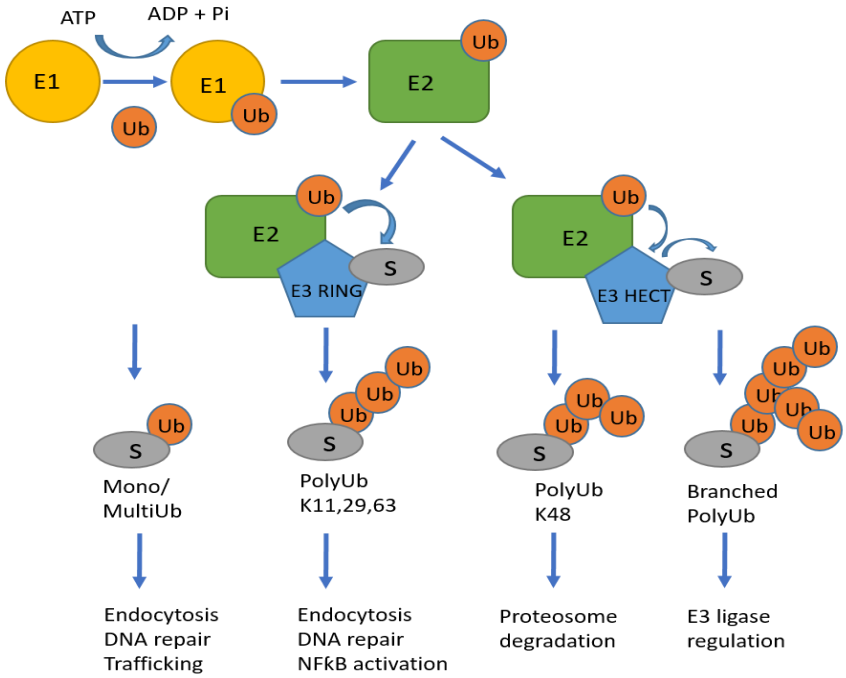


Figure 1.2. Ubiquitin pathway. Ubiquitin will be activated first by E1, a ubiquitin-activating enzyme (E1), using ATP. Then transferred to E2, a ubiquitin-conjugating enzyme, and bind with E3 molecules, a ubiquitin protein ligase, which makes Ubiquitin possible to conjugate to the substrate (Adaptation from Perret et al., 2011).

Ubiquitin contains 7 lysine residues, namely: K6, K11, K27, K29, K33, K48 and K63 (Pickart and Eddins, 2004). The linkage between protein substrate and a specific lysine residue will lead to various pathways such as the inflammatory response, apoptosis, and cell cycle progression (Wang et al., 2012).

1.2. Polyubiquitin chains

A protein can be modified by monoubiquitin or multi-ubiquitin chains (Hicke et al., 2005). Ubiquitin polymers serve as interaction sites for proteins. Already identified are 20 different types of ubiquitin-binding domains (Rahigi and Dikic, 2012). Ubiquitin recognition is the common principle for both degradative and non-degradative ubiquitin functions (Komander et al., 2009; Komander, 2009).

Ubiquitin chains, which are composed of multiple ubiquitin moieties are linked via isopeptide bonds and can be formed on Met1 and all internal Lysine residue of ubiquitin molecule (Lys6, Lys11, Lys27, Lys 29, Lys33, Lys48, and Lys63) (Dikic et al., 2009). Ubiquitination can lead to monoubiquitination, multi monoubiquitination, and homogenous ubiquitin chain (if the same residue is modified during elongation), mixed ubiquitin chain (if different linkages alternate at succeeding position of the chains branched ubiquitin chain and unanchored ubiquitin chain) (Figure 1.3) (Komander and Rape, 2012).

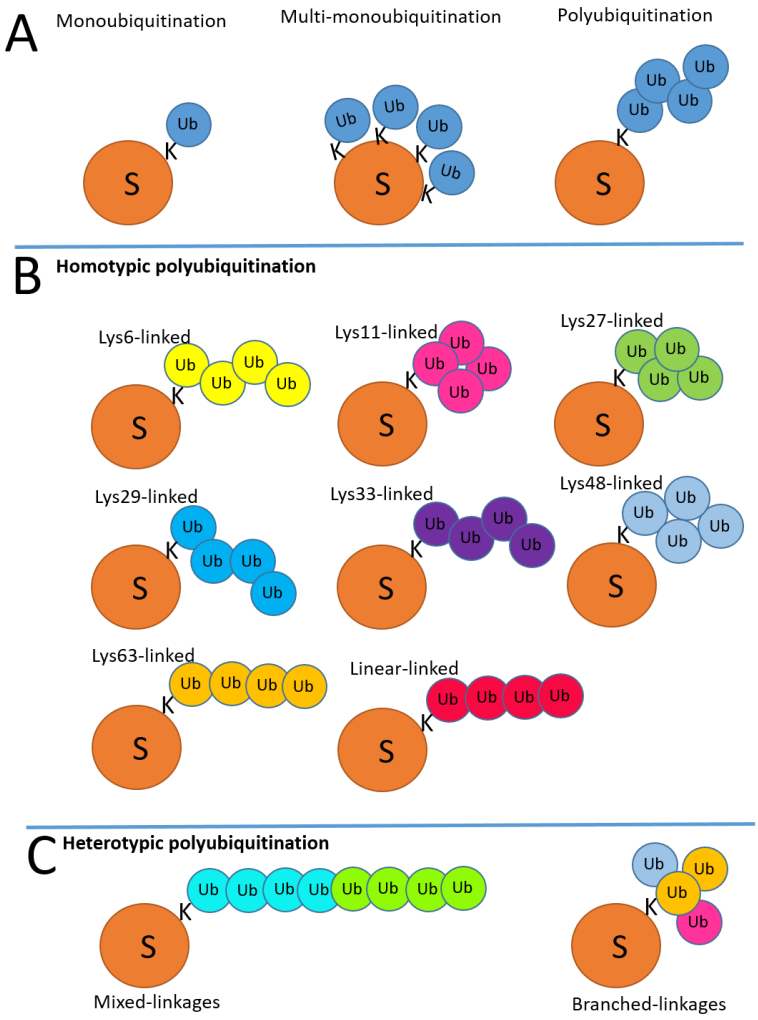


Figure 1.3. Ubiquitin modification
Ubiquitin modification has three general layouts: mono-Ub, Multi-mono-Ub, and polyubiquitination (Adaptation from Komander, 2009).

The role of each protein ubiquitin linkage is different. For example, monoubiquitination (one protein linked with one ubiquitin) has a role in signaling, endocytosis and DNA repair (Rittinger and Ikeda, 2017). In multiple monoubiquitinations, the linkage has a function in signaling and endocytosis. Mixed-linkage polyubiquitination and heterologous modification have a function in proteasomal degradation.

Other functions of homotypic polyubiquitin linkage are known. Lys48-linked polyubiquitin is for proteasomal degradation. Otherwise, Lys63-linked chains have a function in DNA damage repair, cellular signaling, intracellular trafficking, and ribosomal biogenesis (Li and Ye, 2008; Hospenthal et al., 2013). The other ubiquitin chains are: Lys-6 linkages for DNA repair; Lys-11 linkages for endoplasmic reticulum-associated degradation (ERAD); Lys-27 for ubiquitin fusion degradation; Lys-29 for lysosomal degradation and kinase modification; and Lys-33 is involved in kinase modification (Dikic et al., 2009).

Ubiquitin chains have different topologies. Ubiquitin can form eight different homotypic chains. The structure has already been solved for Lys6-Linked diubiquitin (PDB code 2XK5), Lys11-linked diUb (2XEW), Lys48-linked diUb (1AAR), Lys63-linked diUb (2JF5), and Met1-linked diUb chains (PDB code 2W9N) (Husnjak and Dikic, 2012).

1.3. Deubiquitinating enzymes (DUBs)

Ubiquitination is a reversible process, which can be reversed by deubiquitinating enzymes (DUBs). DUBs are proteolytic enzymes which cleave ubiquitin or ubiquitin-like proteins from ubiquitin conjugates or target proteins (Huang and Cochran, 2013) (Figure 1.4). DUBs regulate the activity of proteins (Clague et al., 2012). For this reason, DUBs and ubiquitin have a key role in maintaining protein stability.

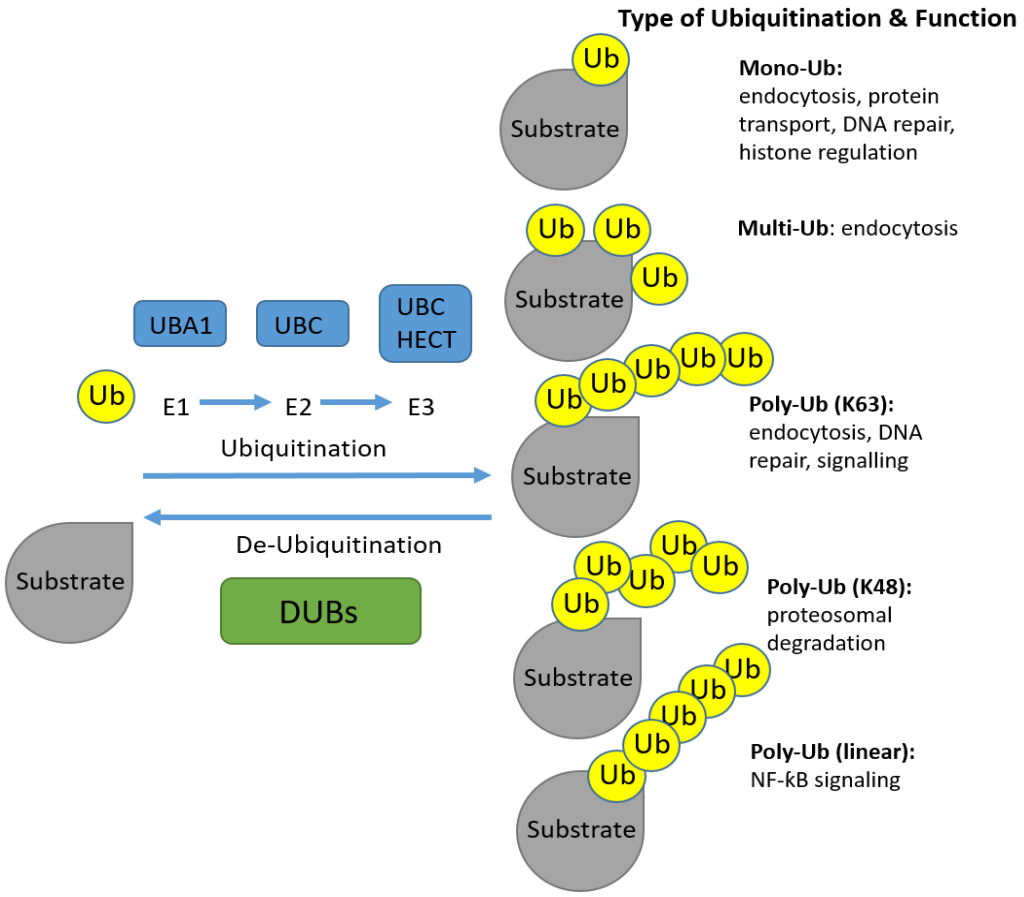


Figure 1.4. Process of ubiquitination and deubiquitination.
A protein can be modified by mono, multi, or polyubiquitination
(Adaptation from Wang *et al.*, 2012)

In the ubiquitin pathway, the roles of some DUBs are to carry out activation of the ubiquitin proproteins by co-translationally cleaving precursors. Ubiquitin recycling is carried out by removing isopeptide-linked ubiquitin from the substrate to free ubiquitin and reversing the ubiquitination or ubiquitin-like modification of protein targets. They are responsible for the monoubiquitin regeneration from poly-ubiquitinated protein (Reyes-Turcu et al., 2009).

DUBs are a superfamily of proteases which can be divided into two groups: cysteine protease and metalloproteases (Komander et al., 2009). Cysteine protease DUBs are ubiquitin specific protease (USP), ubiquitin C-terminal

hydrolase (UCH), Otubain protease (OTU), Josephine and MINDY-1 family (Abdul Rehman et al., 2016). Almost all DUBs are cysteine proteases (Nijman et al., 2005) and the largest class of DUBs are the ubiquitin-specific proteases (USP).

The USP family consists of more than 55 members (Clague et al., 2012). USPs has approximately 300–800 amino acids which contain a catalytic triad, namely Cysteine, Histidine and Aspartate residue (Nijman et al., 2005). USP has drawn special attention as anti-cancer targets and ideal candidates for drug development. Based on the genetic and functional analysis, USP was include in the category of cancer-associated proteases. More than 30 USPs have been linked with cancer indirectly and directly. USP also has a unique biochemical structure that made USP desirable target for anticancer therapies (Pal et al., 2014).

1.4. Ubiquitin Specific Protease (USP) 4, USP 11 and USP 15

Ubiquitin Specific Protease (USP) 4, the main DUB investigated in this study has two paralogues, namely USP11 and USP15. USP4, USP11, and USP15 have a common domain organisation, consisting of a DUSP (domain in USP), two Ubl (Ubiquitin like domain) and a catalytic domain (Figure 1.5). USP4 is most closely related to USP15 and USP11. USP4 and USP15 have 56.9% identity of their

amino acid. USP4 and USP11 have 44.5% identity, and between USP11 and USP15 has 43.2% identity (Harper et al., 2014; Elliott et al., 2011).

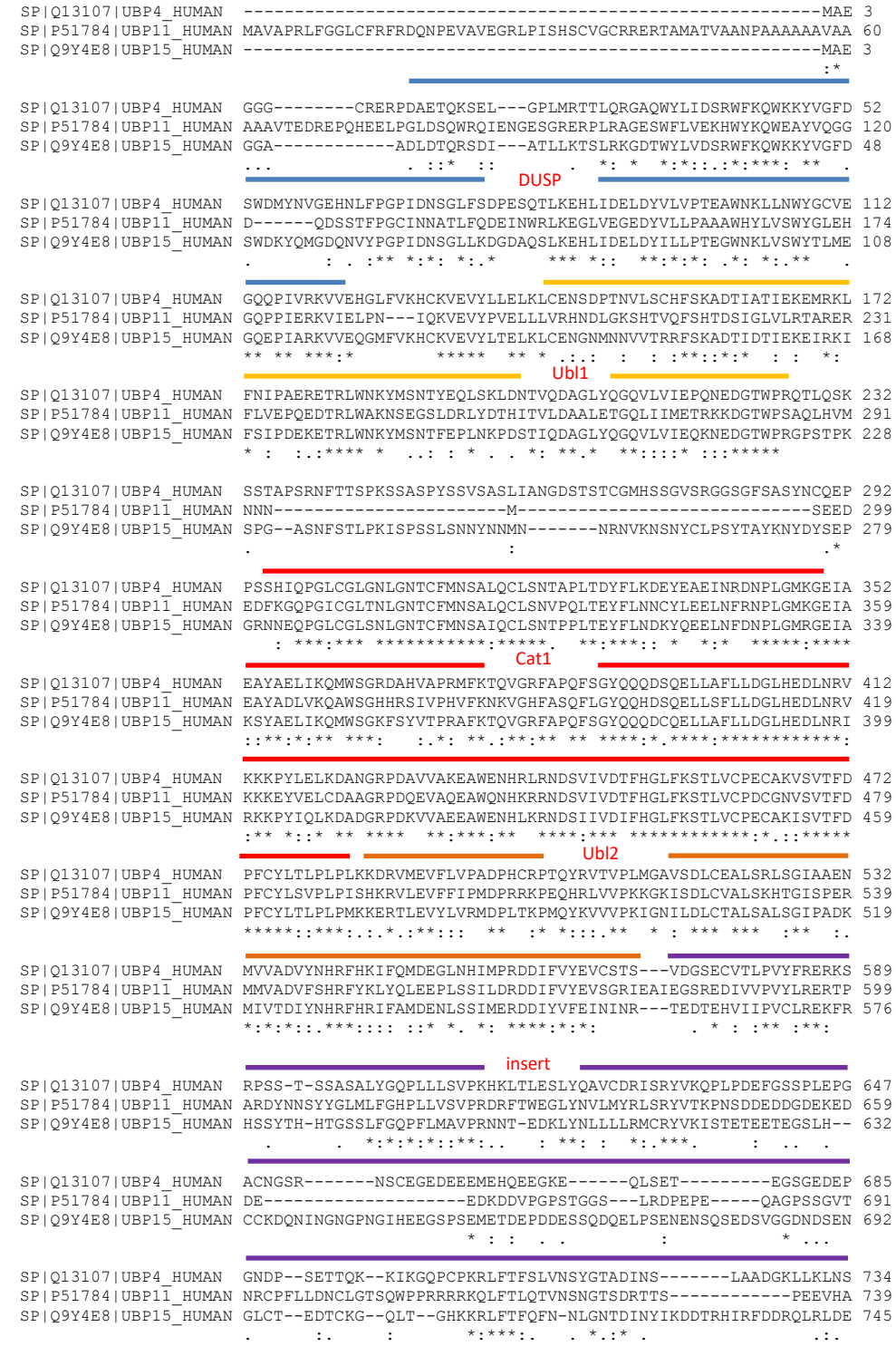


Figure 1.5. Sequence alignment between USP4, USP11, and USP15
All the construct compared based on the structure domain: DUSP (blue), Ubl1 (yellow), Catalytic 1 (red), Ubl2 (orange), insert (purple) and Catalytic 2 (green).

Figure 1.5. (Continued) Sequence alignment between USP4, USP11, and USP15
All the construct compared based on the structure domain: DUSP (blue), Ubl1 (yellow), Catalytic 1 (red), Ubl2 (orange), insert (purple) and Catalytic 2 (green).

USP4 and USP15 have similar functions, such as regulating cell growth, embryonic development, and innate immunity via their interaction with TGF-β, Wnt/β-catenin, and NF-κB pathway proteins. USP4 and USP15 are known to interact with the spliceosome (Vlasschaert, et al., 2015).

Recent research by Clerici, et al., (2014) reveals that there are differences in activity between USP4, USP11, and USP15 when different domains were removed. The activity of USP4 behaves differently compared to USP11 and USP15. The structure of USP4CD (CD: catalytic domain; lacking the DUSP-Ubl domains) has a 15-fold lower activity compared to USP4FL. USP15CD showed threefold lower activity relative to USP15FL. USP11 has a poor solubility of the CD construct, and it was difficult to measure the activity (Clerici et al., 2014).

DUSP-Ubl domain in USP4 is required for ubiquitin dissociation. In the kinetics of USP4-ubiquitin dissociation measurements, USP4FL ubiquitin release is

complete in less than 1 minute, but USP4CD needs almost 1 hour. USP15 has the same behaviour activity with USP4. USP11 has a different one, the DUSP-Ubl domain not able to promote the ubiquitin release (Clerici et al., 2014).

USP4 and USP15 have activity in mRNA splicing. The USP4-DUSP Ubl domains are responsible for the interactions with SART3 that involved in mRNA splicing, but there are no interaction partners for USP15-DUSP Ubl (Harper et al., 2011). USP4 is recruited by U4/U6 small nuclear RNA recycling factor SART3 (squamous cell carcinoma antigen recognized by T-cell 3) to remove the chain of polyubiquitin-Lysine-63 from pre-mRNA Prp3 (processing factor 3) and controls the spliceosome assembly of the splicing process at the distinct stage (Song et al., 2010). SART3 recruited USP15 to regulate deubiquitination of free ubiquitinated histone H2B during transcription from nucleosome (Long et al., 2014).

The structure of USP4, USP11, and USP15 has similar of pattern in cellular regulation. Vlasschaert et al., (2015) reveal in the insert region, USP4 Ser675 and Ser680 are conserved in USP15 but absent from USP11. The phosphorylation site USP4 Tyr539 is conserved in USP15 but substituted by Phe in USP11. In the downstream, USP11 has Tyr551 and Tyr554 whereas His and Phe are universally present in USP4 and USP15, respectively.

The N-terminus of USP11 is longer, more hydrophobic and more disorder compare to USP4 and USP15. The C-terminus of USP15 has rich in aspartic acid,

glutamic acid, and asparagine rather than USP4 and USP11 (Vlasschaert et al., 2015).

1.5. SART3 as a USP4 Binding Partner

SART3 is the squamous-cell carcinoma antigen recognised by T cells-3. SART3 was a 110-kDa nuclear RNA-binding protein. It consists of eight HAT (half a-tetratricopeptide) repeats in the N-terminus followed by NLS (nuclear localisation signal) sequence and two RRMs (RNA recognition motifs) (Figure 1.6) (Gu et al., 1998; Park et al., 2016).

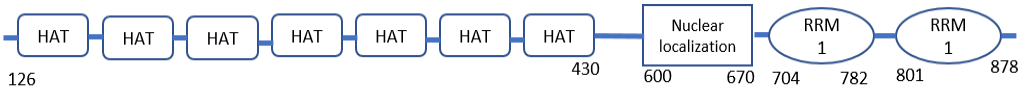


Figure 1.6 Structure of SART3
SART3 consist of the HAT repeats domain, nuclear localization sequence and the two RNA recognition motifs at the C terminus

SART3 is a RNA-binding protein expressed in the nucleus of the majority of proliferating cells, including normal cells and malignant cells (Harada et al., 2012). SART3 has functioned as a substrate-targeting factor for paralogous deubiquitinase USP4 and USP15. SART3 recruits USP4 to ubiquitinated PRPF3 within the U4/U5/U6 tri-snRNP complex, promoting PRPF3 deubiquitination and regulating the spliceosome U4/U5/U6 tri-snRNP spliceosome complex disassembly (Song et al., 2010). Another function of SART3 is to recruit the deubiquitinase USP15 to histone H2B and mediate histone deubiquitination. Also, SART3 regulates gene expression and DNA repair (Long et al., 2014).

Grazette (2015) reveal the structure of SART3's N terminus (residues 96-574) that was a homodimer comprised of a series of anti-parallel α-helices (PDB code: 5JPZ). This dimer binds to two molecules of USP15-DU (DUSP Ubl domain). The DU-finger of USP15 coordinated the SART3 binding interaction. Another crystal structure of SART3 was built with USP4 (PDB code: 5CTR). This structure showed that the linker between the DUSP and Ubl domains of USP4 forms a β-structure, and the residue forming the β-hairpin loop, namely L126, F127, V128, and H130, are the SART3 binding determinant (Park et al., 2016).

1.6. Ubiquitin Specific Protease 4 (USP4) and Ubiquitin interactions

USP4 is one of the deubiquitinating enzymes that has been implicated in the process of receptor trafficking (Zhang et al., 2012; Milojevic et al., 2006). USP4 contains extensive and divergent sequences beyond its catalytic core. USP4 was previously known as ubiquitous nuclear protein (UNP) (Gupta et al., 1993), and identified as a proto-oncogene related to Tre 2/Tre 17 (USP6). USP4 shows a consistently elevated gene expression level in small cell tumors and lung adenocarcinomas, suggesting that it may have a possible causative role in neoplasia (Gray et al., 1995).

USP4 is an oncoprotein that has been found in small cell tumours of the lung (Gray et al., 1995). USP4 also has roles in Wnt signaling (Zhao et al., 2009) and can be bound to the C-terminus of A2A receptor and act as deubiquitinating protein by regulating the cell surface level of the A2A receptor (Milojevic et al.,

2006; Ventii and Wilkinson, 2008). The adenosine A2AR receptor has been shown to be deubiquitinated by USP4 which then leads to cell surface delivery of the receptor. The interaction of USP4 with the A2AR receptor is specific as ubiquitination levels of other receptors such as the beta-2 adrenergic receptor is not affected by overexpression of USP4 (Berthouze et al., 2009).

USP4 is recruited to the spliceosome by complex formation with SART3 (Song et al., 2010). USP4 may have a role in regulating the ubiquitination state and or degradation of the pRb tumour suppressor (Gilchrist and Baker, 2000). USP4 has activity related to cancer and disease (Clague et al., 2013). USP4 also known as UnpEL is ubiquitinated by Ro52 and translocate to the cytoplasmic rod-like structures. UnpEL colocalizes with Ro52 when Ro52 is overexpressed in HEK293 cells (Wada and Kamitani, 2006; Wada et. al, 2006). In carcinoma cells, USP4 promotes TNF-α-induced apoptosis by deubiquitination of RIP1 (Hou et al., 2013). Xiao et al. (2012) explained that USP4 is a negative regulator of TNFα- and IL-1β-induced cancer cell migration. Also, USP4 has an activity to inhibit p53 by deubiquitinating and stabilising ARF-binding protein 1 (ARF-BP1) (Zhang et al., 2011b).

USP4 also has a role in the antiviral immune response. A Recent study by Wang et al. (2013) showed that USP4 interacts with RIG-I (retinoic acid inducible gene) and remove K48-linked polyubiquitination chains from it. Degradation of RIG-I on the host after viral infection represent protective responses.

Ubiquitin specific protease (USP) 4 is an important regulator of different cellular pathways (Clerici et al., 2014). Another study by Uras et al. (2012) showed that USP4 is an enzyme that removes ubiquitin from phospho-inositide-dependent kinase 1 (PDK1). Phosphorylation by PDK1 is important for many growth factor-activated kinases that useful for treat diseases, such as diabetes. On the other hand, a systematic review study showed USP4 is a protease/protease inhibitor gene which related to Crohn's disease (Cleynen et al., 2011). USP4 was also reported regulates DNA repair and cellular survival upon DNA double-strand breaks (DSB) induction. USP4 interacts with CtIP and the MRN complex via its C-terminal insert domain and auto regulatory deubiquitylation mechanism (Wijnhoven et al., 2015).

Research by Hou, et al., (2013) suggests that USP4 may be a potential therapeutic target for head and neck squamous cell carcinoma because of its tumor suppressor role activity. USP4 can interact with receptor-interacting protein 1 (RIP1) and removes K63-linked ubiquitin from it. USP4 also inhibit p53-mediated apoptosis and cell cycle checkpoints, enhance transforming growth factor (TGF)-β respond, inhibit nuclear factor-κB (NF-kB) signaling and antagonising lung cancer migration (Clerici et al., 2014).

A recent study showed that via K48-linked deubiquitinase USP4 physically interacted with interferon regulatory factor 8 (IRF8) function which stabilised IRF8 protein levels in Treg cells (Lin et al., 2017). Another research showed by

removing monoubiquitination from SMAD4, a common intracellular effector for TGF-β family cytokines, USP4 induces activin/BMP signaling (Zhou et al., 2017).

USP4 has known as an oncogene. USP4 is a negative regulator of TNFα- and IL-1β-induced cancer cell migration (Xiao et al., 2012). USP4 has been reported to stimulate the TGFβ-mediated EMT (*Epithelial-mesenchymal transition)* in breast carcinomas (Zhang et al., 2012). The relation between USP4 and cancer was also founded on its role in the Wnt signaling pathway. Wnt pathway is deregulated. It controls multiple developmental processes in human cancer. USP4 interacts with Wnt signaling components: the Nemo-like kinase (Nlk) and the T-cell factor 4 (TCF4) cleave either K48 or K63 with equal efficiency (Zhao et al., 2009). USP4 as ubiquitin-specific proteases are attractive to the cancer drug targets (Sippl et al., 2011). As a cancer drug target, the structure of USP4 is interesting to study.

1.7. USP4 structure

USP4 is 111 kDa deubiquitinating enzyme which cleaves the bond between the Gly residue (G76) on the C-terminal Ubiquitin and the protein target. The structure of USP4 consists of 963 amino acid residues (Figure 1.7). Apart from a functional cysteine protease domain, USP4 harbors an N-terminal domain present in ubiquitin-specific proteases (DUSP) and two ubiquitin-like (Ubl) domains (Zhu et al., 2007; Harper et al., 2011). The role of the Ubl domain

integral to the protease domain is currently unknown while the N-terminal Ubl and DUSP domains act as protein-protein interaction modules (Song et al., 2010; Zhao, et al., 2012).

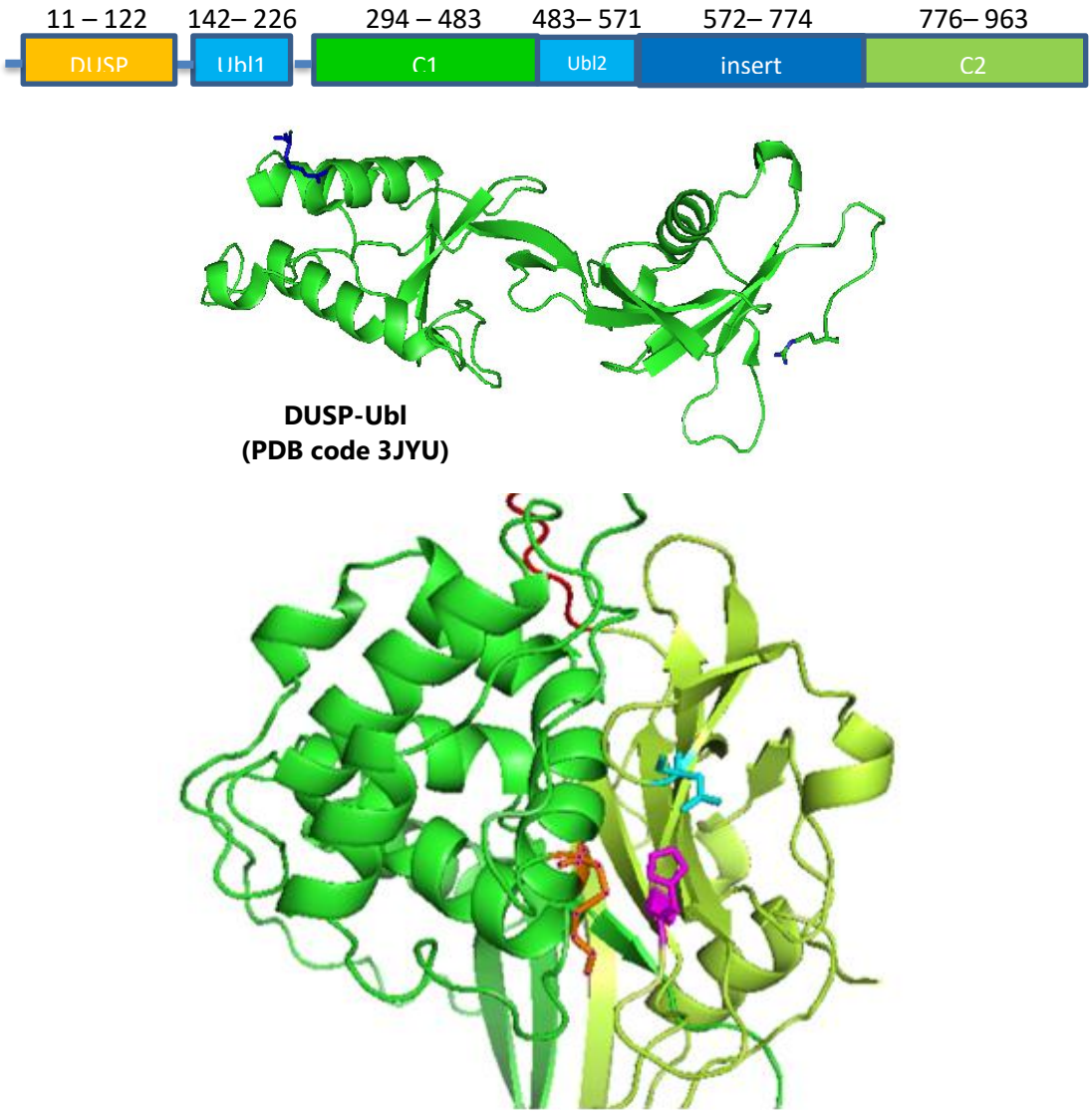


Figure 1.7. Schematic representation of the USP4 domain
USP4 structure consisting of, DUSP-Ubl (left) and D1D2 (C1C2) with catalytic triad: Cysteine (orange), Histidine (pink) and aspartic acid (blue)

The USP4 protein consists of the DUSP-Ubl domains at the N terminus, catalytic core domain consisting of Catalytic1 and catalytic2 subdomain, Ubl2,

and insert, at C terminus. The full-length structure of USP4 is unknown. The PDB database currently only has the structure of the DUSP-Ubl domain (PDB ID: 3JYU) and catalytic core domain (D1D2, PDB ID: 2Y6E) (Clerici, et al., 2014). Other domain structures are still unsolved. There is neither detailed information for USP4 on how its catalytic activity is regulated nor how ubiquitinated proteins are recognised. Clerici et al. (2014) explained that the N-terminal DUSP-Ubl domains are required to achieve USP4's full catalytic activity. A construct without the DUSP-Ubl domains, namely USP4CD, has a weak activity compared to USP4FL (USP4 full length). The DUSP-Ubl domains activate USP4 by promoting ubiquitin release. The catalytic activity of USP4CD will increase when the DUSP-Ubl domain interact with the insert. Interestingly, there is a role of the linker to provide the conformational freedom between DUSP-Ubl and catalytic domain for their interaction and USP4 activation.

1.8. Research design

In this research, several constructs of modified USP4 were engineered to gain additional insights into the USP4 protein structure. First, the USP4FL (USP4 full length) construct was cloned, expressed and purified. The second was a circular permutation construct, namely USP4htt (USP4 head to tail). In this construct, the catalytic2 sub-domain was engineered as the new N terminus, then connected with the catalytic1 subdomain and Ubl2 by a linker region, so that the catalytic1 subdomain and Ubl2 became the new C terminus (Figure 1.8).

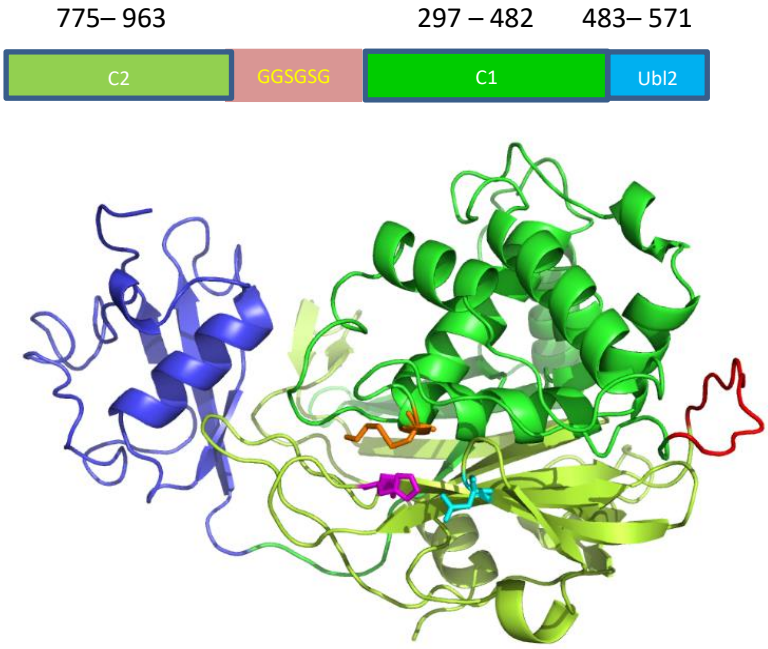


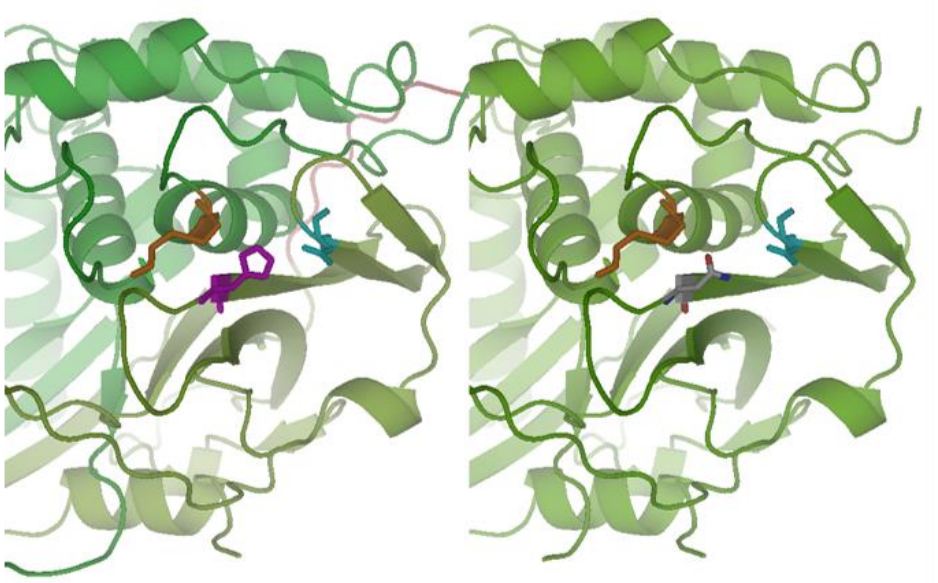
Figure 1.8. Cartoon representative of circular permutation construct: USP4htt
USP4htt construct also contain catalytic triad: Cysteine (orange), Histidine (pink) and aspartic acid (blue) (right)

The main idea creating this construct was based on the hypothesis in the literature at the time that the Ubl2 domain is auto-inhibiting USP4 by residing in the active site mimicking ubiquitin. This should improve the chances of crystallisation trials to solve the structure of the catalytic domain with the Ubl2 domain. The USP4htt is a circular permutated version of the USP4 protease domain and was engineered by deletion of DUSP Ubl1 and the insert domain. The insert domain is a region predicted to be flexible with a function that is still unclear. Because it was difficult to design and predict the linker conformation between the Ubl2 and catalytic2 sub-domain without the insert domain, the opposite direction was chosen by changing the C terminus to a new N terminus in the catalytic2 subdomain. Then the catalytic 2 was connected to the catalytic

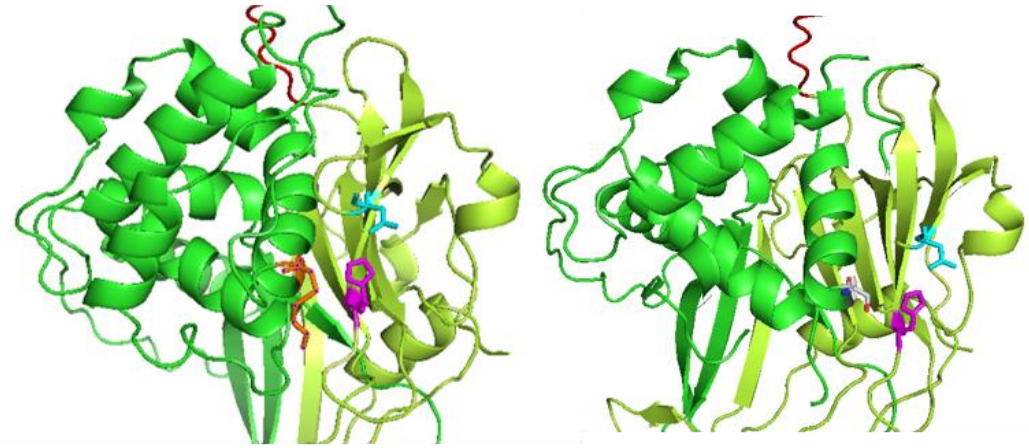
1 subdomain and Ubl2 using a short linker containing serine and glycine and a new C terminus created.

The third, was USP4httΔUbl2, a truncation of USP4htt without Ubl2. Fourthly, a construct of USP4C1C2 is the catalytic core domain of USP4 lacking the insert with the UBL2 domain. This construct consists of catalytic 1 and catalytic 2 subdomains connected by a short linker ASTSK that is present at this position in the related USP8 structure (Harper, *et al*., 2014). The fifth was USP4ΔDU, a construct of the USP4 protease domain lacking the DUSP-Ubl domain. The last, USP4DU, the shortest construct which consists only DUSP and Ubl1 domain.

In addition, USP4 has a catalytic core domain of almost 350 amino acids, which consists of a catalytic triad, namely a cysteine, a histidine, and aspartate (Zhang et al., 2011a) residue. Site-directed mutagenesis was performed to determine the structure of USP4 in complex with Ubiquitin and Ubiquitin variants, by changing its catalytic triad, namely Histidine (H) or Cysteine (C). This mutagenesis has changed the residue 881 from H (Histidine) to N (Asparagine), and residue 311 from Cysteine (C) to Serine (S) (Figure 1.9). Moreover, Asparagine and Serine mimicking with Histidine and Cysteine, respectively. This would increase the chance for crystallisation. It has already been shown that in USP4, C311A mutation results in a loss of the catalytic activity and its ubiquitination by TRIM21 were enhanced (Wada et al., 2006).



Mutation Histidine (purple, left side) to Asparagine (grey, right side)



Mutation Cysteine (orange, left side) to Serine (grey, right side)

Figure 1.9. Mutation of the active site from Histidine (H) to Asparagine (N) and Cysteine (C) to Serine (S) of USP4 catalytic site

Moreover, Ubiquitin contains two Glycine residues at its C-terminus. The glycine 76 (G76) on ubiquitin is usually attached to a substrate lysine residue of the protein target and will be cleaved by the catalytic triad of USP4. Site-directed mutagenesis of the active site of USP4 will prevent cleavage and allow protein complexes to form. The structure of linear diubiquitin has a linker

between two Ubiquitins via Gly76-Met1 (Figure 1.10). Theoretically, this linker will be interacting with the active site of USP4 to form a protein complex. Also, the other ubiquitin has free G76 on its C-terminus which has more flexibility to attach. To decrease this flexibility and to increase the binding affinity between diubiquitin and active site mutant of USP4, residue number 73 on the proximal ubiquitin (Figure 1.10) was changed to a stop codon. This construct was termed diubiquitin L73X.

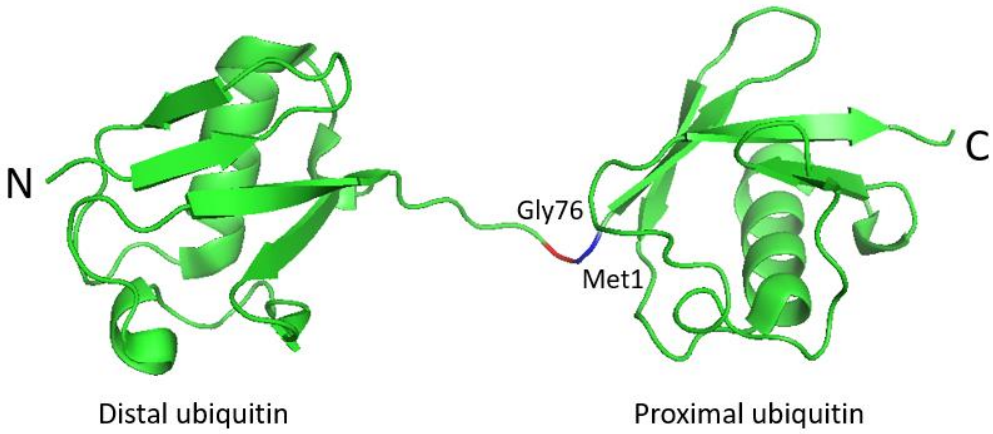


Figure 1.10. Structure of Linear diubiquitin
The linear diubiquitin made by connecting the Gly76 to Met1 of two Ubiquitin. Diubiquitin-L73X created by truncation the Lys73 on the proximal ubiquitin in C terminus to Stop codon

Another investigation to improve the binding site interaction between Ubiquitin and active site mutant of USP4 was to add two Glycine residues to Ubiquitin to mimic an isopeptide linkage (Figure 1.11). The new structure of ubiquitin with triple glycine, namely Ubiquitin-GGG (Ub-GGG) has a purpose of stabilising the interaction for form a protein complex to crystal protein

production. These complexes will be excepted to reveal the molecular mechanism of how USP4 domains interact with ubiquitin.

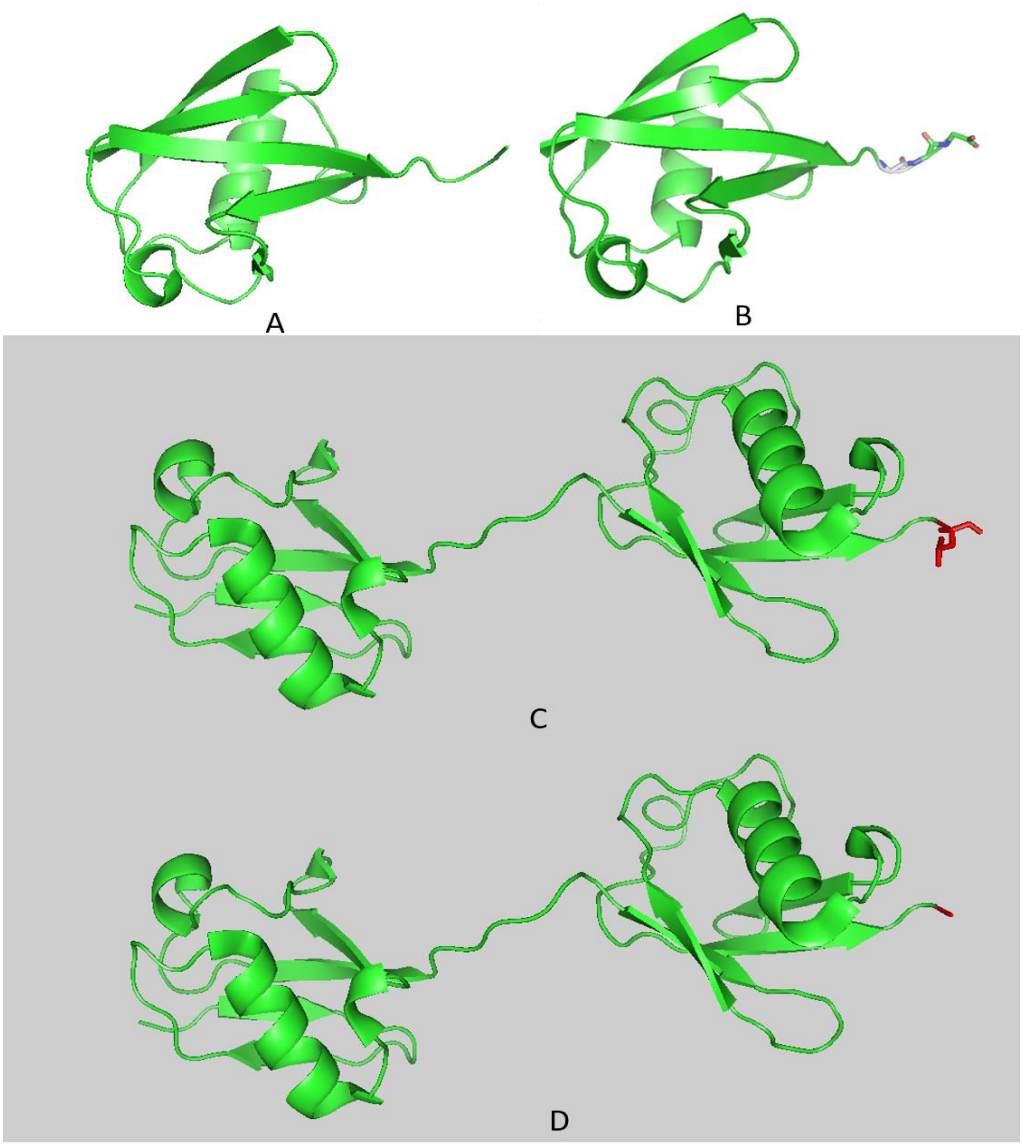


Figure 1.11. The structure of Ubiquitin (A), Ubiquitin-GGG (B): an additional Glycine residue added in the C terminus; linear diubiquitin (C) with Leucine on the C terminus; and diubiquitin L73X with mutated Leucine on the residue 73 into stop codon (D).

In summary, the USP4 modified constructs: USP4FL, USP4htt, USP4C1C2, USP4ΔDU, and USP4httΔUbl2 were mutated, and the interaction with ubiquitin, linear diubiquitin, UbGGG, and Ubiquitin-L73X were characterised and

continued for crystallisation trial. In this research, we also designed USP4 constructs and investigate the binding interaction with ubiquitin variants to characterise the role of each USP4 domain. Together, this will provide novel insights into USP4 structure and function and aid in generating a framework for drug discovery.

1.9. Aims

1. The project will focus on elucidating the USP4 domain structure and its impact on the catalytic function and substrate recognition of USP4.
2. Investigate the interaction of USP4 with ubiquitin variants to determine which USP4 domains are involved in the interaction and how these differ between ubiquitin variants.

1.10. Specific Objectives

The objectives of this project are:

1. Expression and purification of USP4 constructs, including deletion and circular permutated truncation constructs lacking the DUSP-Ubl and Ubl2 domain.
2. Co-expression of active site mutants USP4 and truncation constructs with ubiquitin variants for crystallisation trials.

3. Investigating the enzyme kinetic parameters of USP4 and truncation constructs using a fluorescence assay.
4. Investigating the binding interaction between USP4 and its truncation modified construct with Ubiquitin variant using ESI-mass spectrometry and Isothermal Titration Calorimetry (ITC) assay.

CHAPTER 2

Introduction to General Methods

2.1. Affinity Chromatography

Affinity chromatography is an important separation approach for separating or analysing specific target compounds in samples (Pfaunimiller, et al., 2013). This method is for selective purification of a molecule, or complex mixtures of a group of molecules, based on the highly specific biological interaction between the two molecules. The molecules for example are enzymes, antibodies and other proteins (Urh et, al., 2009; Li et al., 2017).

This method was initially proposed by Porath et al. (1975) using metal ions like Zn (II), Cu (II), Ni (II) and Co (II) towards cysteine, histidine, and tryptophan in aqueous solutions. Proteins can bind specifically to the metal ions coordination sites through certain amino acid residue, but the most used application is a purification of the histidine-tagged fusion protein (Cheung et al., 2012). The main characteristics of His-tag are small, inexpensive to use and have minimal or no effect on the target protein function or structure (Wood, 2014).

The general principle of affinity chromatography is sample application and then the elution of the retained analyte or target compound. A sample that contains the target protein is passed through the affinity column under conditions where the target binds to the immobilized agent. Usually using a buffer which has a pH and composition that will allow strong binding to the immobilized agent in the column followed by awash cycle of any non-retained sample from the column. An elution buffer is then passed through the column

to release the sample or target protein. The protein target then can be collected for further investigation (Li et al., 2017).

The elution buffer may have a different pH, ionic strength, or polarity from the application buffer to make a dissociation of the target protein from the binding agent. Elution of the target protein can be used: a low-pH buffer, a competitive displacement agent or chelating agent. A competitive displacement with Imidazole at the neutral pH is widely used for purifications (Cheung et al., 2012; Li et al., 2017).

The column can be reused again by regenerated using the application buffer and allowing the immobilized-binding agents to return to the initial state. A strong chelating agent (EDTA) will extract the metal ions and disrupts the interaction between protein and ligand. Before the next purification, recharging the column with a metal ion is required (Cheung et al., 2012).

2.2. Size exclusion chromatography

The concept of chromatography size-based separations was prior speculated by Synge and Tiselius in 1950. This concept based on the observation in function of the molecular size. The small molecules could be excluded from the small pores of zeolites (Hong et al., 2012). Size exclusion chromatography is a purification method of the protein which separates molecules based on the size. This method was widely used for purification and analysis of purity,

estimation of molecular weight, and study of interactions: self-association or with other molecules (Irvine, 1997).

The principle of size exclusion chromatography is applied to a small volume of a protein sample to a column containing packing material with pores that are of comparable size to the molecules to be separated. Mobile phase (eluent) is then allowed to pass through the column. The molecular size of solute molecules determines the degree to which it can penetrate the pores. Size exclusion chromatography is essential for analysis and quality control procedures, for example controlling any dimers or conjugates, and commonly used for buffer exchange or desalting. This method is useful for final purification (Hedlund, 2006).

In size exclusion chromatography, the protein sample tends to interact with surface charged sites of the stationary phases. This interaction can result in adsorption of the protein, shifts in retention time, peak tailing, or to change in the 3D conformation of the protein. Reducing the electrostatic interaction can be obtained using increasing the salt concentration or the ionic strength of the mobile phase. Adding the organic modifiers, such as arginine, and manipulate the mobile phase pH could mitigate the secondary interactions. Arginine is preventing the sample to interact with the stationary phase (Hong et al., 2012).

Size exclusion chromatography has three parameters which are used to describe the behavior of the molecule: Vo, Ve, and Vi. Vo is void volume, the volume of interstitial liquid. Molecules with a larger diameter than pore size will completely exclude from the column and elution volume is the same with void volume. Vo usually 30 – 35 % of the total volume (Vt). Ve is elution volume, the volume of eluent collected from the beginning of loading the sample until to the point of its maximal elution. Kd (coefficient of distribution) which represent the fraction of the stationary phase that is available to a given solute, can be described by a function of the void, elution and interstitial volume:

$$Kd = \frac{Ve - Vo}{Vi}$$

Kd always has a value between 0 and 1. Because of the interstitial volume, Vi, is difficult to measure, Kd then replaced by Kav, a coefficient of the partition that can easily be determined for each solute:

$$Kav = \frac{Ve - Vo}{Vt - Vo}$$

Vi, the interstitial volume, was determined by subtracting the void volume from the elution volume. Vt is the total sum of void volume and inner volume (Tayyab et al. 1991; Hedlund, 2006).

2.3. Protein Crystallisation and Structure Determination

Protein crystallisation was firstly published in 1840 on the observation of crystallites in blood preparations by Friedrich Ludwig. In 1864 Felix Hoppe-Seyler confirmed it as hemoglobin which means colorant substance of the blood. Furthermore another family of crystalline protein was found. During that development stage, the procedure to obtain a crystal from protein extraction was warm salt solutions (40-60°C) followed by slow cooling to room temperature (Giege, 2013).

In the early 20th century, the crystallization of other proteins was revealed. Examples are animal and plant globulins (albumins and cananvalin), plant lectin (concanavalin A), some enzymes (catalase, ribonuclease, and urease), toxin diphtheria and hormone insulin. In this period, the researcher noted the importance of salts, organic solvents, pH and temperature for the crystallisation process. The remarkable achievement was in 1934 when John Desmond Bernal and Hodgkin reported the first diffraction pattern of a protein crystal (pepsin) followed by the improvements in crystallization procedure, fabrication of crystal for structure determination and the X-ray crystallography method (Giege, 2013).

The crystallization technique then developed in terms of both process and methodology. Many crystallographers realised that protein crystallisation

could be obtained by chance and art, although some parameter was also important: the solubility of the protein, the type of salt, the temperature, the need for metal ions and the source and amount of protein. Crystallization was no longer limited to isolated proteins but also achieved for protein assemblies, nucleic acid-nucleic acid, and protein complexes (Giege 2013). In the future, protein crystallography and its application may aid the identification of better candidate molecules that are more amenable to high-concentration processing, formulation, and analysis to make quicker, simpler and cheaper of biologics drug development (Brader et al., 2017).

2.3.1. Growing crystals

Crystallisation of a protein often occurs by chance. The importance for crystal formation is growing a high quality well-ordered protein crystal because only these crystals can be useful for determining the molecular structure of the protein (Yan et al., 2016).

Growing protein crystals is initiated by dissolved purified protein in an aqueous buffer containing a precipitant, such as ammonium sulfate or polyethylene glycol, at the concentration just below to precipitate the protein. The water is then removed by controlled evaporation to raise the concentration of protein and precipitant. The precipitation will happen when the combination of protein concentration and precipitant concentration exceeds threshold values.

Formation of the crystal will occur in two stages, nucleation and growth. The initial formation of molecular clusters from which crystals grow, called nucleation, need the concentration of protein and precipitant to be higher than the optimal concentration for slow precipitation (Figure 2.1). If these conditions persist, this will result in the formation of many nuclei that will form an amorphous precipitate or many small crystals. The best strategy starts with a condition corresponding to the nucleation and growth region and then move until nuclei form and growth is observed (Rhodes, 2006; Ilari and Savino, 2017).

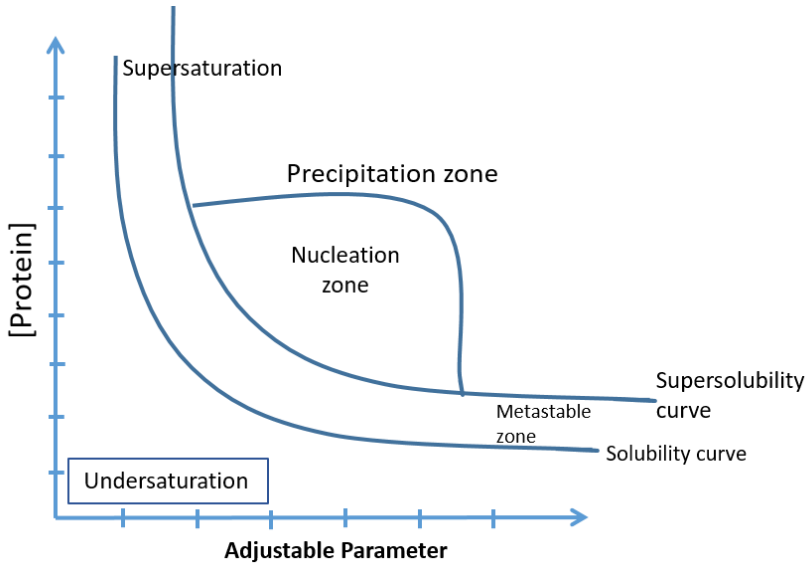


Figure 2.1. Illustration of a protein crystallisation phase diagram (Adaptation from (Adaptation from Chayen, 2005)
In the area of very high supersaturation, the protein will precipitate. Spontaneous nucleation will take place in the area of moderate supersaturation. In the area of lower supersaturation (under nucleation zone) crystals are stable and may grow without further nucleation, this is the best condition for the growth of large, well-ordered crystal; the protein fully dissolved and never crystallise in the undersaturated area

Crystal nucleation develops in two stages: nucleation of new crystal embryos and growth of few nuclei into full-size diffracting crystals. For optimal growth of the crystal, changing the concentration of the precipitation agent, pH or

temperature, the concentration of buffer, and crystallisation technique may be needed (Wlodawer et al., 2017; Smyth and Martin, 2000). Also, the protein needs to be of high purity and yield, homogeneous and be for crystallisation.

A high-quality protein crystal is well ordered, has a good size and diffracts well in the X-ray beam. A number of strategies focus on obtaining a high-quality crystal. The first method is the classical method, such as hanging-drop, sitting drop, and, dialysis. The second method is an advanced method using kinetic control (electric fields, magnetic fields, nucleation inductors, and degree of supersaturation) or transport control (growth in gels, thin capillaries, microgravity, and microfluidics) (Wlodawer et al., 2017).

The most common method of protein crystallisation is hanging drop vapour diffusion and sitting drop method. A concentrated protein solution is combined with a solution of a precipitant and will be concentrated by evaporation. A protein crystal will form in the right condition. Conditions from the initial screen which show the most suitable parameters for crystallisation should be optimised to improve crystal size and form (Rowlett, 2005) .

In the hanging drop or sitting drop method is, a drop containing the protein and precipitants is equilibrated against a reservoir (Figure 2.2). The reservoir then draws water from the drop, driving the protein/precipitant solution mixture to a higher concentration state that enters the supersaturated region

(Gavira, 2016). This method is bringing the protein directly into the nucleation zone by mixing protein with the appropriate amount of precipitant (Dessau and Modis, 2011).

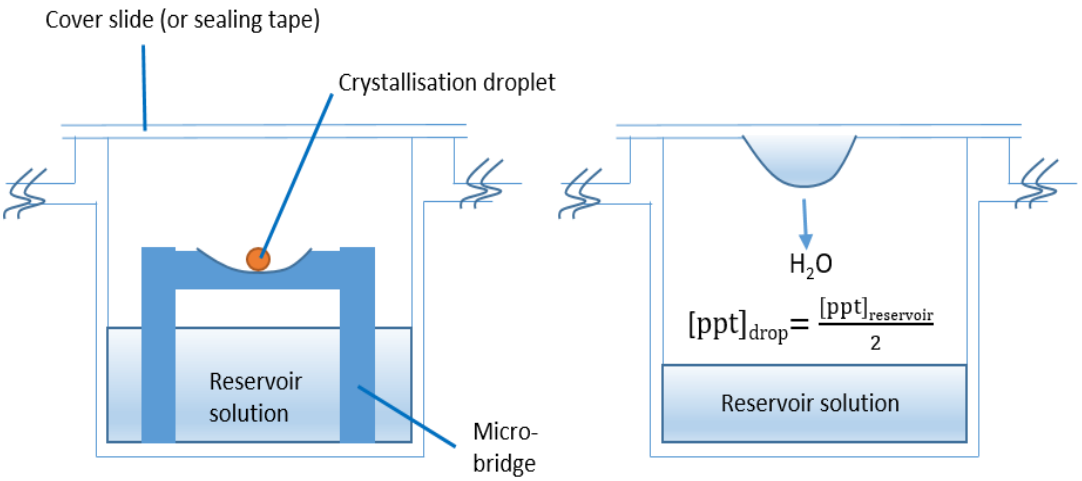


Figure 2.2. Sitting drop vapour diffusion (left) and hanging drop vapour diffusion (right) method in crystallisation (Adaptation from McPherson and Gavira, 2014).
Sitting-drop vapour-diffusion: the drop on the elevated platform, usually 2-10 µL, consists of half stock protein solution and half of reservoir solution with some concentration of polymer precipitant or salt. Before sealing, around 0,5 mL of reservoir solution added to the bottom of the cell. The drop ultimately reaches the reservoir in osmolarity by water equilibration through the vapour phase. Hanging-drop vapour-diffusion: all the same with sitting-drop, except the protein drop is suspended for a cover slip over the reservoir rather than resting on the surface.

Another method for crystallisation trial used the inhibitor agent. The inhibitor agent was expected to bind or interact with the protein and alter the conformation and solubility to form a crystal (Benfenuti and Manganic, 2007). In this research the inhibitor for USP4 was used, namely Vialinin and mitoxantrone.

Protein crystallisation involves three main steps: 1). Determination of the degree of protein purity; 2). Dissolving the protein in a suitable solution; 3). The

solution is brought to super-saturation which actual crystal growth begins (Ilari and Savino, 2017).

Some methods for creating supersaturations are altering the protein itself (by the change of the pH which alters the ionisation state of surfaces amino acid residues). Moreover, also altering the chemical activity of the water (by adding salt); altering the nature of the interactions between the protein molecules and the solvent (adding polymers and ions). Precipitants also needed for optimizing the crystallization, for example, salt, an organic solvent, long chain polymers and low molecular weight polymer and nonvolatile organic compounds (McPherson, 2004).

Many factors are affecting crystallisation. Apart from pH, salt, and the concentrations of precipitants, other factors are the presence or absence of ligands or inhibitors, a variety of salt or buffer, the temperature, or the presence of detergent (McPherson, 2004).

Nowadays, robot and other automated instruments, and a complete integrated system have been developed to increase the throughput of the crystallisation process. They can prepare thousands of screen crystal conditions and performed it with precisely and reliably, less error and better (McPherson and Gavira, 2014).

2.3.2. Cryocooling, X-ray diffraction and structural determination

Once a protein crystal is ready, cryo-cooling is generally needed. The cryo-crystallography technique is performed to reduce the rate of radiation damage. This is essential in data collection and has a vital impact on the data quality obtained and the number of a protein crystals needed. With cryo-cooling a whole data set can often be collected from one crystal with a higher quality of the data. The crystal is mounted in a loop which is less harmful to fragile crystals, and the crystal can then be flash-cooled in liquid nitrogen (Garman and Owen, 2006).

Cooling the protein crystal should avoid the formation of crystalline ice because will be disrupted internal order and interfere the diffraction pattern of the protein crystal. There are two types of cryoprotectant which are mostly used: glycerol (20 – 30 % v/v depending on the condition) and oil in which the crystal is coated. Another agent for cryoprotection can depend on the major component of the mother liquor, such as low molecular weight PEG, MPD, and ethylene glycol (Garman and Owen, 2006).

The protein crystal is then investigated using X-rays. The crystals are exposed to an X-ray beam to collect a diffraction pattern. This process is called X-ray diffraction. X-ray diffraction is caused by the interaction of electromagnetic waves with the matter inside the crystals, and particularly with the electrons.

The diffraction patterns are then collected of a rotating crystal to get a full data set (Smyth and Martin, 2000). The next step after data collection is the processing of the data. Several programs are available for data processing, such as Mosflm, XDS, and HKL-2000 (Karadaghi, 2015). The structure can then be solved by heavy atom (if no similar structure is available) or molecular replacement methods (if a similar model structure is already available).

Molecular replacement (MR) is algorithmic approach uses the structure of a related protein which positioned within the unit cell of the new crystal to obtain the initial phase. For the protein without known homologues, the initial phase achieved by adding Heavy atoms (HA). HA is any atom with more electrons than those normally found in protein (hydrogen, carbon, nitrogen, oxygen and sulfur) (Pike et al, 2016). For example: the residue of the amino acid methionine is replaced by Selenomethione. Selenium replace the usual sulfur of methionine. This substitution provides selenium as built-in heavy atoms and usually did not alter protein conformation or unit cell structure (Rhodes, 2006).

2.3.3. Anomalous scattering

X-rays set the electronic charges around atoms vibrating and generate radiation of the same frequency and propagated in all directions. This is causing diffraction effects. The electron usually vibrates in phase with the incident X-ray beam. However, if the incident photon has an energy near to

the transition energy which could move the atom to an excited state, the electronic vibration gets out of step. The re-radiating energy is having a different phase. Moreover, the intensity of coherent scattering is reduced due to some energy is absorbed to bring about the transition. This effect is known as anomalous scattering (Blow, 2002).

The first evidence for anomalous X-ray scattering was from Mark and Szilard (1925) who showed that diffraction from RbBr crystal differed selectivity depending on the wavelength of irradiating X-rays. The RbBr structure is like rock salt and as Rb+ and Br- ions are isoelectronic and have equivalent scattering factor. Anomalous scattering thereby precisely identifies atomic species which enables the multi-wavelength anomalous diffraction (MAD) and single-wavelength anomalous diffraction (SAD) methods (Hendrickson, 2014).

2.4. Electron spray ionization - mass spectrometry

2.4.1. Introduction to mass spectrometry

Mass spectrometry (MS) is a central analytical technique for protein research and the study of biomolecules (Domon and Aebersold, 2006). MS was first introduced by JJ Thomson, 1912, for isotopic analysis. Mass spectrometry measures the mass-to-charge ratio (m/z) of an ion. The analytes are first ionised and transferred into the gas phase and then separated according to

their m/z ratio in the mass analyser. The detector senses the ion that emerges from the mass analyser (Sins et al., 2015).

Two decades later, mass spectrometry has emerged as the most efficient and versatile tool (Mirza and Olivier, 2008) for protein identification (Baldwin, 2004). Mass spectrometry made polypeptide accessible to analysis rapidly after the development of two techniques: electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI). This method is used to measure the molecular mass of a polypeptide or to determine additional structural features (amino acid sequence or site of attachment and type of posttranslational modification). It became possible to analyse mass spectra of proteins, DNA/RNA, carbohydrates, lipids and polymers (McLafferty, 2008). The development of the method was supported by the development of new mass analyser and complex multi-stage instruments, such as hybrid quadrupole time-of-flight (Q-Q ToF) and tandem time-of-flight (ToF ToF) instruments (Domon and Aebersold, 2006). This development allowed MS to have unique capabilities of specificity, sensitivity, speed, sampling, and automated computer data acquisition (McLafferty, 2008).

2.4.2. Electrospray ionisation mass spectrometry (ESI-MS)

Electrospray is one of the soft ionization methods that will provide information about molecular masses of large biomolecules. The first measurement of

protein by ESI-MS was successful at 1989 by Fenn and coworkers, then followed by the development of another application for characterization of amino acid sequence of protein and peptides (Akashi, 2006). ESI mass spectrometry is an approach based on electrospray ionisation (Leney and Heck, 2017). The sample or biological analytes are sprayed from a non-denaturing solvent. The schematic illustration of ESI shows that electrospray generates ions at atmospheric pressure by spraying sample solution from a metal needle with high voltage (+3 to 4 kV) (Figure 2.3).

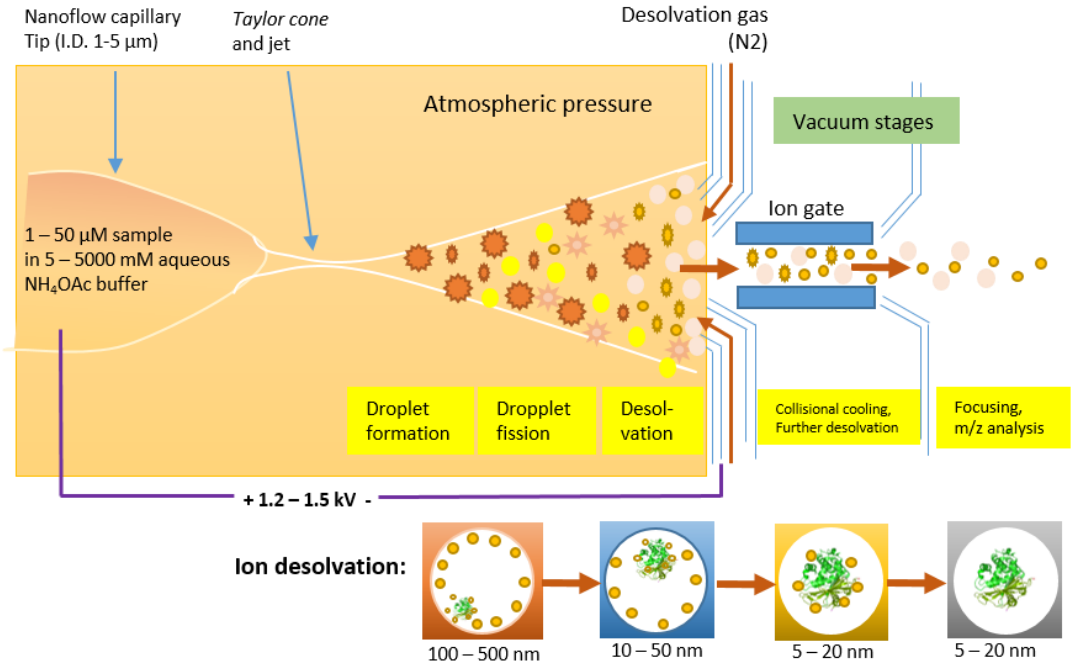


Figure 2.3. Schematic ionisation process on mass spectrometry
(Adaptation from Cubrilovic, 2014)

In spraying process, heated nitrogen gas will help the sample droplets formation. The ionic sample is then transported from the ionization source to

the mass spectrometer under high vacuum through a glass capillary and then mass to charge ratio is analysed (Akashi, 2006; Konijnenberg, et al., 2013).

The protein sample should be in a solution condition which is: neutral pH and without organic solvent. This condition can avoid the denaturation of the macromolecular assemblies. Volatile salts as a buffer such as ammonium acetate or ammonium bicarbonate are used (Akashi, 2006).

Native mass spectrometry is an emerging technology for the analysis of protein complexes. Native MS is a useful tool for refining structural models (Heck, 2008). The molecular mass shown in the mass spectrum can confirm the presence of particular species and provide the information on the stoichiometry of complexes. This data can determine the number of binding sites of a protein and their ligands and for calculating the dissociation constant (K_D) of the complex (Rossi and Taylor, 2011).

The binding of ligands, such as cofactors, nucleotides, lipids, or drug molecules has a characteristic mass shift of protein and protein complex. Ligands are usually associated noncovalently. Their interaction can be investigated by native MS and can reveal the data of the binding stoichiometry, affinity, and cooperativity (Lossi et al., 2016).

2.4.3. Ion Mobility Mass Spectrometry

Ion mobility mass spectrometry (IM-MS) began with the work by Mc Daniel in the late 1950s and 1960s when developed an IMMS instrument to study ion molecule reactions of noble gas and pure hydrogen (Jiang and Robinson, 2013). IM-MS is an analytical technique to separate gas-phase ions based on their shape and size (Lanucara, et al., 2014).

IM-MS is multiple MS measurement using ion mobility separation. IM-MS measure the time for protein (or its various populated structural states) to transverse a weak electrical gradient in a gas-filled chamber. IM-MS can provide information on the stoichiometry, composition, protein contacts and topology of protein complex (Ben-Nissan and Sharon, 2018).

There were three IM-MS technique are used, namely: drift-time ion mobility spectrometry (DTIMS), traveling-wave ion mobility spectrometry (TWIMS) and field asymmetric ion mobility spectrometry (FAIMS). IM-MS can be used to ascertain structural information using sample in small amounts (quantity in nanogram). This technique can determine: conformational dynamics, folding and unfolding intermediates, ligand-induced conformational changes, aggregation intermediates and quaternary structure (Lanucara, et al., 2014).

2.5. Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry is an ideal technique for measuring biological binding interactions. This technique does not need the presence of chromophore or fluorophores or other modification (Pierce et al., 1999). ITC is based on the detection of a heat effect upon binding and used for determining the binding constant (K), the stoichiometry (n), and the enthalpy of binding (ΔH) (Lewis and Murphy, 2005).

Isothermal titration calorimetry (ITC) was first described as a method for simultaneous determination of K*eq* and ΔH by Christansen and Izzat (Christensen et al., 1966). ITC is now routinely used to directly characterise the thermodynamics of biopolymer binding interactions (Freire et al., 1990).

Modern ITC instruments are often employed in a batch or a direct injection mode in which injection of a reagent solution is made to start the experiment. ITC instruments are used in a titration mode in which several incremental injections are made at time intervals during a complete titration experiment. The instrument has high enough sensitivity and fast enough response for most biological/biochemical studies (Figure 2.4) (Freyer and Lewis, 2008).

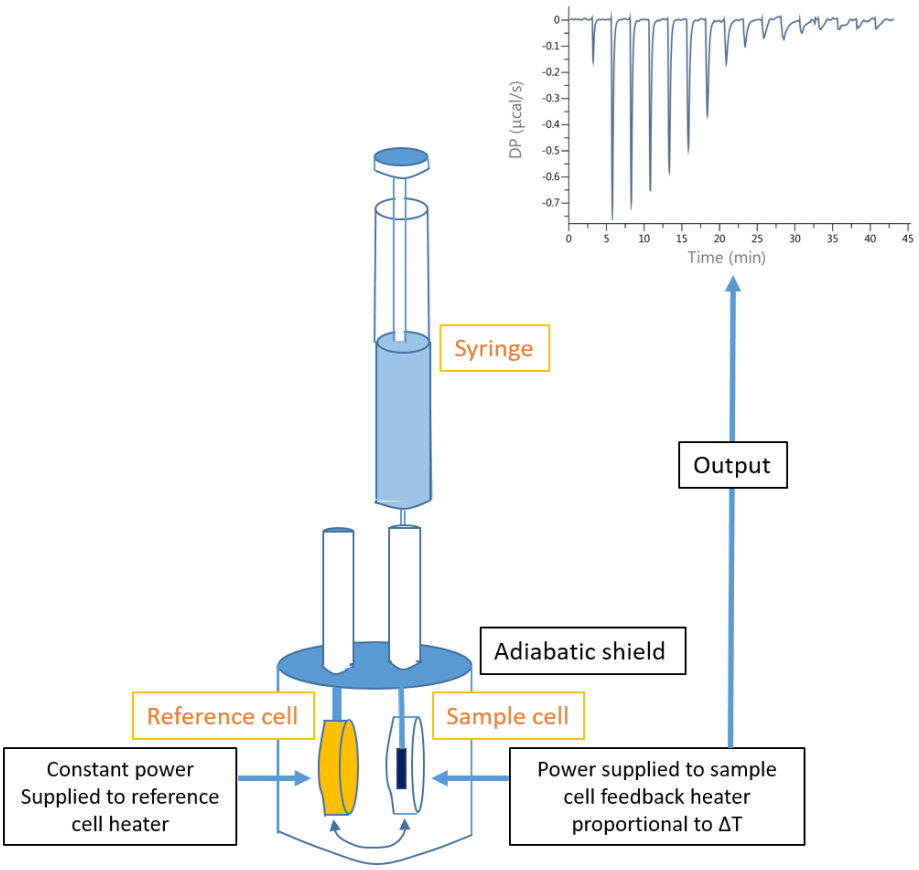


Figure 2.4. Diagram of a typical power compensation ITC.
This diagram shows how power applied by the instrument to maintain a constant temperature between the reference and sample cells is measured resulting in the instrument signal (Adaptation from Freyer and Lewis, 2008).

A typical ITC experiment includes three steps: first, a ligand is titrated into a solution containing the bio-macromolecule (protein) of interest. The second is binding event measured, which was associated with the heat released or absorbed. Moreover, in a third step ITC data are processed (Kd, ΔG, ΔH, ΔS and stoichiometry (n)). ITC is a nondestructive technique, which can be determined the thermodynamic parameter of the interaction without immobilisation, modification, and without molecular weight restrictions (Du et al., 2016).

2.6. Ubiquitin cleavage Enzymatic Assay

Enzymatic assays follow the reaction of substrate consumption or product release typically using pure enzymes. The most practical enzymatic assay is based on synthetic substrates which release a fluorescent or coloured product upon reaction in solution. Many substrates are commercially available and can be used to determine the activity of the enzyme in units (Reymond et al., 2009). For kinetic measurement of enzyme activity in vitro, fluorogenic substrates have been used for a long time. These reagents bring fluorescent groups and the energy emission for their enzymatic conversion to product monitored over time (Baruch et al., 2004).

Deubiquitinating enzymes (DUB) consist of a large family of mostly cysteine hydrolases which cleave ubiquitin-derived substrates of the general structure Ub^{1-72}-Leu^{73}-Arg^{74}-Gly^{75}-Gly^{76}-X, specifically. X can be any number of leaving groups: small thiols and amines, Ub or proteins (Dang, et al., 1998). Fluorescence assays have been developed using Ub-AMC (Ub-amido methylcoumarin) as a substrate for DUB activity. The rate of release of the fluorescent tag from the substrate can be quantified using a fluorescence spectrometer to calculate the enzymatic activity of a DUB (Russel and Wilkinson, 2005). Ub-AMC has high fluorescence sensitivity and hydrolysis efficiency and often used for determined the activities of DUB quantitatively

(Yin et al., 2011). When ubiquitin is cleaved an increase in fluorescence is observed as AMC is released.

CHAPTER 3

Design constructs and Research Outline

3.1. Constructs used

Sixteen USP4 constructs were engineered to elucidate the structure, study the activity, and interaction of USP4 with ligands. All the constructs of USP4 were designed based on the USP4 domain, namely USP4 full-length (USP4FL), USP4 head to tail (USP4htt) as circular permutated version, USP4 core protease domain (USP4C1C2), USP4 head to tail without Ubl2 (USP4httΔUbl2), USP4 without the DUSP-Ubl1 domain (USP4ΔDU) and USP4 DUSP-Ubl1 domain only (USP4DU) (Figure 3.1).

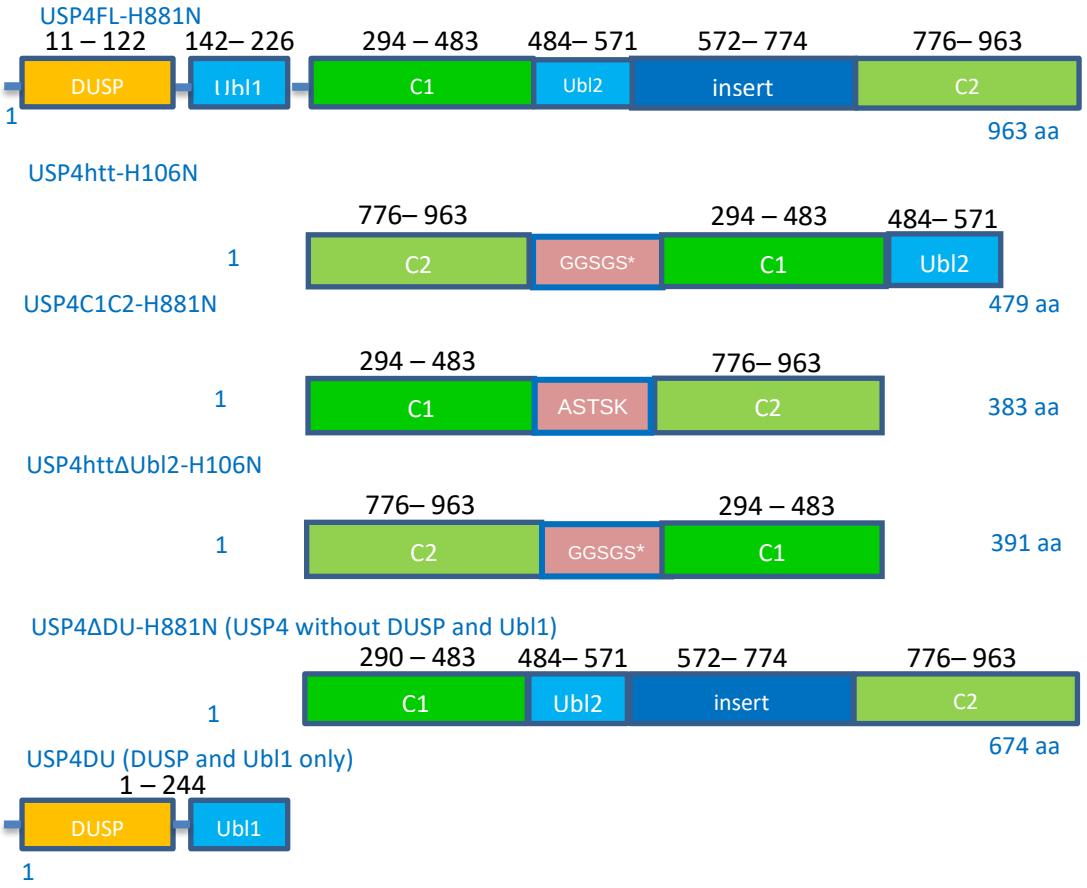


Figure 3.1. Schematic representation of USP4 constructs cloned
*linker consist of = GGSGSGAGSGGSG

USP4FL construct consist of full-length USP4 containing DUSP, Ubl1, the Catalytic domain (C1 and C2), Ubl2, and insertion. USP4htt was designed as circular permutated version starting from residue number 776 to 963 (catalytic-2 sub-domain) and connected by linker join the residue number 294 to 571 (catalytic-1 sub-domain and Ubl2 domain). This construct design by creating a new N-terminus starting from catalytic-2 sub-domain, residues 776. Subsequently, the original C-terminus connected to the original N-terminus by a flexible linker. As a result, the catalytic-2 sub-domain is connected by a linker to the catalytic 1 sub-domain and Ubl2 domain. The Ubl2 forms the new C-terminus after removal of the insertion. USP4C1C2 construct consists of catalytic 1 and catalytic 2 sub-domain only, connected by the linker ASTSK, lacking DUSP, Ubl1, Ubl2, and insertion. USP4httΔUbl2 is the same as USP4htt but without an Ubl2 domain or the same with USP4C1C2 in the reverse direction with the different linker. USP4ΔDU consists of catalytic-1 sub-domain, Ubl2 domain, insertion and catalytic-2 sub-domain. Moreover, the last, USP4DU consists of DUSP and Ubl1 domain only.

In this research, the active site mutant of USP4 construct also created (Figure 3.2. & Figure 3.3). Site-directed mutagenesis was performed to determine the structure of USP4 in complex with the Ubiquitin variant by changing its catalytic Histidine (H) residue number 881 to Asparagine (N) on the catalytic 2 sub-

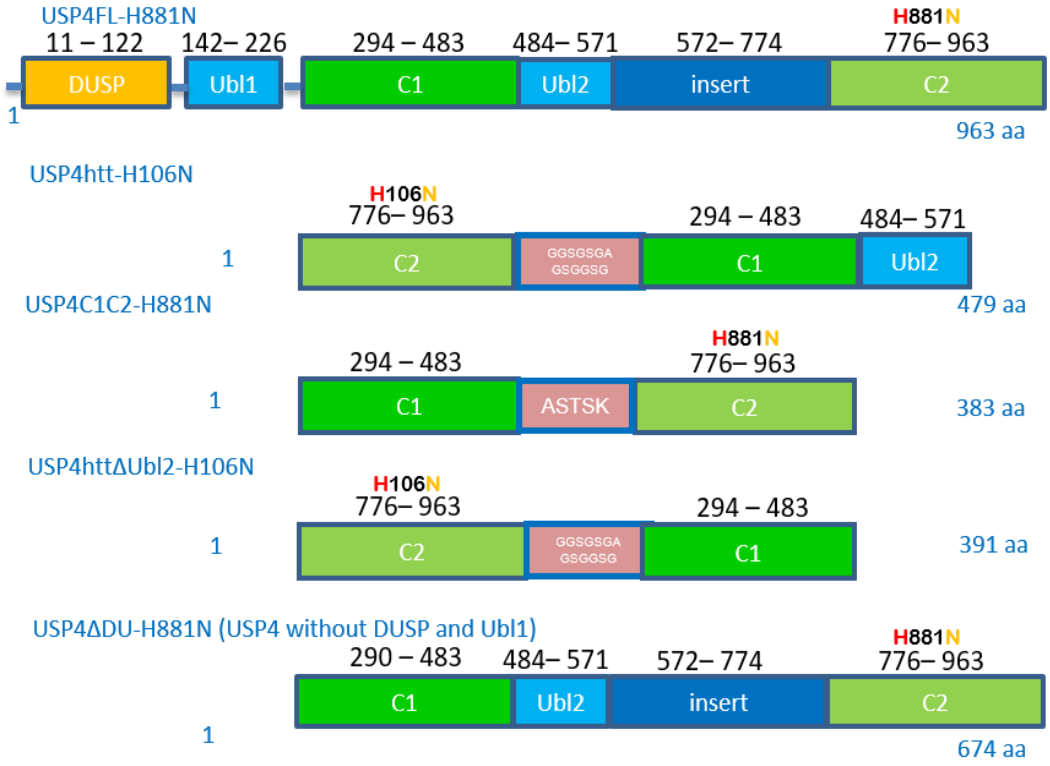


Figure 3.2. Schematic representation of active site mutant H881N of USP4 constructs cloned

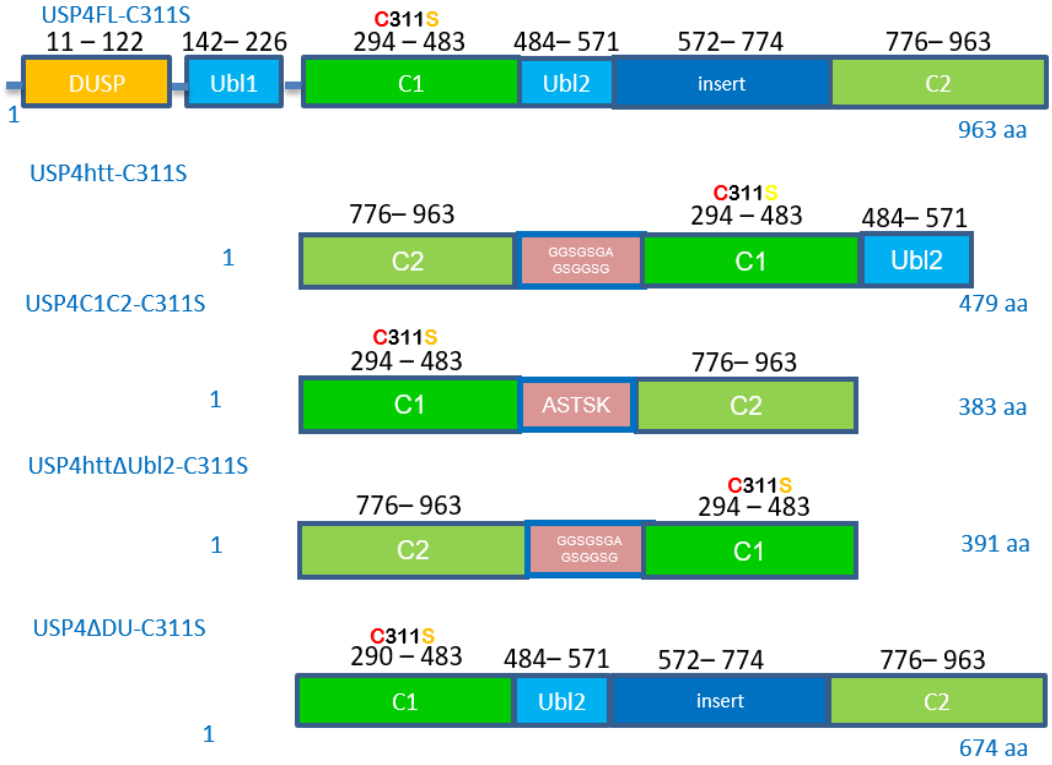


Figure 3.3. Schematic representation of active site mutant C311S of USP4 constructs cloned

domain. Another mutation was changing Cysteine (C) residue number 311 to Serine (S) on the catalytic 1 sub-domain. In all constructs USP4FL, USP4C1C2, USP4htt, USP4httΔUbl2 and USP4ΔDU, the active site mutation H881N and C311S were generated (Table 3.1).

Table 3.1. Summary of active site mutant USP4 constructs design used

Active site mutant	Constructs
Histidine (H) to Asparagine (N)	USP4FL-H881N
	USP4htt-H106N (originally 881)
	USP4C1C2-H881N
	USP4httΔUbl2-H106N
	USP4ΔDU-H881N
Cysteine (C) to Serine (S)	USP4FL-C311S
	USP4htt-C311S
	USP4C1C2-C311S
	USP4httΔUbl2-C311S
	USP4ΔDU-C311S

Both active site mutants H881N and C311S of USP4 combine with Ubiquitin variants, such as Ubiquitin, linear-diubiquitin, Ubiquitin-GGG, and diubiquitin-L73X to form a complex and investigate for crystallisation trial and binding interaction. The binding interaction assay using ESI-Mass spectrometry and Isothermal Titration Calorimetry (ITC).

3.2. Research Outline

The schematic of research outline to elucidate the structure of USP4 show in Figure 3.4.

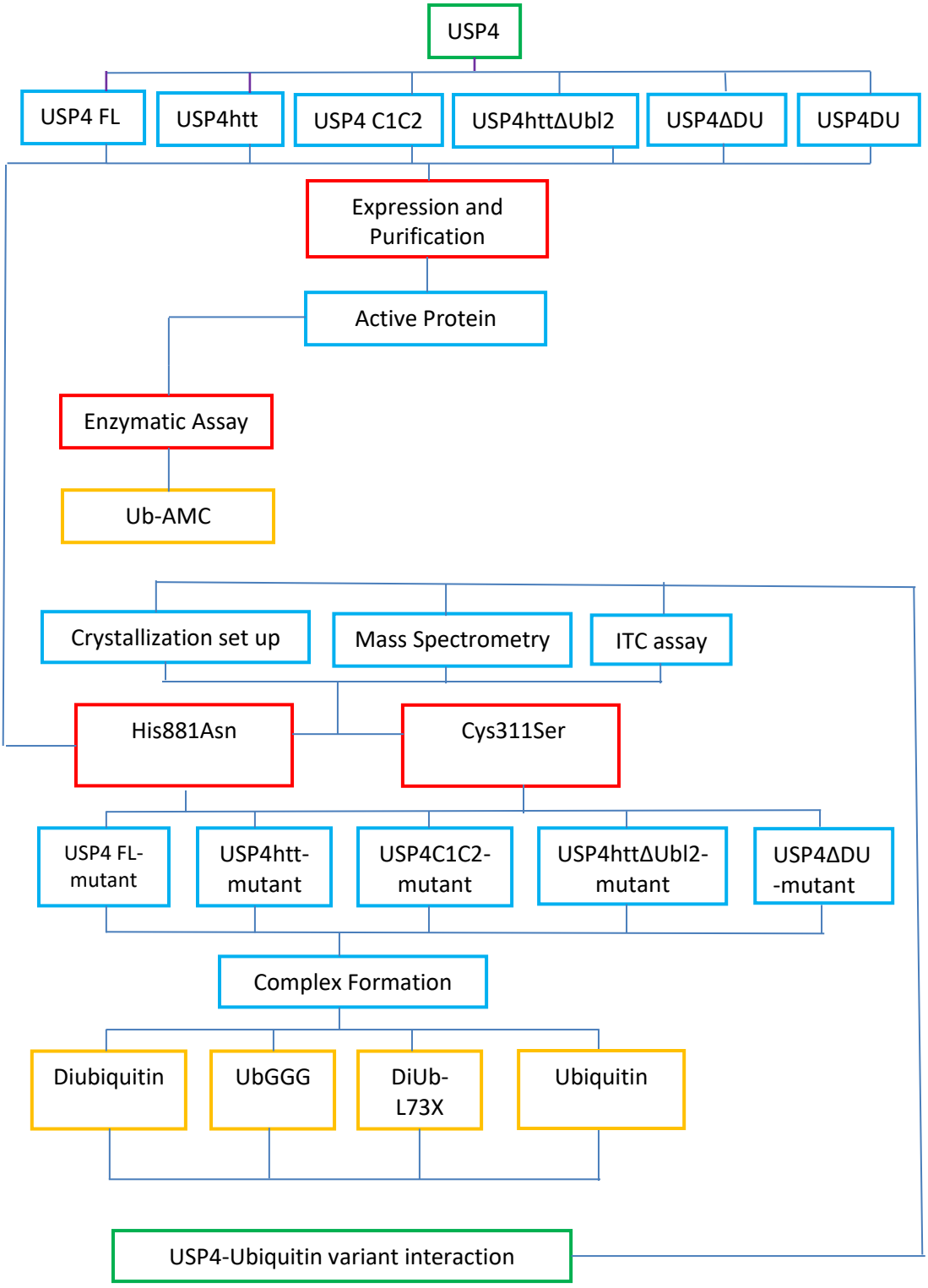


Figure 3.4. Outline of research for elucidating structure of USP4

CHAPTER 4

Material and Methods

4.1. Construct design and cloning

4.1.1. PCR

All constructs were designed by performed PCR using specific forward primer and reverse primer, except USP4DU and USP4ΔDU (Table 4.1). The PCR protocol was running at 95°C for 5 minutes; 95°C for 30 seconds, 60°C for 30 seconds, 72°C for 1 minute for 25 cycles; then 72°C for 10 minutes, continue with the PCR clean-up kit using Sigma-Aldrich and eluted with 50 µL H_2O. The PCR products viewed on 1% agarose gels. Marker 1 kb and 500bp DNA ladder (NEB) was used to determine the PCR product. The chosen PCR product continue to double digestion. Double digestion reactions were used to cut both the PCR product and vector. For digestion reaction, the PCR product digestion contained 16 µL PCR products, and the vector digestion reaction contained 8–16 µL of vector depending on the vector that used. Both incubated at 37℃ for 4 hours. The vector then adds by Antartic phosphatase to prevent re-ligation itself. 1 µL Antartic phosphatase and 2 µL of Antartic phosphatase buffer was added and incubated for 2 hours at 37°C. The vector and PCR product continue to clean up (Gen Elute PCR CleanUp Kit) and ligation process. All the vector of the construct contains Histidine tag. The forward and reverse primer for USP4FL, USP4htt, USP4httΔUbl2, and USP4C1C2 was summarized in Table 4.2 and Table 4.3.

Table 4.1 Summary of USP4 constructs design used

Constructs	Residue number	Restriction site used	Tag	Vector used
USP4-FL	1-963	KpnI - HindIII	N-terminal His tag	pColdI
USP4-htt	776-963 294-571	NdeI - XhoI	C-terminal His tag	pET26b
USP4-C1C2	294-483 776-963	BamHI - HindIII	N-terminal His tag	pProEXHtb
USP4-httΔUbl2	776-963 294-483	NdeI - XhoI	C-terminal His tag	pET26b
USP4-ΔDU*	290-963	BamHI - HindIII	N-terminal His tag	pProEXHtb
USP4-DU*	1-244	NdeI - XhoI	C-terminal His tag	pET26b

*Cloned by Stephen Harper

Table 4.2. Summary the primer for USP4 constructs used

Construct	Forward primer	Reverse primer
USP4FL	GGGGTACCGCGGAAGGTGGAG GCTGC	CCCAAGCTTTTAGTTGGTGTCCATGC TGCA AG
USP4htt	GGAATTCCATATGCACCATCATCATC ATCATAC	CCGCTCGAGTCAGGTGCTGCACAC
USP4httΔUbl2	GGCCATATGCACCATCATCATCAT CATAC	ATGCCTCGAGTCACAGCGG CAGCGG CAG

Table 4.3. Oligonucleotide primers used for the USP4C1C2 construct

Domain	Amino acid residues	Forward Primer (5′-3′)	Reverse primer (5′-3′)
Cat 1 sub-domain	295 (BamHI site) – 478	ATGGATCCTCTCATATACAACCT GGGCTCTGTGG (Primer 1)	TTTACTTGTGGATGCCAGTGGCA GCGT (Primer 2)
Cat 2 sub-domain	776-957 (HindIII site)	GCATCCACAAGTAAAGTGGCCC TGAGA (Primer 3)	GCTAAGCTTCTAGTTGGTGTCCA TGCTGCAAGC (Primer 4)

4.1.1.1. Overlap extension PCR

Overlap extension PCR was used to create the catalytic core of USP4 (USP4C1C2) with four primers (Figure 4.1). First, the two PCR products were prepared using forward primer 1 with reverse primer 2 (as PCR 1 to create catalytic 1 domain) and forward primer3 with reverse primer 4 (as PCR 2 to create catalytic 2 domain). For this step, the method using as described above at 4.1.1. After PCR clean up and confirmation the concentration of PCR product and the size, the following step was second PCR as annealing using PCR 1 and PCR 2 products at temperature 60℃ with 15 cycles without any primer, using Pfu Ultra II enzyme. The third PCR was using forward primer 1 and reverse primer 4 to develop the complete construct of USP4C1C2.

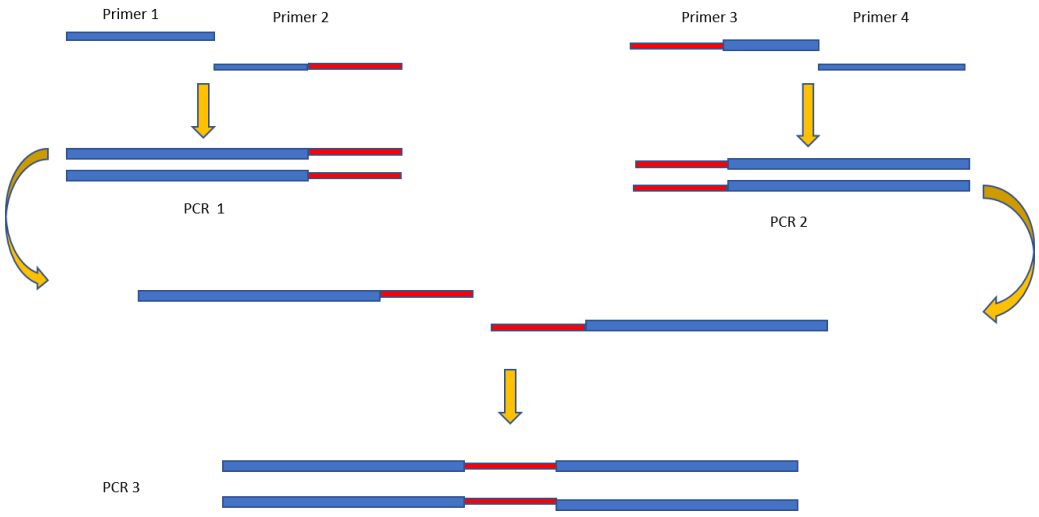


Figure 4.1. Schematic overlap PCR for USP4C1C2

4.1.2. Ligation

Ligation was performed with 1 µL T4 ligase, 1 µL T4 ligase buffer, PCR product (as an insert), and vector and made up to 10 µL. The mixture was incubated at

16°C for overnight. The volume of the insert was calculated based on the concentration of insert and vector, and kb of insert and vector by this equation:

$$\text{Volume insert} = \frac{\frac{[vector]\ x\ Kb\ insert}{Kb\ vector}}{[\text{insert}]}$$

4.1.3. Transformation

The reaction after ligation followed by transformation of the ligation mixture. Two µL of ligation mixture incubated with 50 µL Nova Blue cells on ice for 15 minutes followed by heat shock at 42°C for 1 minute, then kept on ice for 1 minute. 800 µL LB was added in sterile conditions and incubated for 45 minutes at 37°C at shaking incubator 180 rpm. The last step was to plate 200 µL of the culture on specific antibiotic LB plate and incubated for overnight at 37°C.

The antibiotics Ampicillin (100 µg/mL) was added into USP4FL and USP4C1C2 construct. For USP4htt and USP4httΔUbl2 constructs were used Kanamycin (50 µg/mL).

4.1.4. Cracking Digestion Test

Confirmation of transformation result done by cracking test. The single colony after transformation was picked off and put in 5 mL LB media with antibiotic for overnight incubation at 37°C and 180 rpm. These cultures checking by cracking using 20 µL of culture into 5 µL 5X cracking buffer (5 nL 5 M NaOH, 25 g sucrose, 2.5 mL 10% SDS, in 50 mL H_2O, and bromphenol blue). The sample runs on the agarose gel electrophoresis. Sample with the insert should appear higher compared to the control because has more molecular weight.

The chosen sample then purified using Plasmid Miniprep Kit (Sigma-Aldrich). After miniprep, the sample was digested to check that the insert was ligated into the vector. The chosen result from digestion test sent for sequencing (Source Bioscience). After analysing the sequencing data with BLAST program, the chosen sample was transformed into the BL21 cells and was plated in the Chloramphenicol and another selected antibiotic based on the vector. Especially for USP4FL, the ubiquitin was added to the transformation in BL21. The colony from the transformation result was incubated in the LB with the selected antibiotic overnight and then transferred to 20% glycerol LB in -80°C for storage.

4.1.5. Site-directed mutagenesis

Site-directed mutagenesis was used to change the DNA sequence of a plasmid to mutate the amino acid sequence of the protein product. This method was used to mutate the His881Asn on catalytic 2 sub-domain, Cys311Ser on catalytic 1 sub-domain, mutated Lysine residue number 73 to stop codon, and mutated Tyrosine on the residue number 928 of catalytic 2 subdomain on C terminus to stop codon.

The protocol for mutagenesis followed the Stratagene Quick Change Site-Directed Mutagenesis Kit. The reaction consisted of 5µL 10x Pfu Ultra II Buffer, 0.5 µL of DNA template (5 ng/µL), 0.5 µL forward primer, 0.5 µL reverse primer, 1 µL dNTP and 42.5 µL H_2O. Then after mixing, 42.5 µL of Pfu Ultra II was added.

The PCR for mutagenesis was 95˚C for 30 seconds; 95˚C for 30 seconds, 55°C for 60 seconds, and 68˚C for 5 minutes. The setting was cycled 16X after that was 68˚C for 10 minutes. DpnI restriction enzyme was added and incubated at 37ºC for 2 hours. After mutagenesis, following by transformation. After overnight incubation at 37°C, some single colonies were chosen and performed DNA miniprep isolation (Sigma-Aldrich). Then, the sample was sent for sequencing (Source Bioscience) to confirm that the mutagenesis worked. Ten constructs of USP4 mutated for His881Asn and Cys311Ser. The primers used for mutagenesis showed on Table 4.4.

Table 4.4. Primers for mutagenesis

Constructs	Mutation	Forward primer	Reverse primer
USP4FL	His881Asn	GGAGCCATGGGGGTTGGCAAC TACACTGCATATGCGAAGAAC	GTTCTTCGCATATGCAGTGTA GTGCCAAC CCCCATGGCTCC
USP4htt	His106Asn	GGCGCGATGGGCGTGGGCA ACTATACCGCGTATGCG	CGCATACGCGGTATAGTT GCCCACGCCCATCGCGCC
USP4C1C2	His881Asn	GGAGCCATGGGGGTTGGCAAC TACACTGCATATGCGAAGAAC	GTTCTTCGCATATGCAGTGTAGT GCCAACCCCCATGGCTCC
USP4httΔUbl2	His106Asn	GGCGCGATGGGCGTGGGCAACTA TACCGCGTATGCG	CGCATACGCGGTATAGTT GCCCACGCCCATCGCGCC
USP4ΔDU	His881Asn	GGAGCCATGGGGGTTGGCAACT ACA CTGCATATGCGAAGAAC	GTTCTTCGCATATGCAGTGTA GTTGCCAAC CCCATGGCTCC
USP4FL	Cys311Ser	GGAAACCTGGGAAACACCAGCT TCATGAACTCCGCT	AGCGGAGTTCATGAAGCTGGTG T TTCCCAGGTTTCC
USP4htt	Cys311Ser	GGCAACCTGGGCAACACCA GCTTTATGAACAGCGCG	CGCGCTGTTCATAAAGCTGGT GTTG CCCAGGTTGCC
USP4C1C2	Cys311Ser	GGAAACCTGGGAAACACCA GCTTCATGAACTCCGCT	AGCGGAGTTCATGAAGCT GGTGTTTCCCAGGTTTCC
USP4httΔUbl2	Cys311Ser	GGCAACCTGGGCAACACCA GCTTTATGAACAGCGCG	CGCGCTGTTCATAAAGCTGGT GTTG CCCAGGTTGCC
USP4ΔDU	Cys311Ser	GGAAACCTGGGAAACACCA GCT TCATGAACTCCGCT	AGCGGAGTTCATGAAGCTGGT GT TTCCCAGGTTTCC
Diubiquitin	Lys73Stop	CACTTGGTCCTGCGCTAGA GGGGGGGTTAATAAGTTTCC	GGAAACTTATTAACCCCCC CTCTAGCGCAGGACCAAGTG
USP4FL	Cys311Ser and Tyr928Stop	CGTCGAGATGATGAATTTT AGAAGACACCTTCACTTAG	CTAAGTGAAGGTGTCTTCTA AAATTCATCATCTCGACG
USP4C1C2	Cys311Ser and Tyr928Stop	CGTCGAGATGATGAATTTT AGAAGACACCTTCACTTAG	CTAAGTGAAGGTGTCTTCTA AAATTCATCATCTCGACG

4.1.6. Co-expression with Ubiquitin variant

To obtain the protein crystal, a complex formation between active site mutant USP4 modified protein with the Ubiquitin variant were performed, namely co-expression (Table 4.5). Some of this complex were continued for ESI-mass spectrometry analysis (Sub Chapter 4.8). Two µL of the active site mutant USP4 modified construct and 1 µL of selected Ubiquitin variant (Ubiquitin, linear diubiquitin, Ub-GGG, and diubiquitin-L73X) were added in the 50 µL of BL 21 Codon plus cell and continued the procedure as transformation. Antibiotic(s) were added to the plate to confirm the co-expression result was correct. A single colony was chosen to grow in the 100 mL LB media with antibiotic(s) for overnight at 37°C 180 rpm to continue for glycerol stock for expression and purification.

Table 4.5. Co-expression for crystallisation

Constructs	Co-expression with
USP4FL -H881N	USP4FL-H881N with diubiquitin USP4FL-H881N with diubiquitin-L73X
USP4FL -C311S-Y928X	USP4FL -C311S-Y928X with Ub-GGG USP4FL -C311S-Y928X with diubiquitin
USP4C1C2-H881N	USP4C1C2-H881N with diubiquitin USP4C1C2-H881N with Ub-GGG USP4C1C2-H881N with diubiquitin-L73X
USP4C1C2-C311S	USP4C1C2-C311S with diubiquitin USP4C1C2-C311S with UbGGG
USP4C1C2-C311S-Y928X	USP4C1C2-C311S-Y928X with UbGGG USP4C1C2-C311S-Y928X with diubiquitin
USP4htt-H106N	USP4htt-H106N with diubiquitin USP4htt-H106N with Ub-GGG USP4htt-H106N with diubiquitin-L73X
USP4htt-C311S	USP4htt-C311S with UbGGG USP4htt-C311S with diubiquitin
USP4httΔUbl2-H106N	USP4httΔUbl2-H106N with diubiquitin
USP4httΔUbl2-C311S	USP4httΔUbl2-C311S with Ub-GGG
USP4ΔDU-H881N	USP4ΔDU-H881N with Ubiquitin
USP4DU	USP4DU with SART3*

*=co-purify

4.2. The Protein Expression and Purification

The protein expression starts from overnight cultures were grown from glycerol stocks with antibiotic(s) in 100 mL LB at 37℃ 180 rpm shaking incubator. The next day, the cultures then transfer to 5 X of 800 – 1000 LB media with appropriate antibiotic(s), then continue incubation at 37℃ 180 rpm until OD reached between 0.6 – 0.8, and induced with IPTG 0.5 mM. After induction, the incubation continues with different temperature depending on the construct for certain time (overnight to 60 hours). Every construct need specific condition for expression (Table 4.6). The cell then was harvested and lysis. Cells were harvested by centrifugation for 30 minutes at 4600 rpm. The supernatant was removed and the pellets were store at -20°C in 50 mL tubes.

Table 4.6. Expressions conditions

Construct	Cell type	Media	Induction temperature	IPTG concentration	Expression time (hour)
USP4FL	BL21 CP	2YT	10	0.5	60
USP4htt	BL21 CP	LB	18	0.5	20
USP4C1C2	BL21 CP	LB	37	0.5	4
USP4ΔDU	Arctic express	LB	15	0.5	36
USP4httΔUbl2	BL21 CP	LB	18	0.5	20
USP4DU	BL21 CP	LB	37	0.5	4
Ubiquitin	BL21 CP	LB	37	0.5	4

For active site mutant constructs, the expression was the same as wild-type construct.

4.2.1. Lysate Preparation

The procedure for lysate preparation was: cells thawed and resuspended in buffer A (about in 10 mL per gram). Cells were sonicated on ice for 20 seconds off and 20 seconds on for a total 2 minutes. After sonication, centrifugation at

14500 rpm at 4°C for 45 minutes was performed to separate supernatant and pellet. The supernatants filtered with 0.45 µM filter and ready for purification.

4.2.2. Affinity chromatography purification

All the USP4 modified construct has a His tag and purified using an affinity chromatography method (Table 4.7). A Ni column (GE Healthcare) 5 mL and AKTA Prime or AKTA Purifier GE Healthcare used. The column was prepared by washed with water 2x column volume (CV), EDTA 5x CV, water 2x CV, then charged with $NiSO_4$ 5x CV, water 2xCV, and Buffer A 5x CV. Protein sample then was loading to the Ni column. Buffer B contains a high concentration of Imidazole to elute the protein interest. The fraction was collected at 2.5 mL and continue by SDS PAGE analysis. The protein then concentrated using a centrifugal ultrafiltration unit with suitable molecular weight membrane.

Table 4.7. The types of buffer for protein purification of all construct

Construct	Affinity chromatography	Size exclusion chromatography	Co-express with
USP4FL	Buffer A: 300 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 20 mM Imidazole Buffer B: 300 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 500 mM Imidazole	300 mM NaCl 50 mM Tris-Cl 1% Glycerol	
USP4htt	Buffer A: 150 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 20 mM Imidazole Buffer B: 150 mM NaCl, 50mM Tris-Cl, 5% Glycerol, 500 mM Imidazole	150 mM NaCl 50 mM Tris-Cl 1% Glycerol	
USPC1C2	Buffer A: 300 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 20 mM Imidazole	300 mM NaCl 50 mM Tris-Cl 1% Glycerol	

Table 4.7. (Continued) The types of buffer for protein purification of all construct

	Buffer B: 300 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 500 mM Imidazole		
USP4ΔDU	Buffer A: 300 mM KCl, 50mM Tris-Cl, 10% Glycerol, 20 mM Imidazole Buffer B: 200 mM KCl, 50mM Tris-Cl, 10% Glycerol, 500 mM Imidazole	150 mM NaCl 50 mM Tris-Cl 1 % Glycerol	
USP4httΔUbl2	Buffer A: 150 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 20 mM Imidazole Buffer B: 150 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 500 mM Imidazole	150 mM NaCl 50 mM Tris-Cl 1 % Glycerol	
USP4DU	Buffer A:100 mM NaCl, 50mM Tris-Cl,10% Glycerol, 20 mM Imidazole Buffer B:100 mM NaCl, 50mM Tris-Cl,10% Glycerol, 500 mM Imidazole	300/100 mM NaCl 50 mM Tris-Cl 1 % Glycerol	SART3 (co-purify)
USP4FL-H881N USP4FL-C311S USP4FL-C311S-Y928X	Buffer A:150 mM NaCl, 50mM Tris-Cl,5% Glycerol, 20 mM Imidazole Buffer B:150 mM NaCl, 50mM Tris-Cl, 5% Glycerol, 500 mM Imidazole	300 (100) mM NaCl 50 mM Tris-Cl 1 % Glycerol	Diubiquitin Ub-GGG diUbiquitin-L73X diubiquitin Ub-GGG
USP4htt-H160N USP4htt-C311S	Buffer A: 150 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 20 mM Imidazole Buffer B: 150 mM NaCl, 50mM Tris-Cl, 5% Glycerol, 500 mM Imidazole	300/150/100 mM NaCl 50 mM Tris-Cl 1 % Glycerol	Ubiquitin Diubiquitin Ub-GGG Diubiquitin-L73X Ub-GGG Diubiquitin
USP4httΔUbl2-H106N	Buffer A: 150 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 20 mM Imidazole Buffer B: 150 mM NaCl, 50mM Tris-Cl, 5% Glycerol, 500 mM Imidazole	150 mM NaCl 50 mM Tris-Cl 1 % Glycerol	Diubiquitin
USP4C1C2-H881N USP4C1C2-C311S USP4C1C2-C311S-Y928X	Buffer A: 150 mM NaCl, 50mM Tris-Cl, 10% Glycerol, 20 mM Imidazole Buffer B: 150 mM NaCl, 50mM Tris-Cl, 5% Glycerol, 500 mM Imidazole	300/150/100mM NaCl 50mM Tris-Cl 1% Glycerol	Diubiquitin Ub-GGG Diubiquitin-L73X Diubiquitin Ub-GGG
USP4ΔDU-H881N	Buffer A:150 mM NaCl, 50mM Tris-Cl,10% Glycerol, 20 mM Imidazole Buffer B:150 mM NaCl, 50mM Tris-Cl,5% Glycerol, 500 mM Imidazole	100mM NaCl 50mM Tris-Cl 1% Glycerol	Ubiquitin

4.2.3. Anion exchange

Protein samples for anion exchange were prepared concentrated until 5 mL and then adding with 45 mL buffer A (20 mM Tris). At least the pH of the buffer two units above the predicted pI by protparam of the protein sample. This technique used Resource Q 6 mL (GE Healthcare). The column was pre-equilibrated with buffer A, and the sample protein was loaded on AKTA purifier using a 50 mL super loop. 30 mL buffer A was used to wash the unbound protein then the protein was eluted using buffer B (20 mM Tris, 1 M NaCl) as a gradient for 20 CV. The protein fractions were collected and analysed using SDS-PAGE. Only the construct of USP4ΔDU used this method for purification.

4.2.4. Size exclusion chromatography

The final purification used size exclusion chromatography. This purification depending on the size of the protein product and use different buffer (Table 4.7). For the small protein (3 – 70 kDa) was used HiLoad 16/60 Superdex 75, and for big protein (10 kDa – 600 kDa) was used HiLoad 16/60 Superdex 200 (GE Healthcare) (Table 4.8). Before application, the column equilibrated in the appropriate buffer for 1 column volume. Then the sample was loaded to the column up to 5 mL volume. The purification was running by the installed method. The fraction of protein interest collected in the tube and after checked on the SDS PAGE gel concentrated and tested the quantity using Nanodrop

(A280) to determine the protein concentration (mg/mL) and confirm by the extinction coefficient factor of each protein.

Table 4.8. Gel filtration columns

Column	Proteins used
HiLoad 16/60 Superdex 75	Ubiquitin, Ubiquitin-diubiquitin, diubiquitin-L73X
HiLoad 16/60 Superdex 200	USP4FL, USP4C1C2, USP4htt, USP4httΔUbl2, USP4ΔDU, USP4DU, Complex USP4FL with Ubiquitin variant, Complex USP4htt with Ubiquitin variant, Complex USP4C1C2 with Ubiquitin variant, Complex USP4ΔDU with Ubiquitin variant, Complex USP4httΔUbl2 with Ubiquitin variant Complex USP4DU with SART3

4.3. Ubiquitin and ubiquitin variant expression and purification

Ubiquitin and linear diubiquitin got from Dreveny group stock, and Ubiquitin-GGG was made by Stephanie Ward (2015). Ubiquitin and ubiquitin variant expression and purification was prepared for ITC analysis (Sub chapter 4.9). The expression starts from overnight 100 mL culture of Ubiquitin, 10 mL of culture was transferred to 1 L of LB broth with antibiotics (Kanamycin and Chloramphenicol) and incubated at 37˚C with 180 rpm shaking until OD = 0.7. Then induction by addition of 500 µL 1M IPTG for 4 hours at 37˚C with 180 rpm shaking. Cells were pelleted, resuspended and lysed by sonication in low salt buffer Buffer A (20 mM ammonium acetate pH 5.1) and centrifuged. After centrifugation, acetic acid added until the pH drops between 4.5 and 5.0. The solution becomes milky white. The supernatant was centrifuged again for 10

minutes at 14500 rpm 4°C and pH of the supernatant adjusted until 5.1 with sodium hydroxide.

Ubiquitin and ubiquitin variant used the cation exchange chromatography for purification. The sample then was loaded onto the cation exchange column (SP resin) and eluted with a linear gradient of increasing concentration of ammonium acetate (Buffer B: 0.5 M ammonium acetate pH 5.1). The purification continued by size exclusion chromatography using superdex 75 (GE Healthcare) with the buffer 150 mM NaCl, 20 mM Tris pH 5.1 and 1 % Glycerol. Protein was concentrated and stored at -80°C.

4.4. SDS-PAGE

The SDS-PAGE analysis was used to confirm the molecular weight of protein samples and the purities of single or complex protein. Gels with 0.75 mm thickness prepared. SDS-PAGE consists of two components, the resolving, and stacking (Table 4.9). The resolving gel at 15% mostly used. Only for ubiquitin variant, the concentration of 18% used.

Table 4.9. Composition of SDS-PAGE gels

Reagent	Resolving gel volume (mL)	Stacking gel volume (mL)
H_2O	3.7 – 6.3*	5.4
30% Acrylamide/Bis-acrylamide	8 – 10.67*	2
1.5 M Tris pH 8.8	4	-
0.5 M Tris pH 6.8	-	2.5
10 % SDS	0.16	0.1
10 % APS	0.16	0.1
Temed	0.016	0.01

*the range for 15 – 20% of gels

Samples prepared by mixing 3X sample buffer (reducing) with the protein samples. The sample and the protein marker (NEB) were boiled for 5 minutes

at 95°C before loading into the gels. The gels were run at 200 V for about 45 minutes until the band reached the bottom of the gel. Gels were stained with staining solution (50% isopropanol, 10% glacial acetic acid, 0.01 % Coomassie Brilliant Blue) then followed by a destaining solution with 10% glacial acetic and 50% isopropanol.

4.5. Protein concentration calculation

The concentration of purified protein determined by Nanodrop 1000 (Thermo Scientific). 2 µL sample of protein was used with blank from gel filtration buffer as a correction. The value from Nanodrop then divides by extinction coefficients factor of the protein sample. Extinction coefficient factor is taken from Protparam analysis. The final value of protein in mg/mL or µg/µL.

4.6. Enzymatic Activity Assay using Ubiquitin-AMC (Ub-AMC)

Enzymatic activity assay using Ub-AMC (Pickering and Davies, 2012) was performed to calculate the specific activity of the active protein (USP4FL, USP4htt, USP4C1C2, and USP4ΔDU) (Figure 4.2).

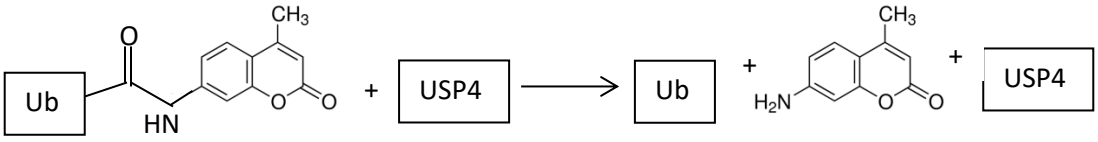


Figure 4.2. Reaction of Ub-AMC with USP4 enzymes

First, the standard curve of AMC was determined. Serial dilution of AMC was prepared from 0.1; 0.2; 0.3; 0.4; 0.5; 0.6; 0.7; 0.8; 0.9; and 1 µM in the total volume of 30 µL with AMC buffer (150 mM NaCl, 50 mM Tris-Cl pH 7.5, 1 mM DTT). Control was AMC buffer only. All measurements were in

duplicate/triplicate and performed in 384-well white plates (Nunc). The plate was read with EnVision 2104 multilabel plate reader using emission 426/428 nm and excitation 355 nm. An AMC standard curve produced by plotting concentration vs the fluorescence of AMC, and first equation (y=bx+a) linear regression made.

The enzymatic activity assay for USP4FL performed at a concentration of 50 nM with 0.1 μM, 0.25 μM, 0.5 μM, and 0.75 μM ubiquitin-AMC (7-amido-4-methyl coumarin) as the fluorogenic substrate (Viva Bioscience). For USP4htt performed at a concentration of 500 nM with 0.75 μM ubiquitin-AMC. USP4C1C2 performed at 20 nM with 0.75 μM ubiquitin-AMC, and USP4ΔDU was at 50 nM with 0.75 μM ubiquitin-AMC.

Reactions were carried out in buffer 150 mM NaCl, 50 mM Tris-Cl pH 7.5, 1 mM DTT, in 384-well white plates (Nunc) with 30 μL volume. The plate was read with EnVision 2104 multilabel plate reader using emission wavelength 426/428 nm and excitation wavelength 355 nm. The measurement was taken every minute for 60 minutes, and the curve time vs fluorescent unit was obtained. Then, using the AMC standard curve equation, the fluorescence unit data translates into the AMC released. The fluorescent unit data only taken from the initial rate, a linear line which described the substrate was cleaved by the enzyme, usually around first 10 - 20 minutes. Based on this calculation, the second curve of the time vs AMC released (in μM) and the second equation

(y=bx+a) of USP4FL, USP4htt, USP4C1C2, and USP4ΔDU were generated. The enzymatic activity assay was determined using the equation of the time vs AMC released (in μM) result by divided the b (slope) with the amount of protein (in mg). Specific activity calculated as $\mu M min^{-1} mg^{-1}$.

4.7. Crystallisation trials

The crystallisation trials performed to the active site mutant of various USP4 constructs, both H881N and C311S, combine with ubiquitin variant, namely ubiquitin, diubiquitin, Ubiquitin-GGG, and diubiquitin-L73X. Crystallisation trials also performed using inhibitor agents such as Vialinin and mitoxantrone for USP4C1C2 construct, and SART3 in complex with USP4DU.

The sitting drop vapor diffusion was crystallisation method used, and the Mosquito robot was used to set up the crystallisation trials. The concentration of complex protein was around 2.5 – 5 mg/mL in the 400 nL drop with a various screen such as Morpheus, PACT premier, PGA, Structure, MIDAS, Classic suite, Clear Strat, MemGold, Proplex, and JCSG plus (Molecular Dimensions). The tray was incubated on 10℃ and observed every day until several weeks or months later.

4.8. Native Mass Spectrometry

Native mass spectrometry analysis was performed to characterise the complex of protein and protein-protein interaction. This technique is done by Neil Oldham and Gemma Cook (School of Chemistry, University of Nottingham)

using SYNAPT^{TM} HDMS (Waters, Altrincham, UK) as a time-of-flight (TOF) mass analyser. The sample for this assay was prepared as Table 4.10, consist of a single and complex protein. The protein sample was desalted using 150 mM ammonium acetate pH 6.8 and injected into the instrument. The instrument was operated using MassLynx^{TM} 4.1 software.

Table 4.10. Samples for ESI-mass spectrometry

Constructs	Samples
USP4FL	Complex USP4FL with Ubiquitin
	Complex USP4FL-H881N with diubiquitin
	Complex USP4FL-H881N with diubiquitin L73X
USP4C1C2	Complex USP4C1C2-H881N with Ub-GGG
	Complex USP4C1C2-H881N with diub-L73X
USP4ΔDU	Complex USP4ΔDU with Ubiquitin
USP4DU	Complex USP4DU-SART3

4.9. Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC) is an experiment to investigate the binding interaction, especially between the protein-protein interactions. The sample for ITC were active site mutant USP4FL-C311S, USP4C1C2-C311S and USP4ΔDU-C311S with the ligands: Ubiquitin, diubiquitin, and Ub-GGG.

The experiment performed at MicroCal PEAC-ITC machine (Malvern) (Table 4.11). The sample for protein diluted to 20 µM to 50 µM. The sample for ligand was prepared 200 µM to 500 µM. The ligands (Ubiquitin, Ub-GGG, and diubiquitin) titrated into sample cell in 19 x 2 µL aliquot and the spacing between injections was 180 seconds with the stirring speed was 750 rpm. The temperature set to 25°C. The same protocol performed in a buffer control by

injecting the ligand into a buffer (150 mM NaCl, 50 mM Tris pH 7.5 and 1 % Glycerol). Using Malvern software, the binding parameter, such as constanta dissociation and stoichiometry of interaction will be calculated.

Table 4.11. Details of ITC experiment between USP4 and Ubiquitin variant

Macromolecule	**Molecular weight (kDa)**	**Ligand**	**Molecular weight (kDa)**
USP4FL-C311S	111	Ubiquitin	8.5
		Ubiquitin-GGG	8.6
		Diubiquitin	17
USP4C1C2-C311S	46	Ubiquitin	8.5
		Ubiquitin-GGG	8.6
		Diubiquitin	17
USP4ΔDU-C311S	79.5	Ubiquitin	8.5
		Ubiquitin-GGG	8.6
		Diubiquitin	17

Chapter 5

Result and Discussion: Cloning, Protein Expression and Purification of Active Proteins (USP4FL, USP4htt, USP4C1C2, USP4ΔDU, USP4httΔUbl2, and USP4DU)

This chapter shows the results and discussion about the process of the design, expression, and purification of USP4 constructs. Each USP4 construct was characterised in terms of its expression and purification and evaluated for its enzymatic activity. These experiments have the purpose of understanding the function and role of each domain in USP4 and comparing the stability and activity between the domains.

5.1. Cloning, Protein Expression, and Purification of USP4FL

The first construct was USP4FL which has 111.1219 kDa and 3.128 kb. This construct had 963 amino acids and was cloned into the pColdI DNA vector with a Histidine tag, and the total amino acid was 984. The PCR result showed the length of the PCR product was about 3 kb. The cracking test was done to choose the correct construct and continue to digestion test with KpnI and Hind III. The pColdI DNA, as a vector, has a 4.407 kb of DNA. The digestion result indicated the construct of USP4FL inserted in the vector of pColdI. The sequencing result confirmed that the construct was corrected and created.

For expression of USP4FL, the construct of USP4FL continue to the transformation in *E. coli* BL 21 codon plus. On the transformation process, Ubiquitin was added. Co-expression with ubiquitin increased the yield of USP4FL (Catanzariti, et al., 2004).

The first purification of USP4FL was Ni^{2+} affinity chromatography. USP4FL has a poly-histidine tag and will bind specifically to Nickel, as specific ligands in the

matrix (Tomlinson et al., 2007). The Ni^{2+} column purification of USP4FL showed a broad single peak, indicating plenty of protein in solution was obtained (Figure 5.1). The protein USP4FL started eluting from 10% buffer B. The SDS-PAGE gel showed some protein degradation of the protein target. USP4FL has 111.1 kDa, and the result showed that there was an intense band below 212 kDa and above of 66.4 kDa, which refers to the protein target.

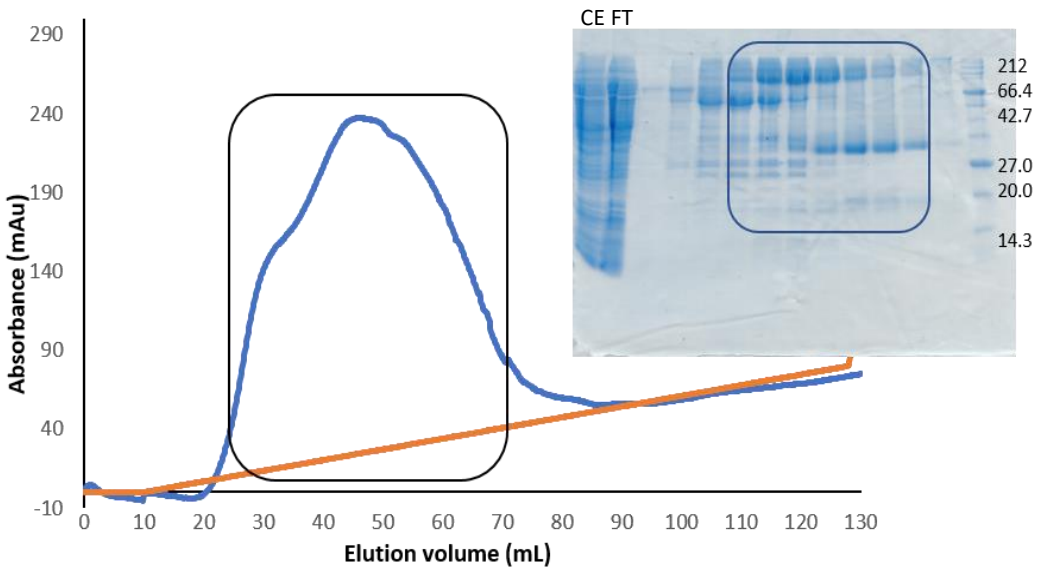


Figure 5.1. Affinity chromatography and SDS-PAGE gel of USP4FL.
The chromatogram showed that protein elutes from 10% Buffer B. SDS-PAGE gel from Ni column showed a band below 212 kDa marker (refer to USP4FL with 111 kDa) and some degradation product. (CE= crude extract; FT=flow trough).

The gel filtration purification of USP4FL showed that the protein elutes on the first peak with the elution volume of 72.8 mL. This volume equals to that of a standard curve at the prediction of molecular weight of about 97.8 kDa. The value was lower than the actual molecular weight, which indicated that the protein folded compactly. The SDS-PAGE gel also confirmed that the product contains some

impurities (Figure 5.2). The second peak and associated molecular weight indicated that the DUSP Ubl domain, at about 29 kDa, was the degradation product. The total pure protein after concentrating was only 0.7 mg/L culture.

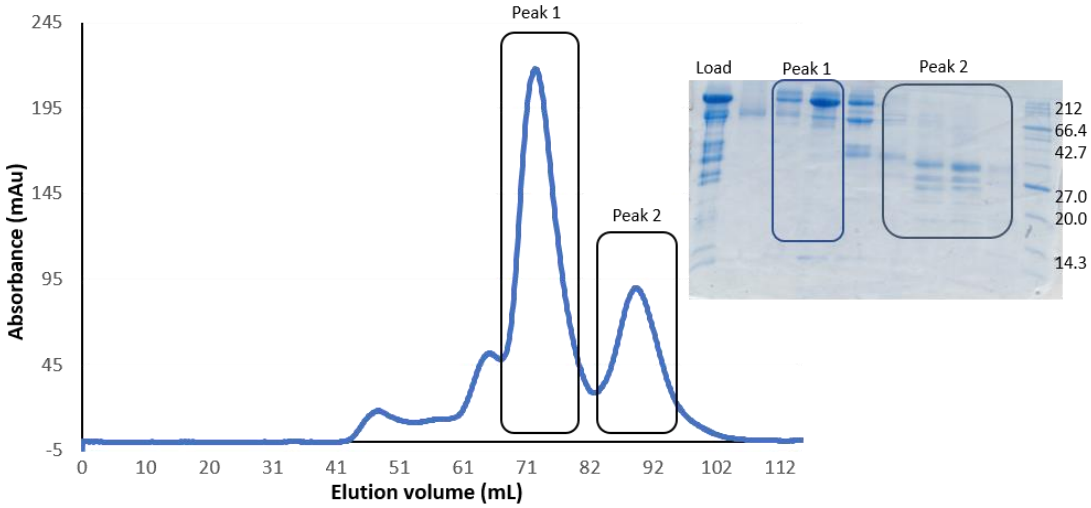


Figure 5.2. Chromatogram of USP4FL Gel Filtration purification
The chromatogram of Superdex 200 HiLoad 16/60 showed several peaks. The protein target has elution volume at 72.8 mL, on the first peak, which equals to 97.8 kDa based on the standard curve. SDS-PAGE gels showed a band at around 111 kDa which is a USP4FL with a little degradation product appears.

Then enzymatic activity assays were performed to confirm the activity of USP4FL. The USP4FL protein indicates an active protein by the capability to cleave the substrate Ub-AMC (7-amido-4-methyl coumarin) to free AMC and Ubiquitin. AMC was a fluorescent substance. The fluorescence of AMC will increase over time by the activity of USP4FL. The enzymatic activity result showed USP4FL was active (Figure 5.3).

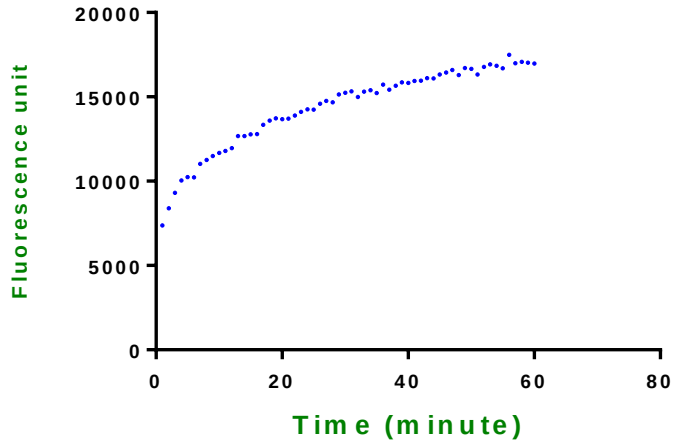


Figure 5.3. Enzymatic activity assay graph of USP4FL with Ub-AMC.
USP4FL concentration was 50 nM, and Ub-AMC was 1 μM. The plate was read every minute for 1 hour in duplicates.

An AMC standard curve was needed to calculate the specific activity of the protein (Appendix). AMC standard curve was obtained by preparing a serial concentration of AMC and check its fluorescence. The standard curve was linear between X and y concentration of AMC. Based on the AMC standard curve, the equation then calculated, it was $y=5253x+20.87$ ($R^2=0.96$) (Equation 1). The specific enzyme activity was calculated from the initial rate Figure 5.3. The initial rate was taken from zero unit until the fluorescence did not linearly increase, for about first 15 minutes. The increase of linearity on the graph means the enzyme successfully cleave the substrate (Ub-AMC) to be free Ub and AMC. By the time, the number of AMC increase along with the linearity of the graph. When the graph start stationary, it indicates that the substrate has been all converted by the enzyme. Equation 1 and the linear sector of the curve were used to determine the graph of time vs AMC released (Figure 5.4).

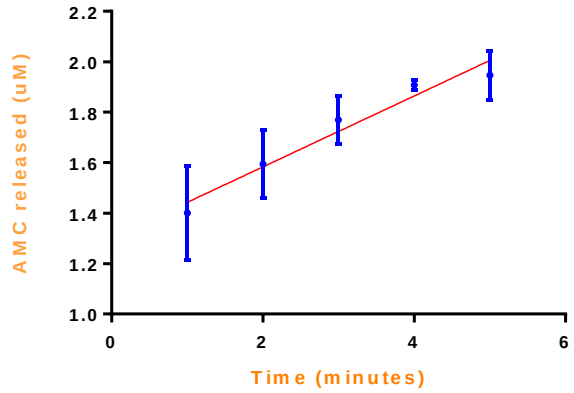


Figure 5.4 Linear range of progress curve to obtain initial rate of USP4FL
USP4FL was 50 nM, and Ub-AMC was 1 μM. Data on duplicate

Based on the graph of time vs AMC released from USP4FL (Figure 5.4), the next equation was calculated. It was Y=0.14*X + 1.30 (R^2 = 0.96) (equation 2). The slope (b) of equation 2 was then divided by the amount of enzyme to obtain the specific enzyme activity. The specific activity of USP4FL was 843.5 μM/min/mg enzyme.

The kinetic parameters of USP4FL is then determined using the serial concentrations of substrate Ub-AMC, from 300 nM to 1.2 μM, the chosen concentration of USP4FL was 50 nM. The experiment was run for 60 minutes (Figure 5.5). From the raw data, the initial rate was determined, about first 20 minutes of the reaction and then plotted into AMC standard curve as a product (Figure 5.6) to obtain kinetic parameter including K_m, V_{max}, and k_{cat}.

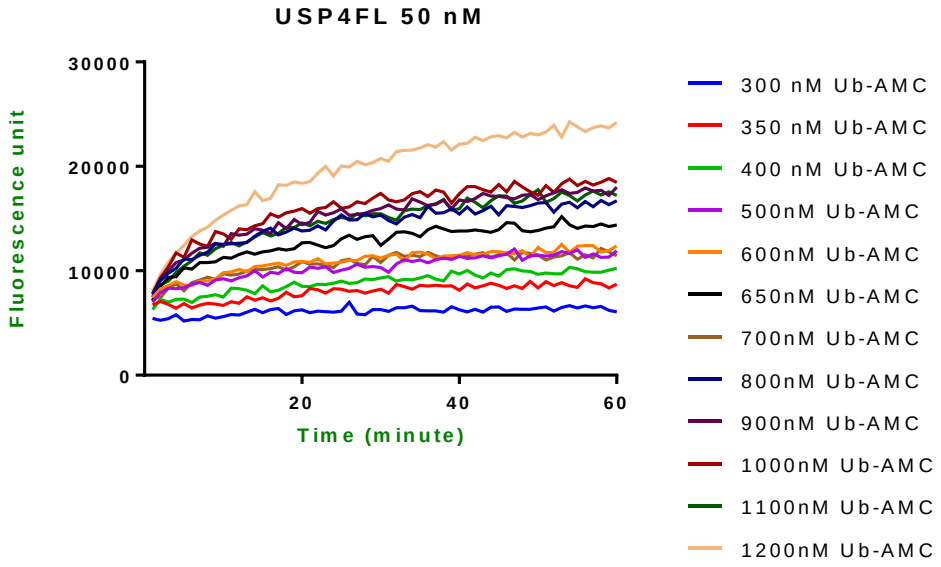


Figure 5.5. Sample of raw data from the kinetic assay with USP4FL at 50 nM and the Ub-AMC, from 300 nM until 1200 nM

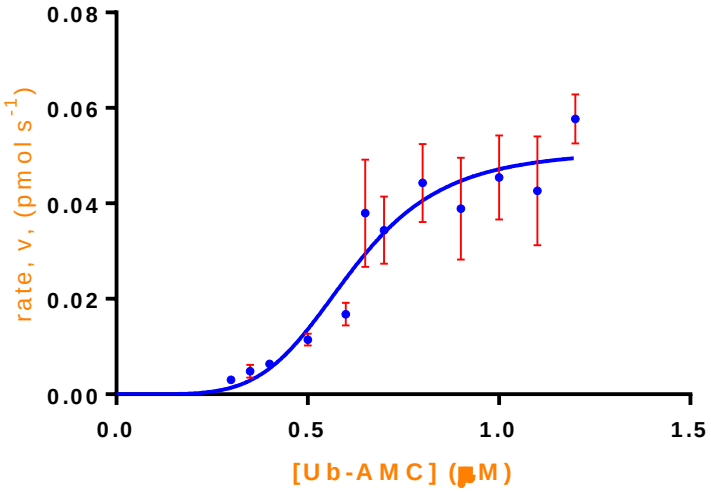


Figure 5.6. Graph of the rate of reaction against substrate concentration for USP4FL. The data were measured in duplicate.

Compared to the data obatained by Clerici et al., (2014), the K_m and k_{cat}/K_m value from this experiment (Table 5.1) were slightly higher, such as 0.68 ± 0.081 μM and 55.76 ± 6.83 unit rather than 0.15 μM and 5.1 ± 0.23 unit. However, the value of k_{cat} was 0.037 ± 0.000097 s^{-1} is less than 0.30 ± 007s^{-1}.

Table 5.1. The kinetic parameters for USP4FL

Parameter	k_{cat} [s^{-1}]	K_m [μM]	$k_{cat}/K_m .10^3$[M^{-1} s^{-1}]	V_{max} [pmol/s]
Mean ± SD	0.037 ± 9.7 10^{-5}	0.68 ± 0.081	55.76 ± 6.83	0.056 ± 0.0001

5.2. Cloning, Protein Expression, and Purification of USP4htt

The USP4htt structure consisted of 479 amino acids or about 1,236 kb of nucleotides. The cloning process started with PCR product of USP4htt and only produced at 65˚C annealing temperature. The ligation with pET26b, as a vector, was done, but the DNA concentration after digestion test for USP4htt was low (5.9 ng/μL). The cloning step then continued by transformation. The miniprep DNA isolation after transformation process of the USP4htt constructs indicated high-value concentration (123.4 ng/μL). The mini-digest of the USP4htt confirmed that the PCR product was inserted into the pET26b vector. Sequencing showed that this new construct of USP4htt was correctly engineered.

USP4htt needed optimisation for protein expression and purification. The highest yield after gel filtration purification (S200 HiLoad 16/60) was 0.3 mg/L of culture when it was co-purified with ubiquitin. The initial Ni column and gel filtration purification of the USP4htt showed minimum impurities (Figure 5.7). The protein product of USP4htt was 50.5 kDa. The result from purification showed the band above a marker 42.7 kDa, which refers to USP4htt. The gel

filtration result showed decreasing impurities with the clearer band at above 42.7 kDa.

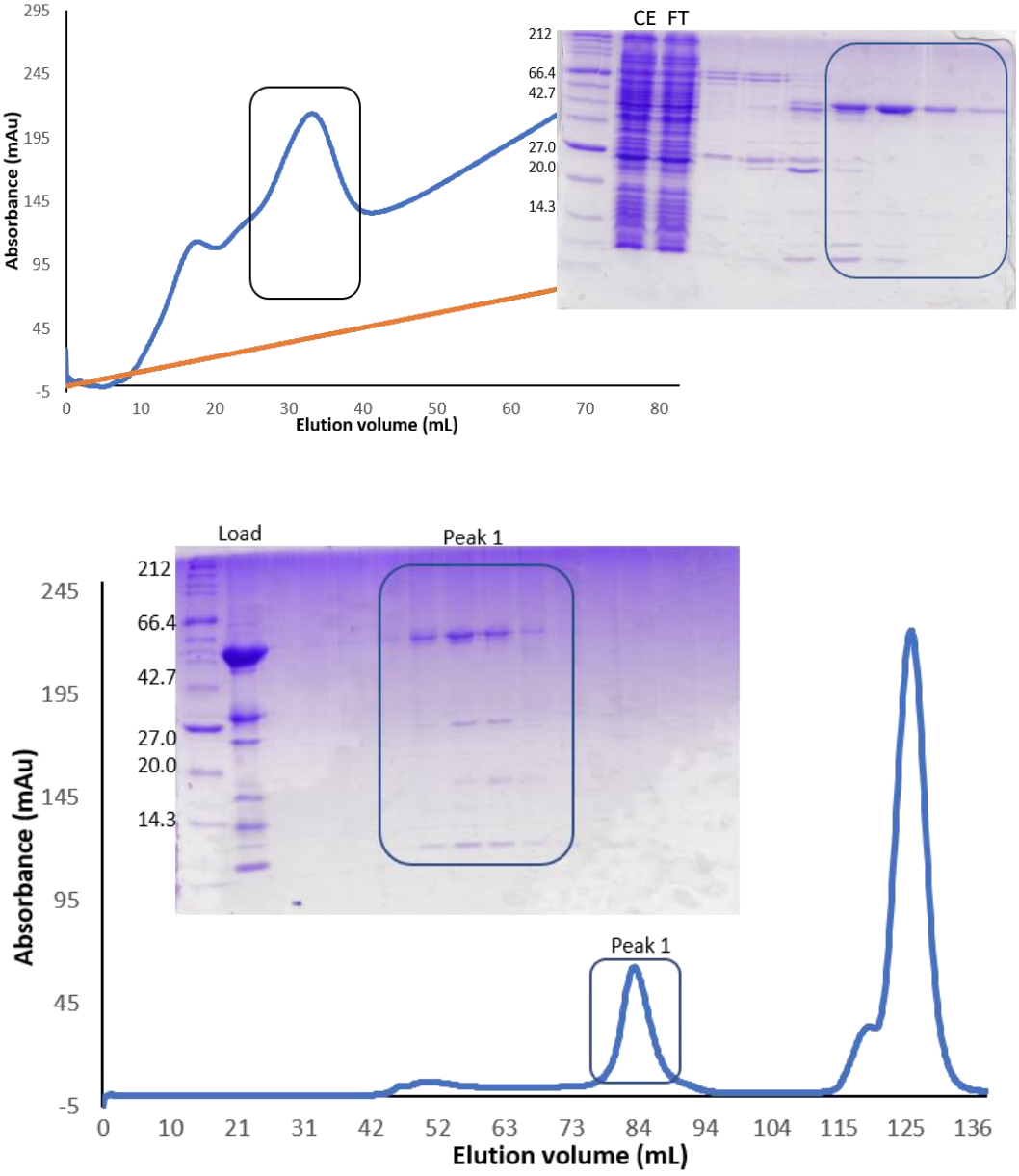


Figure 5.7. Affinity chromatography and gel filtration purification of USP4htt
Chromatogram from Ni column purification (top) showed after 20% of buffer B the protein USP4htt was eluted. SDS-PAGE gel indicated USP4htt has about 50.5 kDa was expressed. Gel Filtration purification in Superdex 200 HiLoad 16/60 (bottom) showed the elution volume was 82 mL correlate with 44.8 kDa at the standard curve. The SDS-PAGE gel of gel filtration of USP4htt concludes that the protein was quite pure in low concentration. (CE=crude extract; FT=flow trough).

Gel filtration Superdex200 HiLoad 16/60 of USP4htt has elution volume at 82 mL. This value was equal with 44.8 kDa of molecular weight at the standard curve. Slightly less than actual 50.5 kDa. The total protein yield of USP4htt was 0.3 mg from 5 L cultures or about 0.05 mg/L cultures.

Elution volume of USP4htt higher than USP4FL. This data in accordance with the principle of size exclusion chromatography, the molecule with bigger size will elute first followed by the small one. In this study, USP4FL with 111 kDa elute at 72.8 mL before the USP4htt at 82 mL whose 50.5 kDa of molecular weight.

The enzymatic activity assays of USP4htt also was performed to confirm the activity. An AMC standard curve to calculate the specific activity of the USP4htt protein was same as the standard curve for USP4FL. The specific enzyme activity of USP4htt was calculated from the initial rate, from 0 until 10 minutes. In this area, the curve showed linear, the enzyme in progress converting the substrates and showed by the increasing of the fluorescence (Figure 5.8). The following calculation using equation 1 and linear sector of the curve to determine the graph of time vs AMC released (Figure 5.9).

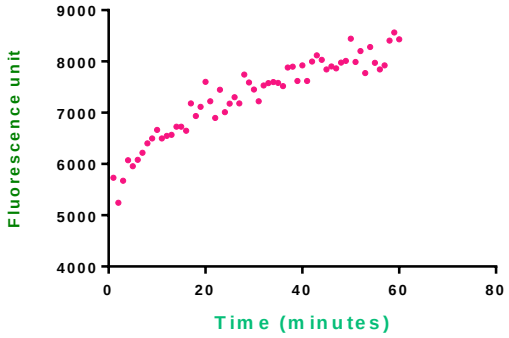


Figure 5.8 Enzymatic activity assay graph of USP4htt with Ub-AMC
USPhtt was 500 nM and Ub-AMC 1 µM. The plate was read every minute for 1 hour in duplicate.

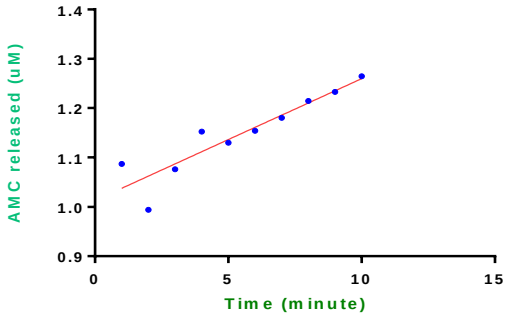


Figure 5.9. Linear range of progress curve to obtain the initial rate of USP4htt.
USP4htt was 500 nM, and Ub-AMC was 0.75 µM

Based on the graph of time vs AMC released from USP4htt (Figure 5.9), the equation of each it was known. USP4htt was Y=0.025*X + 1.01 (R^2= 0.84). The slope (b) of this equation divides by the amount of enzyme was specific enzyme activity. The specific activity of USP4htt is 32.5 µM/min/mg enzymes. The specific activity of USP4htt was lower more than 25 times compared to USP4FL.

5.3. Cloning, Protein Expression, and Purification of USP4httΔUbl2

The construct of USP4httΔUbl2 was created by putting the catalytic 2 and catalytic 1 subdomains in the reverse direction and connecting them with the

flexible linker. The construct is the circular permutated USP4htt but without the Ubl2 domain. The construct has the same N-terminal domain as the USP4htt construct, but the catalytic 1 sub-domain is now at the C-terminus.

This construct was cloned into the same vector as USP4htt, namely pET26b, and digested using the same enzymes (NdeI and XhoI). After ligation, transformation, and miniprep, the sequencing results showed USP4httΔUbl2 construct was created correctly.

The USP4httΔUbl2 based on Protparam calculation has molecular weight 40.7 kDa. The purification was using Ni column and gel filtration method. The Ni column and gel filtration purification showed a high yield of product with a little degradation (Figure 5.10). On the Ni column purification, the protein target started to elute at 18.5% Buffer B and only has one single peak. The SDS PAGE gel after Ni column showed a thick band with low impurities.

The gel filtration chromatogram showed the protein target eluted at 87.2 mL. Compared with a standard curve, the elution volume of USP4httΔUbl2 corresponded to 28.9 kDa. The value was lower than the actual size of the protein 40.7 kDa. This data showed that USP4httΔUbl2 has a compact structure. However, the SDS PAGE gel showed the prediction of molecular weight around 40.7 kDa. SDS PAGE gel indicated that the protein was quite pure with a little impurity on the sample. The total protein after gel filtration

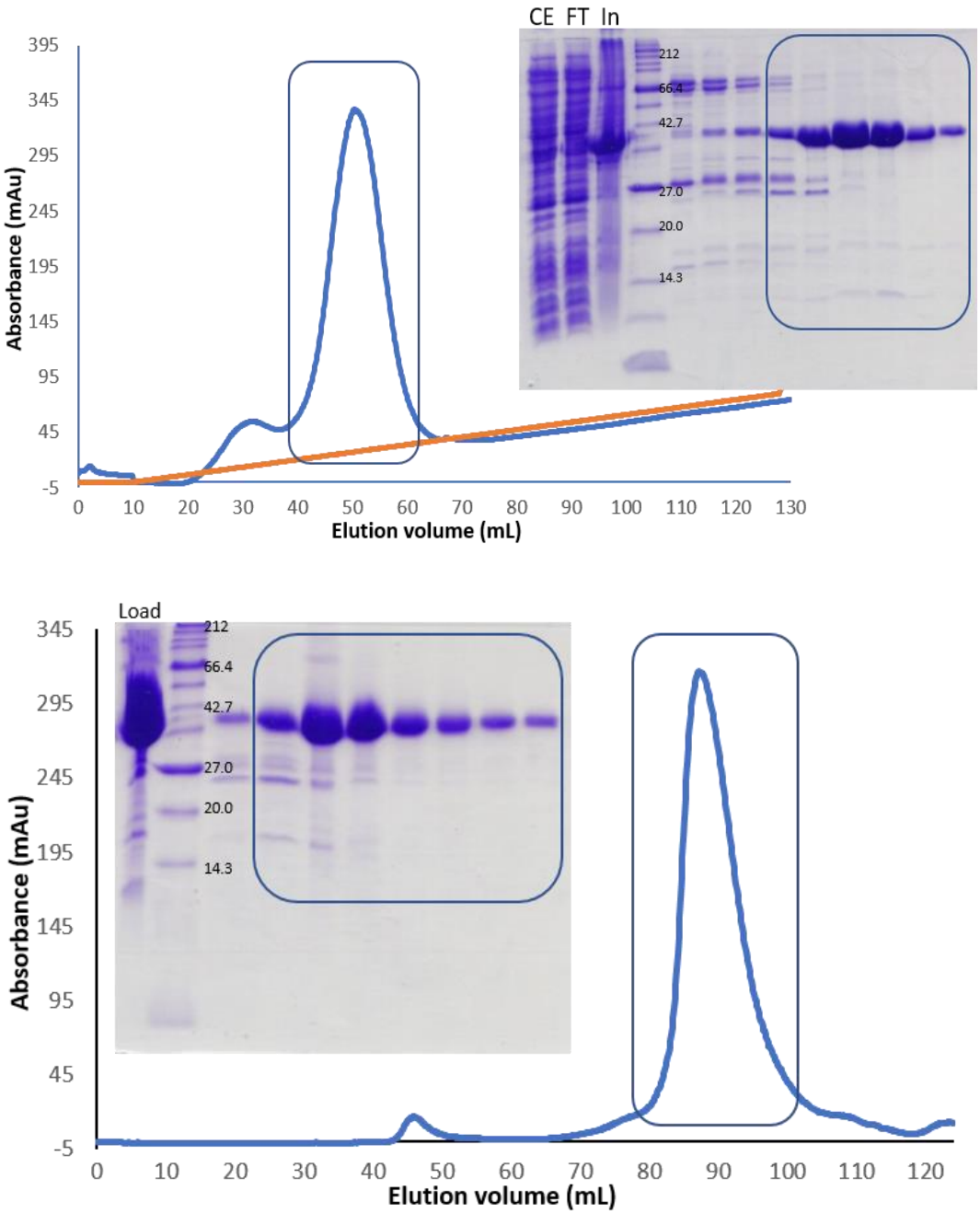


Figure 5.10. Affinity chromatography and gel filtration purification of USP4httΔUbl2
Chromatogram from Ni column purification (top) shows a single peak corresponding to the protein product. SDS-PAGE showed a high expression of the product USP4httΔUbl2 at 40.7 kDa. CE=crude extract; FT=flowtrough; In=Insoluble; M=Marker;
A chromatogram from gel filtration SuperDex 200 HiLoad 16/60 (bottom) with a single high peak at 87.2 mL as elution volume correspondent with 28.9 kDa based on the standard curve. SDS-PAGE gel shows the protein is about 40.7 kDa with less degradation product.

purification was 14.5 mg from 4 L cultures. The yield of this purification was 3.63 mg/L culture.

5.4. Cloning, Protein Expression, and Purification of USP4C1C2

The construct of USP4C1C2 consists of the catalytic domain only, namely catalytic 1 and catalytic 2 subdomains only. The construct of USP4C1C2 was engineered by overlap extension PCR adapted from Warrens et al., (1997). The construct was designed from four different primers to produce the first and the second PCR product. The first product was a catalytic 1 subdomain and the second was a catalytic 2 as subdomains in the structure of USP4FL. The catalytic 1 and catalytic 2 have 549 bp and 543 bp, respectively. The third PCR was combined the catalytic 1 and catalytic 2 to be 1 kb of the product. The product will connect the catalytic 1 and catalytic 2 with the linker ASTSK. This construct ligated on the pProExHtb vector with Histidine tag. Sequencing analysis confirmed that USP4C1C2 had engineered.

The Ni^{2+} affinity column purification of USP4C1C2 showed two peaks on the chromatogram. The protein target was in the second peak after elution with Buffer B at 23.5 % (Figure 5.11). SDS-PAGE gel showed the very clear single band at above a marker 42.7 kDa. The actual size of USP4C1C2 was 46.4 kDa.

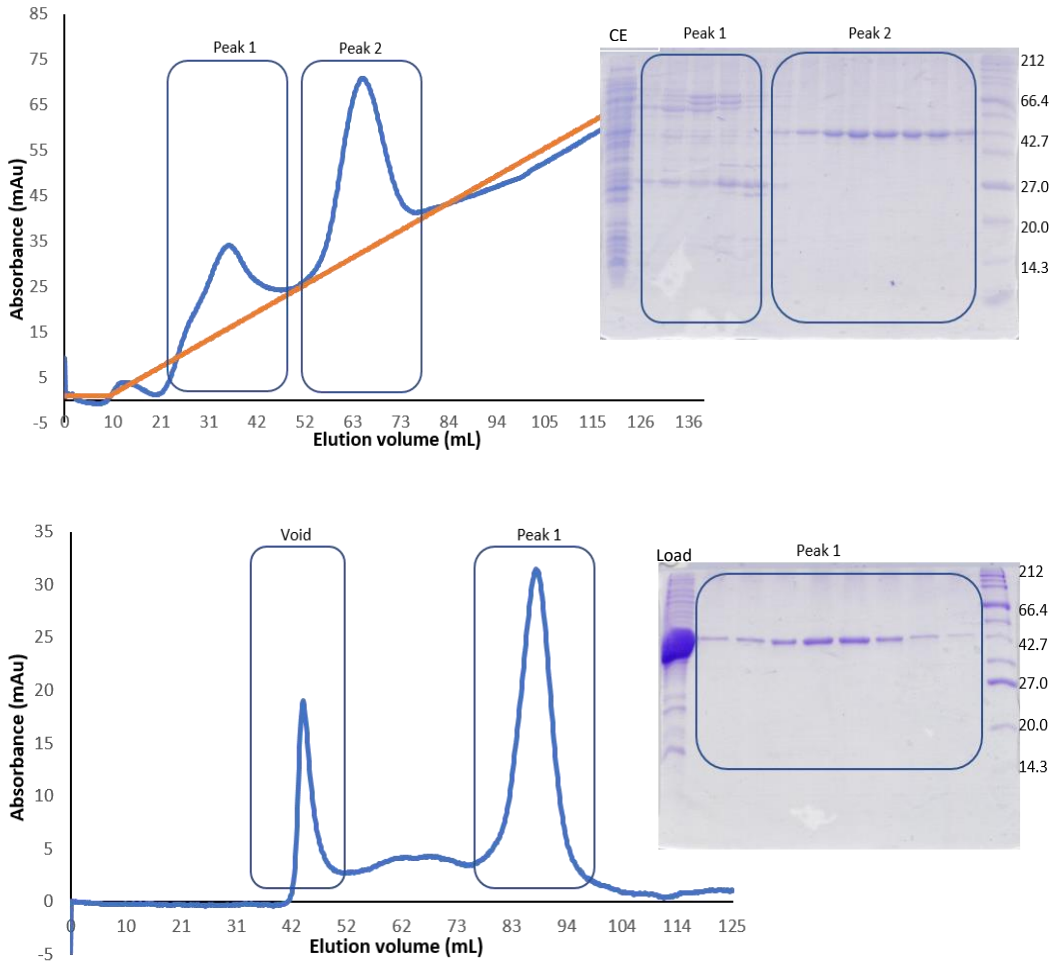


Figure 5.11. Protein expression and purification of USP4C1C2.
Nickel column chromatogram (top) showed that the protein was eluted from 25 % Buffer B on the second peak. The SDS-PAGE gel showing the second peak was an almost pure protein of USP4C1C2. Gel filtration Superdex 200 HiLoad 16/60 chromatogram (bottom) of USP4C1C2 indicates a high peak with elution volume 87.8 mL. This equal to 27.5 kDa based on a standard curve. The SDS-PAGE gel of USP4C1C2 showed the single protein was purified. (CE = crude extract)

The chromatogram of gel filtration purification showed a void volume and a single peak which elute at 87.8 mL. Compared to the standard curve, this volume refers to about 27.5 kDa. This molecular weight was less than actual size, indicates that USP4C1C2 was folded compactly. The SDS PAGE gel data showed that the USP4C1C2 protein was purified. SDS PAGE gel showed a clear single band at around 42.7 kDa. After collection and concentration, the total of

USP4C1C2 protein was 2.4 mg from 5 L cultures or about 4.8 mg/mL in 500 µL. The yield was 0.48 mg/L cultures.

The ubiquitin-AMC enzymatic activity assay was performed to investigate the activity of USP4C1C2 to cleave of Ub-AMC to Ub and AMC using the same method as for USP4FL and USP4htt. From the graph time vs fluorescence unit, the initial rate was taken for 13 minutes (Figure 5.12) and based on the calculation with a standard curve, the equation for USP4C1C2 was Y=0.11X + 2.35; R^2=0.98). The specific enzyme activity of USP4C1C2 was 4126.2 µM/min/mg enzymes (Figure 5.13).

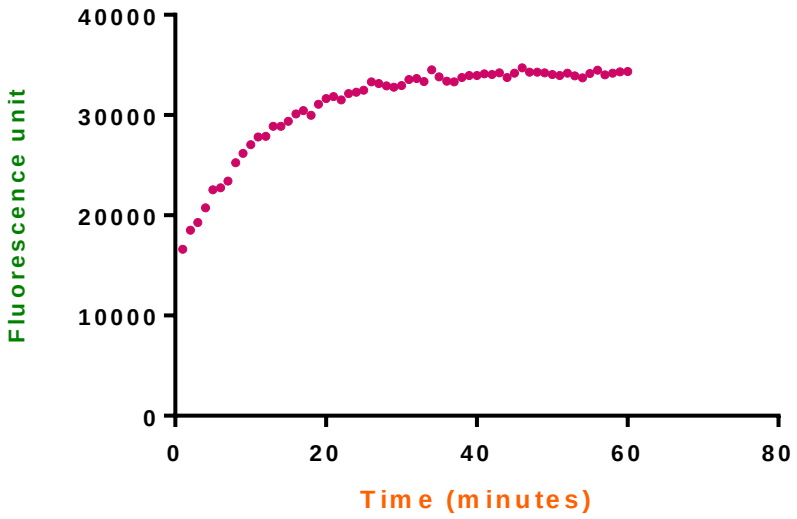


Figure 5.12. Graph of fluorescent of active protein USP4C1C2 with Ub-AMC USP4C1C2 concentration was 20 nM and Ub-AMC were 0.75 µM. All measured in duplicate and the plate was read every minute for 1 hour.

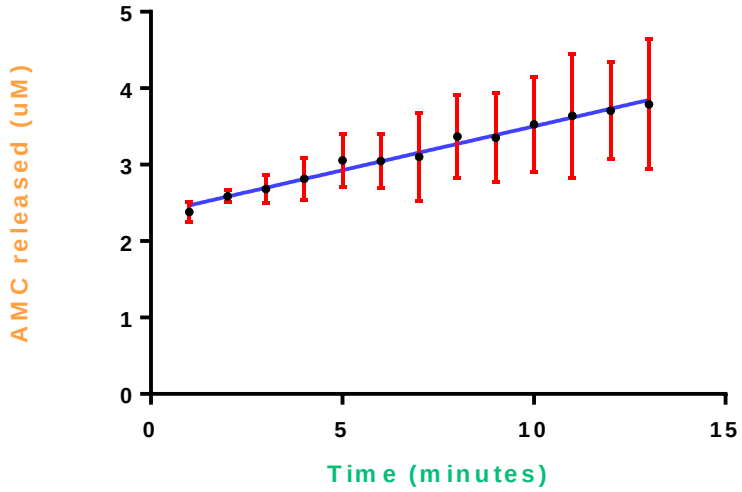


Figure 5.13. Enzymatic activity of active protein USP4C1C2 with Ub-AMC
The specific enzyme activity for USP4C1C2 was 4126.2 µM/min/mg enzymes.

USP4C1C2 only contains the USP4 catalytic core consisting of catalytic 1 and catalytic 2 sub-domains. The value of specific enzyme activity USP4C1C2 was higher almost 5 times compared to USP4FL and higher more than 125 times to USP4htt. The circular permutation construct USP4htt has the lowest specific activity.

5.5. Protein Expression and Purification of USP4ΔDU

The construct of USP4ΔDU, lacking the N-terminal DUSP and Ubl1 domain, was created by removing the DUSP domain and Ubl1 domains. USP4ΔDU construct has a Histidine tag and a sequence for TEV protease cleavage. Paulina Cygan (2009) had already made this construct and an expression protocol established in the Arctic express strain. The protein expression of USP4ΔDU needs 30 hours incubation at 15°C at 170 rpm in shaking incubator.

The purification method of USP4ΔDU involved the affinity chromatography, anion exchange chromatography, digestion by TEV protease and gel filtration as final purification.

The first purification result of USP4ΔDU by Ni column showed some impurities. The chromatogram of Nickle column showed a double peak and the protein target eluted after 24.3 % buffer B on the second peak. SDS PAGE gel indicates the high quantity of protein target (Figure 5.14).

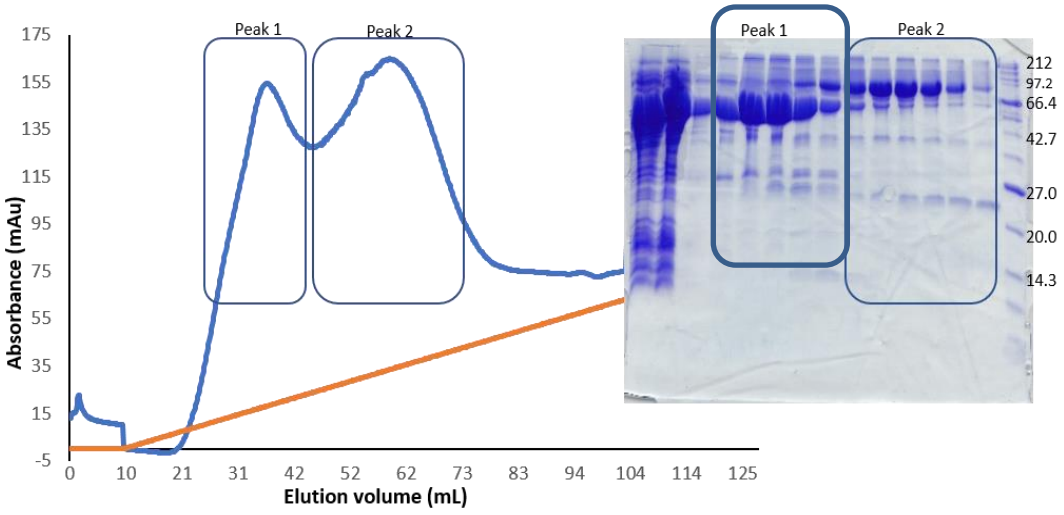


Figure 5.14. Chromatogram of Ni column purification of USP4ΔDU
Chromatogram showed a double peak which second big peak was the protein target. SDS-PAGE gel result showed bands that correspondent to USP4ΔDU at about 79.4 kDa.

Purification was continuing to the second purification, namely anion exchange. After anion exchange chromatography, the protein showed less of the degradation band (Figure 5.15). The product from anion exchange was a single wide band at a below than 97.2 kDa on the marker that correspondent to 79.4 kDa as the molecular weight of protein target.

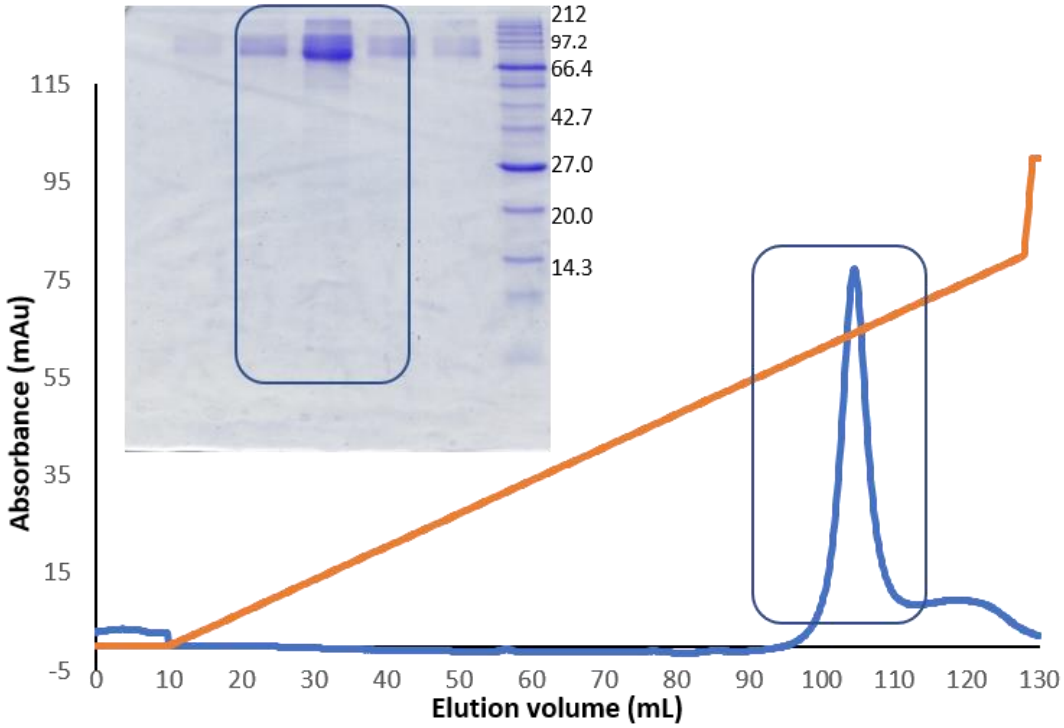


Figure 5.15. Chromatogram of anion exchanges purification of USP4ΔDU
Chromatogram showed a single high peak that contains the protein target. SDS-PAGE gel result showed bands that correspondent to USP4ΔDU (79.4 kDa)

After collection and concentration, the protein was digested with TEV protease to cleave the ENLYFQG sequence. SDS-PAGE gel showed that protein was digested well by TEV protease. The final purification of USP4ΔDU was gel filtration S200 16/60 (Figure 5.16), and the protein eluted at 67.2 mL. This volume equal to 156.5 kDa on the standard curve, higher than the actual size 79.4 kDa, about twice that indicates the protein was on dimer conformation when elutes in the solution. After collection, the protein concentration of the USP4ΔDU was 1.5 mg/mL in 300 µL volume. The total protein was 0.45 mg from 5L cultures. So, the yield was 0.09 mg/L cultures.

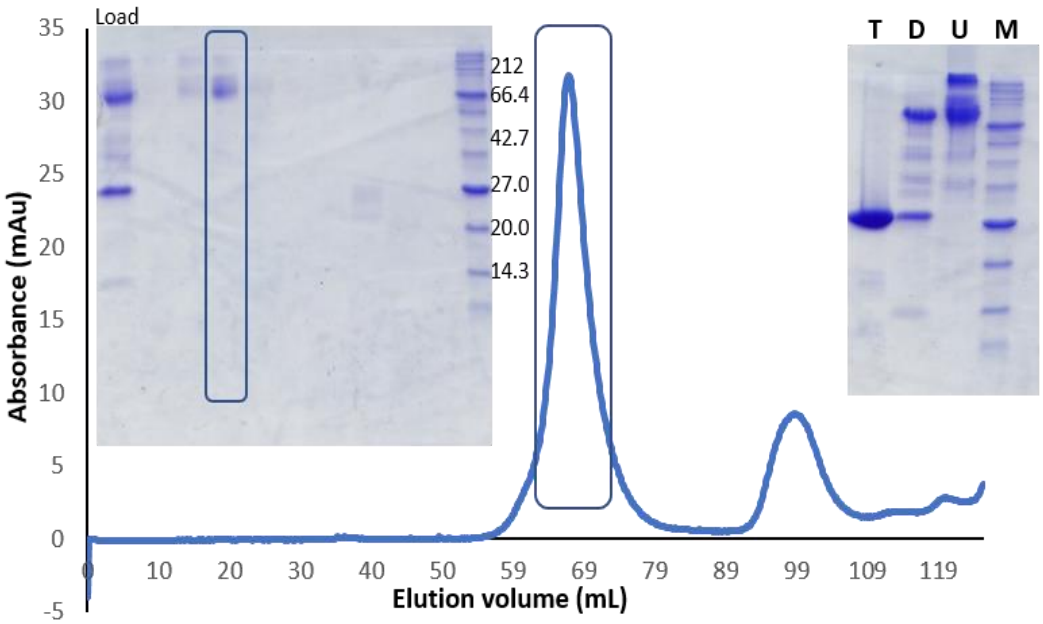


Figure 5.16. Chromatogram of gel filtration purification of USP4ΔDU
Chromatogram of superDex 200 HiLoad 16/60 showed a single high peak that contains the protein target at elution volume 67.2 mL. Equal with 156.5 kDa on the standard curve, that indicates the protein elutes as a dimer. SDS-PAGE gel result showed bands that correspondent to USP4ΔDU (79.4 kDa). The yield of this protein was 0.09 mg/L cultures. Before eluted on the gel filtration, the sample cleavage by TEV protease. SDS-PAGE on the right showed that sample was digested (D). Undigested sample (U) and TEV protease (T) as a control. Marker (M)

Bands from the SDS-PAGE gel of USP4ΔDU indicated a pure protein. A single big band showed above the marker band of 66.4 kDa which refer to 79.4 kDa of protein target. The multi-stages of protein purification increased the purities of protein.

The ubiquitin-AMC enzymatic activity assay was performed to check the activity of USP4ΔDU to cleave of Ub-AMC to Ub and AMC (Figure 5.17). Based on the calculation, the equation was $Y=0.085X+1.16$ ($R^2=0.99$). The USP4ΔDU has lower specific enzyme activity compared to the USP4C1C2. The specific

enzyme activity was 719.4 μM/min/mg of enzyme. USP4C1C2 has an activity about 6 times higher than USP4ΔDU.

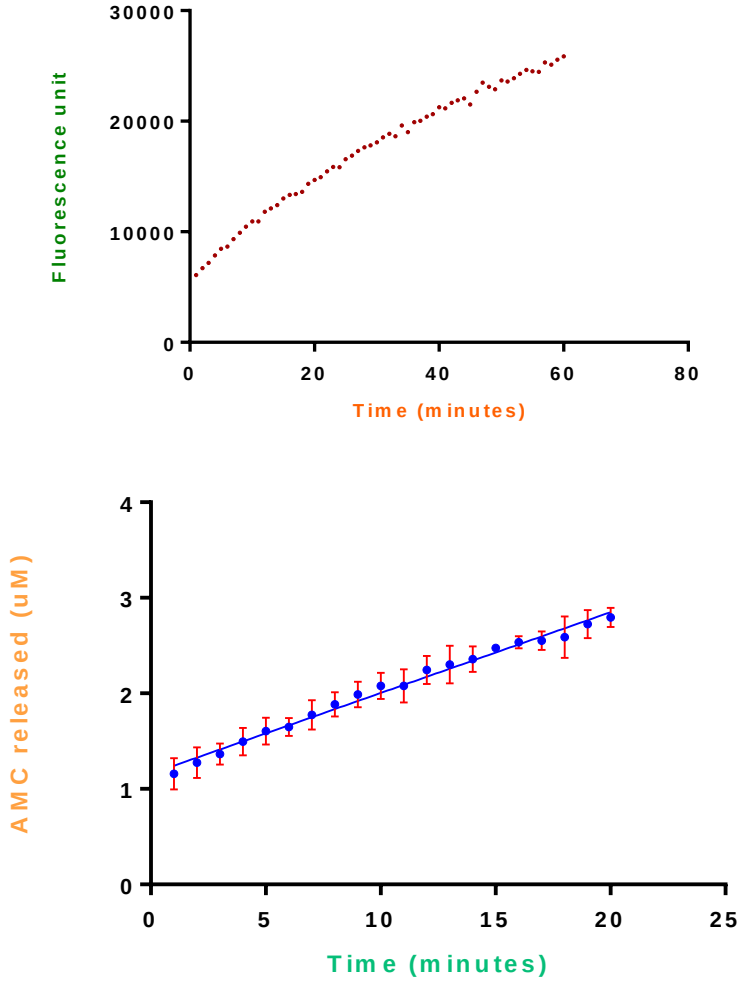


Figure 5.17. Enzymatic activity of active protein USP4ΔDU with Ub-AMC USP4ΔDU concentration was 50 nM, and Ub-AMC was 0.75 μM. All measured in triplicate and the plate was read every minute for 1 hour (Top). The specific enzyme activity for USP4ΔDU was 20.9 μM/min/mg enzymes (Bottom).

5.6. Protein Expression and Purification of USP4DU

Another modified construct based on the USP4 structure was USP4DU. This construct created by Affif Grazete (2015) consists of DUSP domain and Ubl1 domain only without a catalytic domain, Ubl2 and insert. The construct of

USP4DU was expressed well. From 3L cultures produces a plenty of protein. The purification of USP4DU was done by Nickel affinity chromatography and followed by the gel filtration method.

Ni column purification showed that the protein target eluted from 15% buffer B. SDS PAGE gel confirm the protein target with the high band at above protein marker 27 kDa. The protein USP4DU has 29.2 kDa (Figure 5.18).

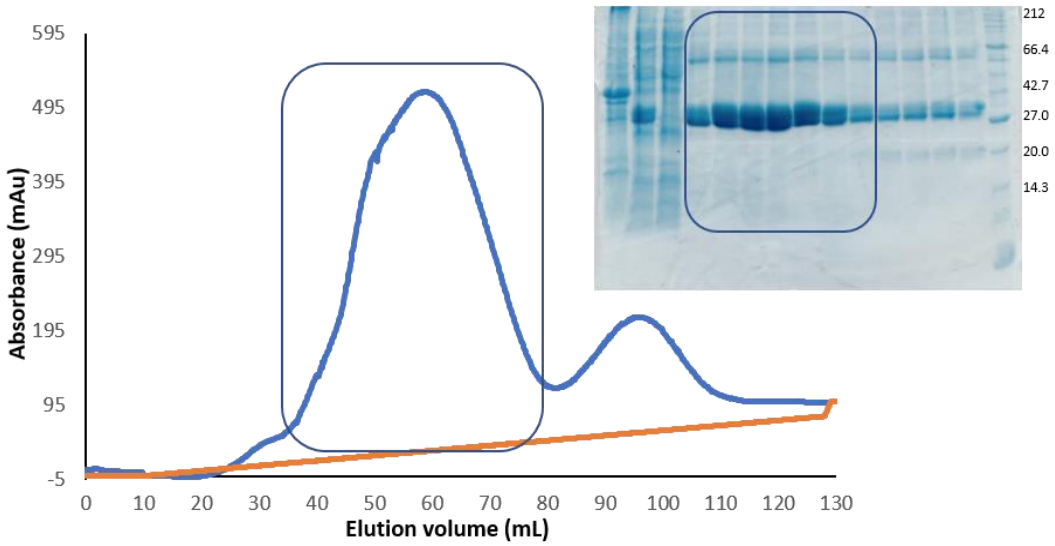


Figure 5.18. Chromatogram of affinity chromatography of USP4DU
Chromatogram showed a big high peak followed by small peak. The first peak contains the protein target. SDS-PAGE gel result showed bands that correspondent to USP4DU approximately at 29.2 kDa.

The chromatogram of gel filtration purification showed a high quantity of product. There was a double peak, but the protein target was on the second peak at the elution volume was 85.6 mL. This elution volume refers to 33 kDa based on the standard curve showed that USP4DU was on the monomer formation. This size almost equal to the actual size of the protein USP4DU. SDS

PAGE gel showed a high quantity of protein product at the right molecular weight (Figure 5.19).

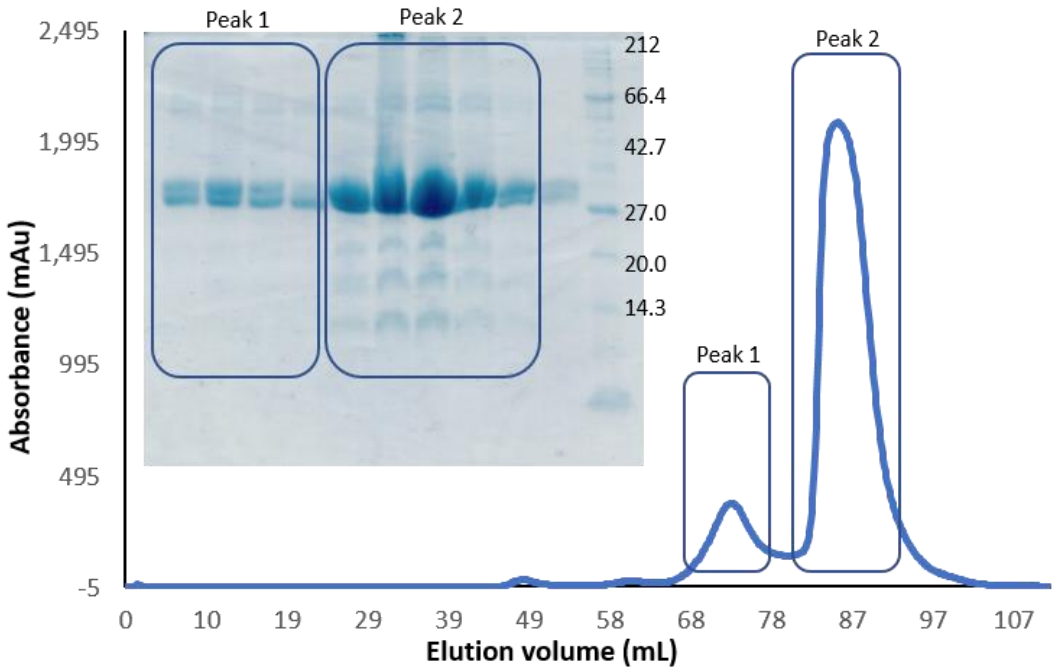


Figure 5.19. Chromatogram of gel filtration purification of USP4DU
Chromatogram of superdex 200 HiLoad 16/60 showed a low peak followed by high peak. The second peak contains the protein target. SDS-PAGE gel result showed bands that correspondent to USP4DU (29.2 kDa). The elution volume was 85.6 mL and equal to 33 kDa based on the standard curve. The yield of the protein was 2.6 mg/L cultures.

5.7. Discussion

In summary, the construct of USP4FL, USP4C1C2, USP4ttΔUbl2, and USP4DU relatively well expressed compared to the USP4htt and USP4ΔDU (Table 5.2). The construct which contains DUSP Ubl1 and at least, with catalytic core domain, C1C2, have a higher yield. This domain, the DUSP-Ubl1 and C1C2 influence the protein expression and purification. Both domains have a role for the stabilisation protein expression. The constructs USP4htt and USP4ΔDU have less yield. USP4htt and USP4ΔDU have no DUSP Ubl1 domain.

Table 5.2. Comparison of yield of USP4 modified construct after gel filtration purification using Superdex 200 HiLoad 16/60

The construct	Yield (mg per L culture)
USP4FL	0.7
USP4htt	0.05
USP4C1C2	0.48
USP4httΔUbl2	3.63
USP4ΔDU	0.09
USP4DU	2.6

Optimisation to increase the yield of protein expression and purification was performed for some constructs. USP4FL was optimised using expression in a different strain. The USP4FL was expressed in the LOBSTR-BL21 (DE3)-RIL strain. This is another model for expression, and the objective is to increase the purity and yield of USP4FL. LOBSTR-BL21(DE3)-RIL was ideal for a low-expressing protein target. This strain has modified of ArnA and SlyD, the most contaminant in *E.coli*, which have reduced affinities to Ni resins (Anderson, et al., 2013). However, the yield of USP4FL did not increase significantly.

USP4htt purification also applied some modification. First, as mention in the result, the purification was combined with ubiquitin to stabilise the protein product. Secondly, the buffer in Ni^{2+} column purification was changed to 150mM NaCl, 50 mM Tris-Cl, 20 mM Imidazole and 10% glycerol for buffer A. For buffer B was 300 mM NaCl, 50 mM Tris-Cl, 500 mM Imidazole and 5% glycerol. Previously was 300 mM of NaCl in Buffer A. Reducing the concentration of salt increased the protein solubility. Thirdly, to reduce the

oxidation damage and increase the yield, DTT was added to a concentration of 1 mM in the fractions when concentrated the protein before running gel filtration. Based on this optimisation, the total protein yield of USP4htt increasing three times, from 0.1 mg to be 0.3 mg from 5 L cultures or about 0.05 mg/L cultures.

The elution volume data showed the constructs followed the principle of size exclusion purification. The bigger molecular size of protein will elute first. The interesting data showed by USP4ΔDU construct (Table 5.3). Elution volume of USP4ΔDU was 67.2 mL equal with 156.5 kDa based on the standard curve. This size around double compares to actual its molecular weight at 79.4 kDa. It can be indicating this construct has a dimer formation in the protein expression. The other constructs were in monomer conformation in the solution. The construct of USP4DU has an agreement with USP11DU and USP15DU. All the DU constructs were in monomer (Harper, et al., 2011; Harper, et al., 2014).

Table 5.3. Comparison of elution volume of USP4 modified construct based on gel filtration purification Superdex 200 HiLoad 16/60

The construct	Elution volume (mL)	Mass prediction (kDa)	Actual size (kDa)
USP4FL	72.8	97.8	111
USP4htt	82	44.8	50.5
USP4httΔUbl2	87.2	28.9	40.7
USP4C1C2	87.8	27.5	42.7
USP4ΔDU	67.2	156.5 (dimer)	79.4
USP4DU	85.6	33	29.2

Specific enzyme activity data showed that USP4C1C2 has the highest activity, followed by USP4FL, USP4ΔDU and USP4htt as the lowest. Clerici, et. al., (2014) mentioned that USP4C1C2 has the enzymatic activity. The minimum domain for activity was catalytic core, USP4C1C2.

USP4FL has DUSP-Ubl1 and catalytic core. On the other hand, USP4htt, the construct without DUSP-Ubl1 and insert, has the weakest activity (Table 5.4). This data correlates with Clerici, et al., (2014) that the DUSP Ubl1 domain has effects increasing the cleavage activity. DUSP-Ubl domain increases the Ubiquitin dissociation from the USP4 by changes the conformation of the catalytic domain.

Table 5.4. Specific Enzymatic Activity of USP4 construct

The Construct	Enzyme activity (μM/min/mg)
USP4FL	843.5
USP4htt*	32.5
USP4C1C2	4126.2
USP4ΔDU	719.4

*= single experiment

USP4htt does not contain the DUSP-Ubl1 domain and the insert domain. DUSP-Ubl1 and insert domain of USP4 has a role on the catalytic activity. The insert has a role in a part of the USP4 binding to Ubiquitin. In this study, the purification of USP4htt was quite difficult. The protein tends to degradation. The data for specific enzyme activity was only from a single experiment. This construct still needs optimisation to increase the yield.

The kinetic parameter data showed that this USP4FL was an active enzyme. There was a sequence difference between this USP4FL construct and Clerici et al., (2014). This USP4FL was consist of 963 amino acid, and USP4FL of Clerici's construct only contains 925 residues on C terminus. Possibly this difference caused the differences in kinetic parameters. The differences type of substrates (TAMRA-Ubiquitin) and the concentration was used could affect the result.

Comparison with the homologues USP11 and USP15 showed that in USP11, the USP11FL has higher Km than USP11C1C2 (originally USP11CatΔUbl2 (Harper, et al., 2014). USP15FL also has higher Km than USP15ΔDU (originally USP15CD) (Clerici, et al., 2014). The Km data of USP11FL and USP15FL were higher compare to the USP11C1C2 and USP15ΔDU, respectively. All these data confirmed that the full-length construct of USP11FL and USP15FL have the highest activity.

CHAPTER 6

Result and Discussion: Characterization of active site mutant of USP4 and Crystallisation trials (USP4FL, USP4htt, USP4httΔUbl2, USP4C1C2, and USP4ΔDU)

This chapter focuses on the expression and purification of the complex of USP4 active site mutants with Ubiquitin variants, such as Ubiquitin, UbGGG, and linear diubiquitin. The protein complexes were subjected to Crystallisation trials to characterise the structure of USP4.

6.1. Protein expression and purification of USP4FL H881N and USP4FL C311S with Ubiquitin Variants

Co-expressions were performed of an active site mutant of USP4FL with the different Ubiquitin variants. USP4FL-H881N was co-expressed with Ubiquitin, diubiquitin, and diubiquitin-L73X, while USP4FL-C311S has an additional mutation to remove flexible residues at the C-terminus. USP4FL-C311S has Tyrosine (Y) residue number 928 mutated to stop codon (Y928X). USP4FL-C311S-Y928X was co-expressed with diubiquitin and UbGGG.

6.1.1. Expression and Purification of active site mutant USP4FL-H881N with Ubiquitin variants

The expression and purification of an active site mutant USP4FL-H881N with ubiquitin was performed using the same method as USP4FL wild type. A Ni column was used for a first purification step. The chromatogram from Ni^{2+} affinity column showed a double peak, indicative of at least two protein species. At 15.1 % of buffer B, the protein started to elute. SDS-PAGE gel result confirmed it contained some degradation products (Figure 6.1). The protein of USP4FL, a 111 kDa, indicated by the thick band on the top of the gel is located

to below the 212 kDa marker. Ubiquitin, 7.5 kDa, was another product on the bottom of the gel.

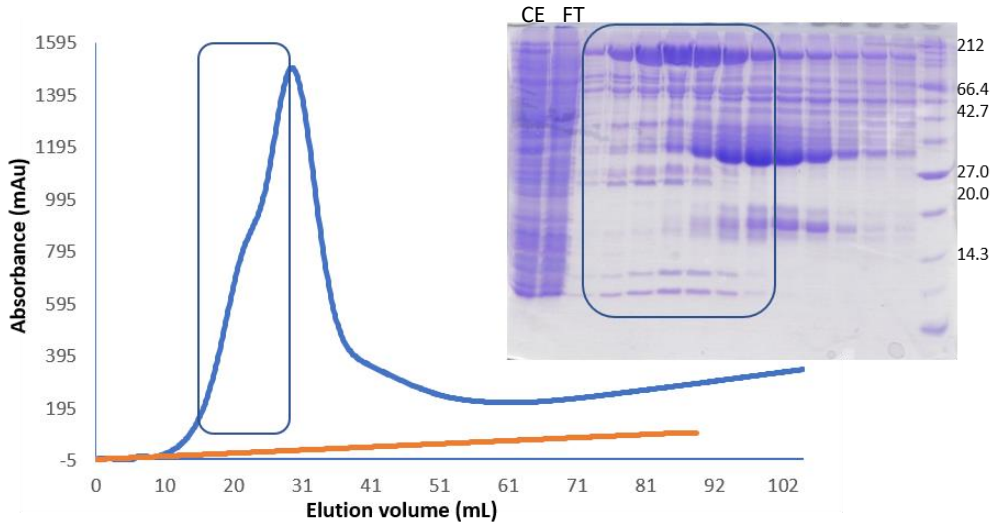


Figure 6.1. Nickel column purification of USP4FL-H881N with ubiquitin
Chromatogram of active site mutant USP4FL-H881N with Ubiquitin and SDS-PAGE gel from Nickel column purification showed the protein complex and some degradation product. The protein complex elutes from 15.1% Buffer B. The fraction on the blue box were collected and concentrated on continuing for gel filtration purification.
(CE = crude extract; FT= flow trough).

The chromatogram from the second step gel filtration purification of USP4FL-H881N – ubiquitin showed two peaks (Figure 6.2). The first peak was protein complex of active site mutant USP4FL-H881N and ubiquitin, and the second one was a degradation product, probably the DUSP-Ubl domains. The protein complex eluted at 66.2 mL which correspondent to 156.5 kDa of the standard curve. The elution volume indicates complex USP4H881N with Ubiquitin may contain monomer USP4FL-H881N with tetramer Ubiquitin. There was a polyubiquitin chain in the protein complex USP4FL-H881N with Ubiquitin. The active site mutant USP4FL-H881N with Ubiquitin pu rification had a yield of 2.6 mg from 5 L cultures (0.52 mg/L culture).

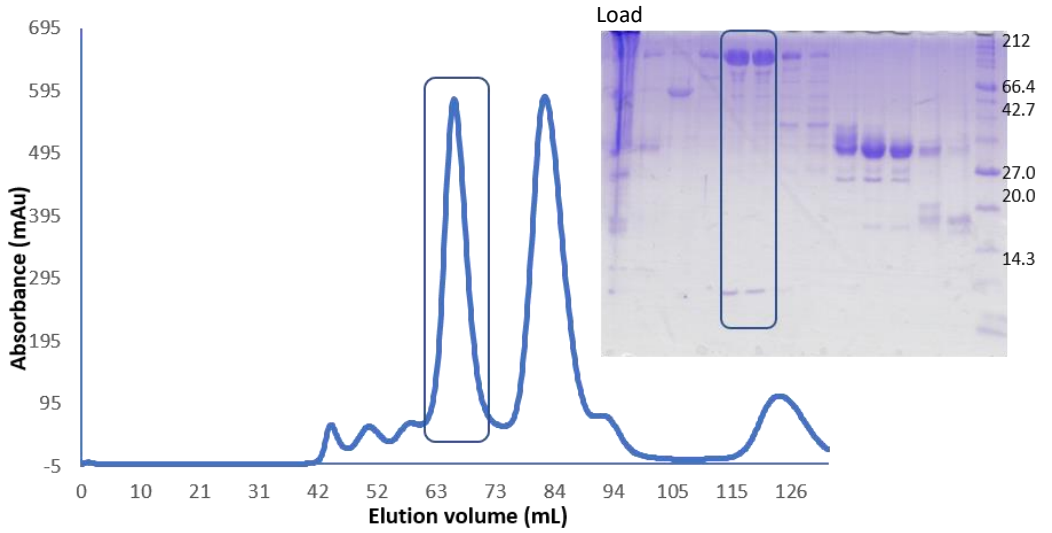


Figure 6.2. Size exclusion purification of USP4FL-H881N with ubiquitin
Chromatogram after gel filtration purification of active site mutant USP4FL-H881N with Ubiquitin showed the double peak using superdex 200 HiLoad 16/60. The result indicates USP4FL-H881N and ubiquitin expressed well. The elution volume was 66.2 mL which corresponds to 156.5 kDa using a gel filtration standard curve. SDS-PAGE gel showed purer band compared to Ni column purification with the correct size of molecular weight refer to protein marker.

Further complexes for active site mutant of USP4FL-H881N combined with diubiquitin and diubiquitin-L73X were also prepared. The Ni column purification of the complex USP4FL-H881N with diubiquitin showed small peaks in the first elution followed by single high peak (Figure 6.3). The protein complex was in the second peak after buffer B elution at 23.5%. The SDS PAGE gel confirmed that the first peak contained some protein degradation products and the second peak was the target protein. The USP4FL-H881N, 111 kDa, was below the protein marker band of 150 kDa, and diubiquitin, 17 kDa, was on the bottom at about 15 kDa.

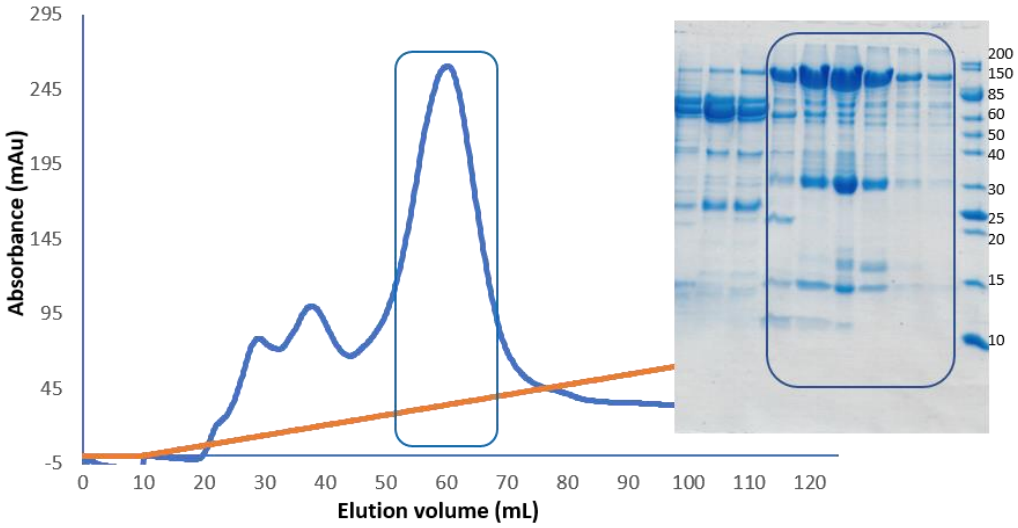


Figure 6.3. Nickel column purification of USP4FL-H881N with linear diubiquitin Chromatogram after nickel column purification of active site mutant USP4FL-H881N with linear diubiquitin and SDS-PAGE gel showed the protein complex and some protein degradation products. The protein complex elutes at 23.5 % buffer B.

The gel filtration purification showed three peaks on the chromatogram, and the highest peak was the protein complex. The elution volume of complex USP4FL-H881N with diubiquitin was 61.3 mL which equal to 257.5 kDa of the standard curve. It can be interpreted that protein complex of USP4FL-H881N with diubiquitin form a dimer of USP4FL-H881N and dimer diubiquitin (Figure 6.4). The final concentration of the protein complex active site mutant of USP4FL-H881N with diubiquitin was 10.85 mg/mL in 200 µL. The yield was 2.6 mg from 5 L cultures or approximately 0.52 mg/L culture. The protein complexes were used in Crystallisation trials and ESI-mass spectrometry analysis (Chapter 7).

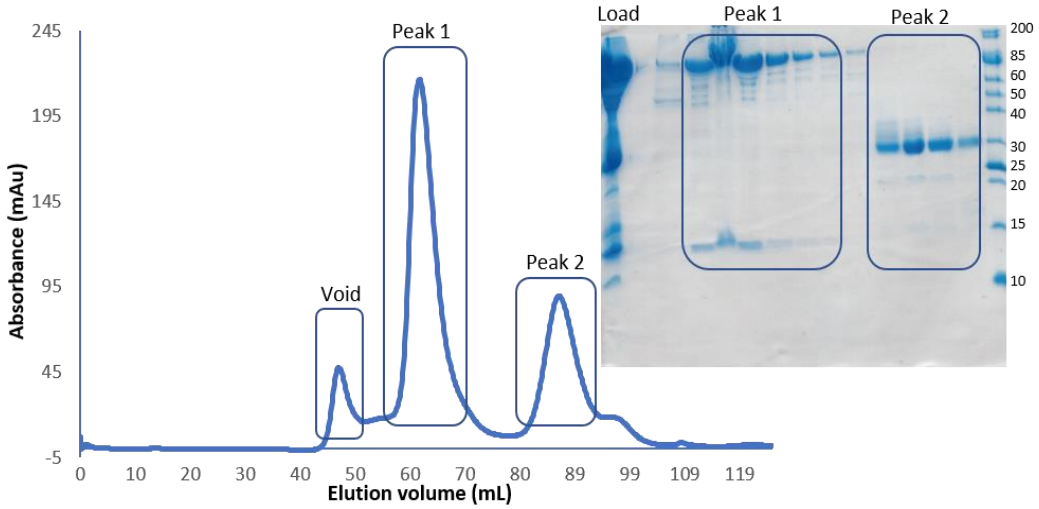


Figure 6.4. Size exclusion purification of USP4FL-H881N with linear diubiquitin
Chromatogram after gel filtration purification active site mutant USP4FL-H881N with linear diubiquitin using superdex 200 HiLoad 16/60 showed elution volume at 61.3 mL which correspondent to 257.5 kDa with a standard curve. SDS-PAGE gel on peak 1 indicates USP4FL-H881N with diubiquitin

The last protein complex prepared was USP4FL-H881N with diubiquitin-L73X. SDS-PAGE gel result of the Ni column purification step indicates the protein complex was expressed. Both USP4FL-H881N appear on the top of the gel-based below marker protein of 212 kDa a diubiquitin-L73X band was higher than 14.2 kDa protein marker.

The gel filtration purification showed two peaks and the complex was in the first peak. The protein complex eluted at 68.9 mL which correspondent to a standard curve of 135.5 kDa (Figure 6.5). This molecular weight indicates a monomer USP4FL-H881N bind with monomer diubiquitin-L73X to form a protein complex. The second peak was another protein, probably the degradation of DUSP-Ubl domains because it was about 30 kDa. The complex

of USP4FL-H881N and diubiquitin-L73X had the concentration 18.49 mg/mL in 200 µL. The yield was 3.7 mg from 5 L cultures or around 0.74 mg/L culture.

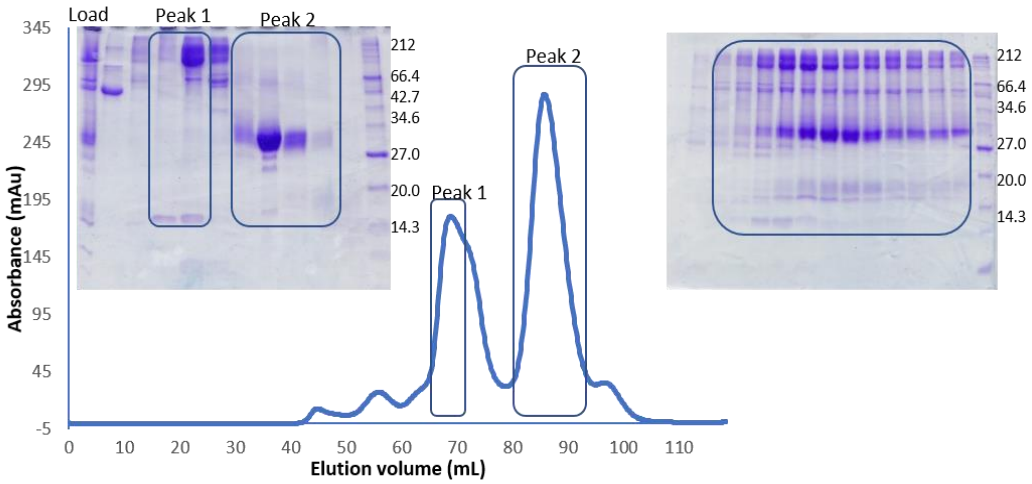


Figure 6.5. Gel filtration chromatogram of USP4FL-H881N with diUb-L73X.
SDS-PAGE gel result from Ni column fractions showed bands that correspondent to USP4FL-H881N (111,12 kDa) and an unclear band of diubiquitin-L73p on the bottom gel (17kDa) *(right)*. Chromatogram gel filtration S200 HiLoad 16/60 showed double peak which the first one contains the protein target, and the elution volume was 68.9 mL correspondent to 135.5 kDa at the standard curve. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4FL (111,12 kDa) and a band of diubiquitin-L73X on the bottom gel (17 kDa). The concentration was 18.49 mg/mL in 200 µL. The yield was 3.7 mg from 5 L cultures (mg per L of culture).

Purification process for USP4FL-H881N with diub-L73X was followed by a second gel filtration using analytical column S200 10/300 (GE Healthcare) for mass spectrometry analysis (Figure 6.6) (Explain in chapter 7). This purification has a purpose to desalt the sample, and the final concentration was 5.9 µM.

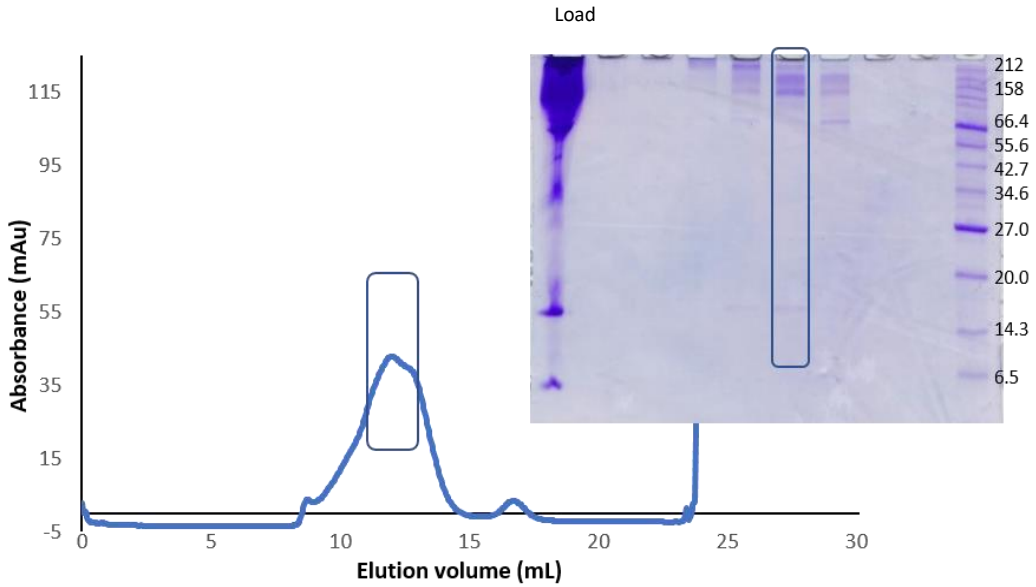


Figure 6.6. Gel filtration chromatogram of the USP4FL-H881N with diUbL73X
Chromatogram using analytical column S200 10/300showed a big peak which contains both proteins. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4FL (111,12 kDa) and a band of diubiquitin-L73X on the bottom of the gel (17 kDa).

6.1.2. Expression and Purification of active site mutant USP4FL-C311S in complex with Ubiquitin Variants

The other USP4-FL mutant generated was C311S and deleting the C-terminus by mutating Tyr928 to a stop codon on the C terminus. The double mutant of USP4FL-C311S-Y928X with diubiquitin was used to form a protein complex. The mutation of USP4FL-Y928X removed serine and glycine residues on the C-terminus of USP4FL to decrease the flexibility of the C-terminus and increase the opportunity of forming crystal protein complex of USP4FL-C311S-Y928X-diubiquitin.

The two proteins were co-expressed and then the protein complex USP4FL-C311S-Y928X with diubiquitin obtained by Ni column and gel filtration

purification. The nickel column chromatogram showed a wide double peak with the second peak higher than the first one. The SDS-PAGE gel revealed the presence of some degradation protein. The USP4FL-C311S-Y928X was at the expected molecular weight 111.12 kDa, and a weaker diubiquitin band was at the bottom of the gel in agreement with the protein marker (Figure 6.7).

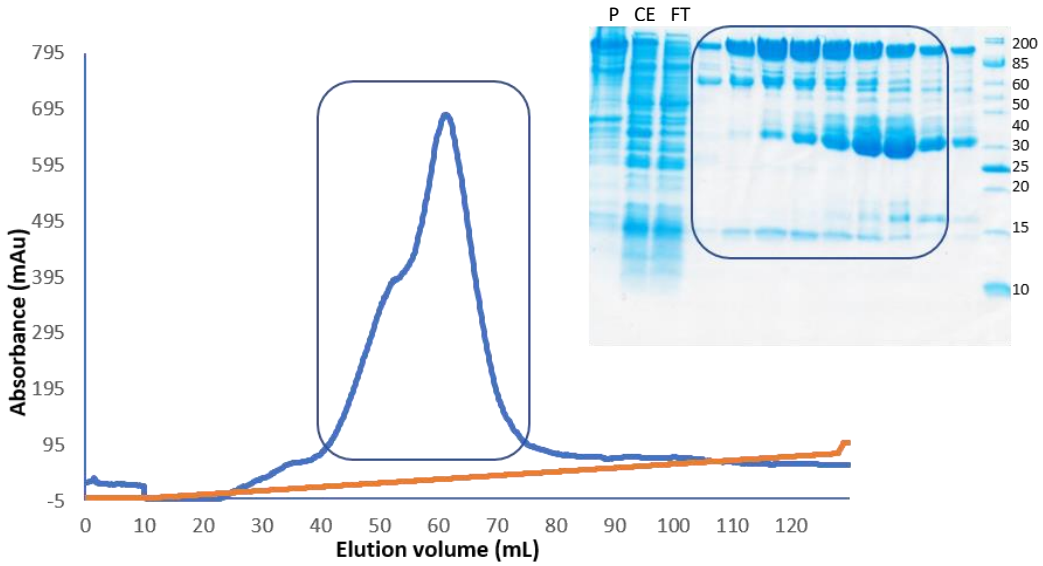


Figure 6.7. Nickel column chromatogram of the USP4FL-C311S-Y928X and diubiquitin co-expression.
The chromatogram showed a single wide peak. The protein complex start elutes at 20.5 % Buffer B. (P = Pellet; CE=crude extract; FT=flow trough)

The purity of USP4FL-C311S-Y928X seems increased although the gel filtration chromatogram did not show a single peak on the elution after void volume. The complex of USP4FL-C311S-Y928X with diubiquitin elutes at a volume 62.1 mL equal to 240.6 kDa of the standard curve. This molecular weight indicates a dimer USP4FL-C311S-Y928X bind with monomer diubiquitin.

SDS-PAGE gel showed less degradation, a strong band of USP4FL-C311S-Y928X, and weaker diubiquitin band at the bottom of the gel (Figure 6.8). The concentration of protein complex was 6.1 mg/mL in 300 µL of volume protein from 5 L preparation, so the total protein was 1.8 mg. The yield was 0.36 mg/L cultures.

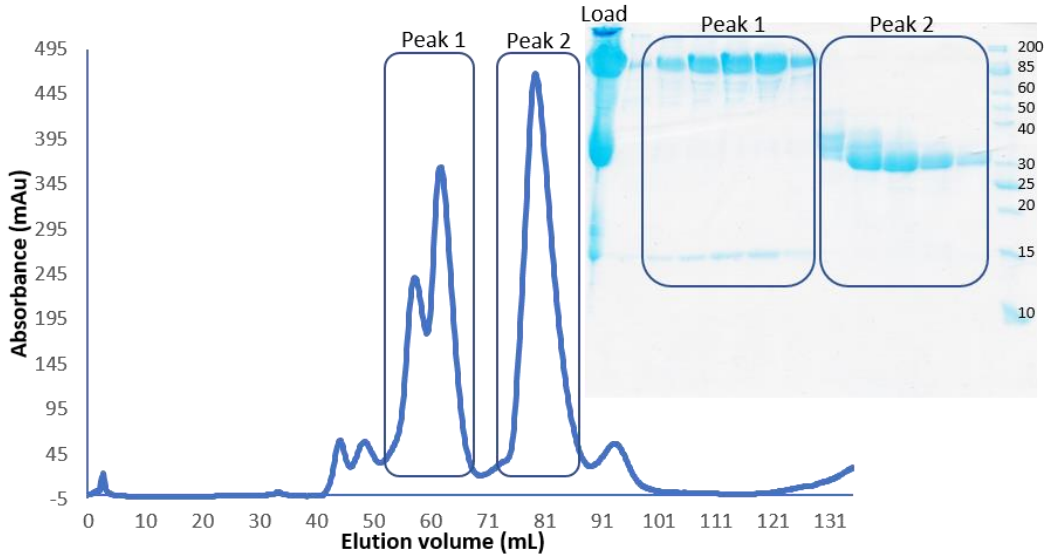


Figure 6.8. Gel filtration chromatogram of USP4FL-C311S-Y928X with diubiquitin
Gel filtration of superdex 200 HiLoad 16/60 showed the protein complex was on the peak 1, consist of USP4FL and diubiquitin on the 62.1 mL of elution volume correspondent to 240.6 kDa of the standard curve. Peak 2 was correspondent to DUSP-Ubl domain.

The second protein complex was the complex of USP4FL-C311S-Y928X with Ub-GGG. This protein complex was successfully co-expressed and purified. The Ni column purification showed the protein complex elutes early upon elution with Buffer B, at about 2%. The SDS-PAGE gel showed there was a degradation product (probably the USP4 N-terminal domains), but the protein target has the intense band. This band was above protein marker 85 kDa (Figure 6.9) at an expected molecular weight of 111 kDa.

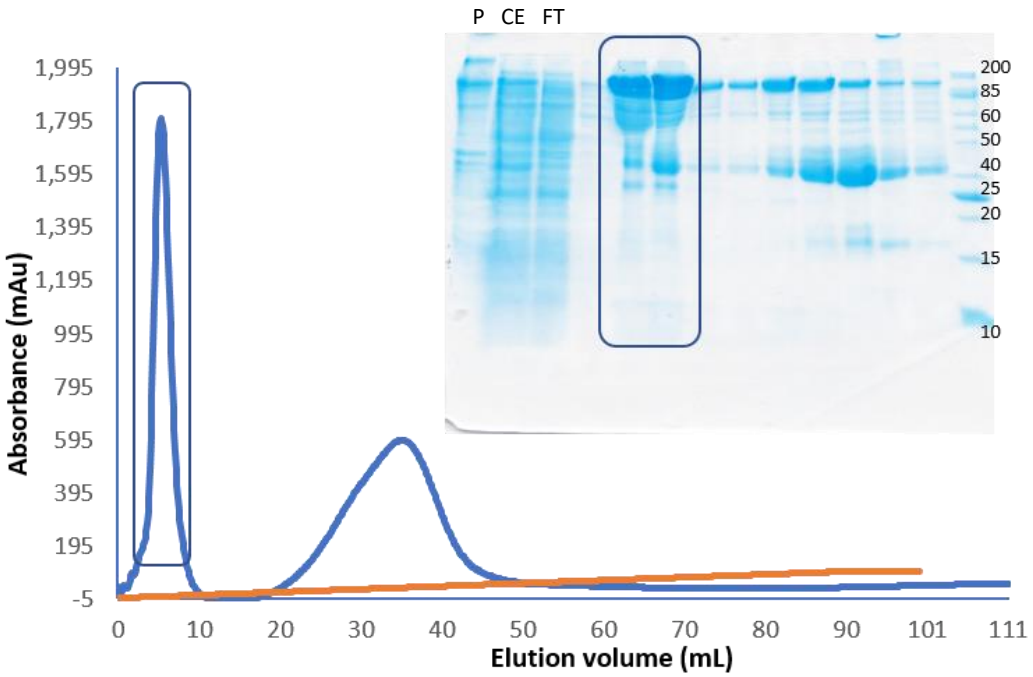


Figure 6.9. Nickel column chromatogram of USP4FL-C311S-Y928X with Ub-GGG.
The protein complex elutes early at 2 % of Buffer B. SDS PAGE gel showed the protein complex was around below 200 kDa marker on the top and less amount of Ub-GGG in the bottom of the gel. (P=pellet; CE=crude extract; FT=flow trough).

The gel filtration purification indicates pure protein of USP4FL-C311S-Y928X and UbGGG. After a void volume peak, there were another two smaller peaks, and then a single high peak elutes at 62.9 mL which correlates the molecular weight 225 kDa of the standard curve. This data showed that complex USP4FL-C311S-Y928X with UbGGG form a dimer USP4FL-C311S-Y928X with monomer UbGGG.

SDS-PAGE gel of USP4FL-C311S-Y928X with UbGGG showed the protein complex well expressed (Figure 6.10). Compared to the complex with diubiquitin, complex USP4FL-C311S-Y928X gave the concentration about 5.8 mg/mL in the same volume, and the total protein was 1.74 mg from 5 L

cultures. The yield was 0.35 mg/L culture. There was no significant difference in total protein yield between the complex of active site mutant of USP4FL-C311S-Y928X with diubiquitin or with UbGGG. The yield was 0.36 mg/L and 0.35 mg/L cultures respectively.

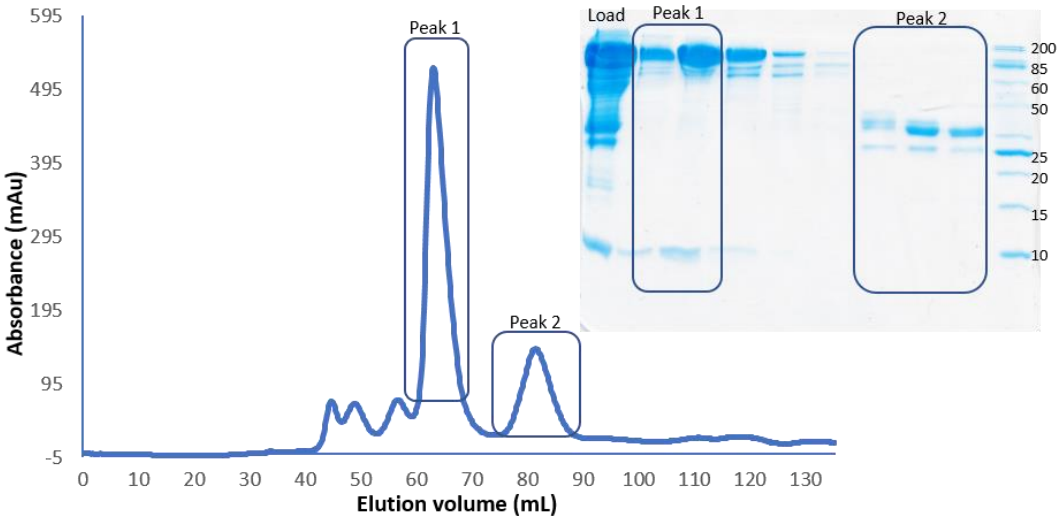


Figure 6.10. Gel filtration chromatogram of USP4FL-C311S-Y928X with Ub-GGG
Purification of USP4FL-C311S-Y928X with UbGGG using Superdex 200 HiLoad 16/60 elutes the protein target in peak 1 at volume 62.9 mL, with a correspondent with 225 kDa of the standard curve. This size indicates dimer USP4FL-C311S-Y928X bind with monomer Ub-GGG. Peak 2 was on USP4FL purification usually was a DUSP-Ubl domain.

6.1.3. Expression and purification of active site mutant USP4FL-C311S

The USP4FL-C311S mutant was also expressed on its own only. This construct was expressed and purified separately without any ligand. The Ni column purification showed two peaks. The second peak of the chromatogram usually was equal or higher and contained more degradation product of the DUSP Ubl1 domains. SDS-PAGE gel confirmed that the USP4FL-C311S was expressed

(Figure 6.11). The intense band at the top of the gel showed that protein was expressed (predicted Mw=111,112 kDa).

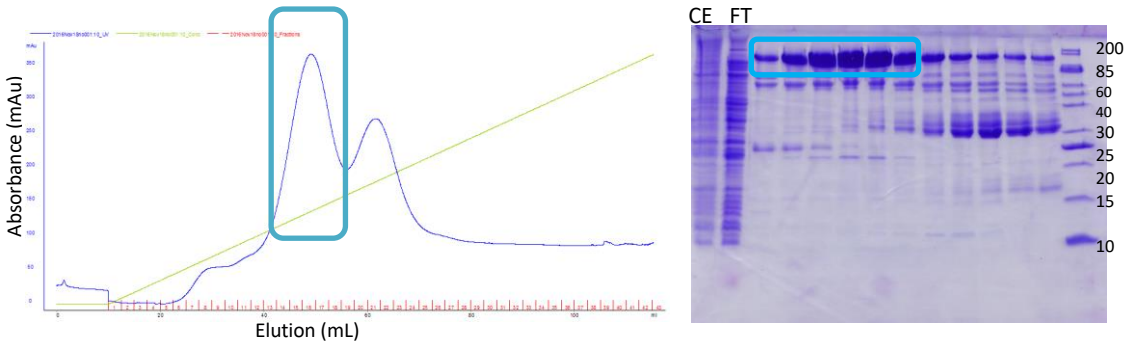


Figure 6.11. Nickel column chromatogram of USP4FL-C311S
The chromatogram showed that protein sample elutes at 22% Buffer B. SDS-PAGE gel indicates a thick band at the top about 111 kDa and some degradation product.

The gel filtration purification only showed a single peak rather than two as the other USP4FL gel filtration purification. The elution volume was 68 mL equal with a standard curve at 146 kDa compared with 111 kDa as actual size. SDS-PAGE gel showed that the purity of protein target increased. Compared to the purification of USP4FL-H881N, the USP4C311S mutant gave less of the degradation product (Figure 6.12). There was no degradation of N terminus DUSP Ubl domain on the SDS-PAGE gel result. The protein concentration was 25.6 mg/mL in 300 µL from 5 L cultures, so the total protein was 7.7 mg (1.54 mg pure protein/L of culture). This protein sample was prepared for ITC assays (Chapter 7).

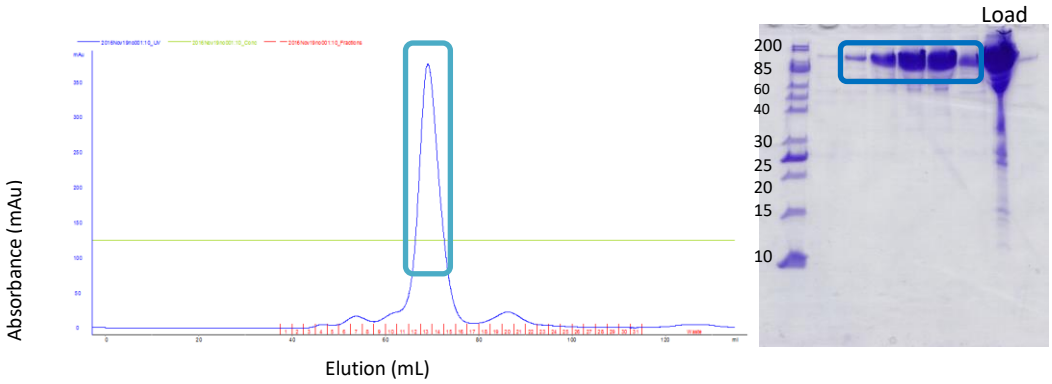


Figure 6.12. Gel filtration chromatogram of USP4FL-C311S
The chromatogram of superdex 200 HiLoad 16/60 showed a single peak elutes at 68 mL correlate with 146 kDa of the standard curve. Almost equal with actual size 111 kDa. SDS PAGE gel showed less degradation product without no contamination of N terminus DUSP domain.

The data of elution volume and molecular weight prediction showed an active site mutant USP4FL could be monomer or dimer in complex with Ubiquitin variant (Table 6.1) depend on the ligand (Ubiquitin variant) and the type of mutation (H881N or C311S-Y928X). The construct active site mutant USP4FL-C311S can produce the highest yield. In the mutation H881N, complex with Ubiquitin and diubiquitin has not affected the yield, but complex with diubiquitin-L73X will increase the yield ~0.5-fold. Double mutation C311S-Y928X in complex with diubiquitin and Ub-GGG also reveal the similar yield.

Table 6.1. Elution volume and molecular weight prediction of protein complex USP4FL-H881N and C311S-Y928X with Ubiquitin variant

Constructs	Yield (mg/L culture)	Elution volume (mL)	Mass Prediction (kDa)	Actual size (kDa)	Interpretation
USP4FL-H881N with Ubiquitin	0.52	66.2	156.5	111 + 8.5	Monomer USP4FL-H881N bind with pentamer Ubiquitin

Table 6.1. (Continued) Elution volume and molecular weight prediction of protein complex USP4FL-H881N and C311S-Y928X with Ubiquitin variant

USP4FL-H881N with diubiquitin	0.52	61.3	257.5	111 + 17	Dimer USP4FL - H881N bind with dimer diubiquitin
USP4FL-H881N with diubiquitin-L73X	0.74	68.9	135.5	111 + 17	Monomer USP4FL H881N with Monomer diubiquitin L73X
USP4FL-C311S-Y928X with diubiquitin	0.36	62.1	240.6	111 + 17	Dimer USP4FL C311S-Y928X with monomer diubiquitin
USP4FL-C311S-Y928X with UbGGG	0.35	62.9	225	111 + 8.5	Dimer USP4FL C311S-Y928X with monomer UbGGG
USP4FL-C311S	1.54	68	146	111	Monomer USP4FL C311S

6.2. Protein expression and purification of USP4htt-H106N and C311S active site mutants

The active site mutant of USP4htt-H106N was co-expressed and purified to form a complex with diubiquitin, UbGGG, and diubiquitin-L73X. Also, the USP4htt-C311S active site mutant was co-expressed and purified to form a complex with diubiquitin and UbGGG. All these USP4-ubiquitin complexes were used for Crystallisation trials.

6.2.1. Protein Expression and Purification of USP4htt-H106N active site mutant with Ubiquitin Variants

Initially, the active site mutant USP4htt-H106N was co-expressed and purified with diubiquitin. The USP4htt protein on its own is not very stable (Chapter 5).

It was hoped that the protein could be stabilised when in complex with ubiquitin variants. The active site mutant of USP4htt-H106N is inactive in cleaving diubiquitin (H106N in this construct corresponds to H881 in full-length USP4).

The Ni^{2+} affinity column chromatogram of the active site mutant USP4htt-H106N co-expressed with diubiquitin showed a single peak. The complex started to elute from 21 % of buffer B. The SDS-PAGE gel indicated the protein complex result as showed on the gel in around 50.5 kDa indicated by protein marker (Figure 6.13). The diubiquitin also showed in the bottom of the gel according to the 14.3 kDa of a protein marker, the actual size of diubiquitin was 17 kDa. Both proteins appeared in the same lane.

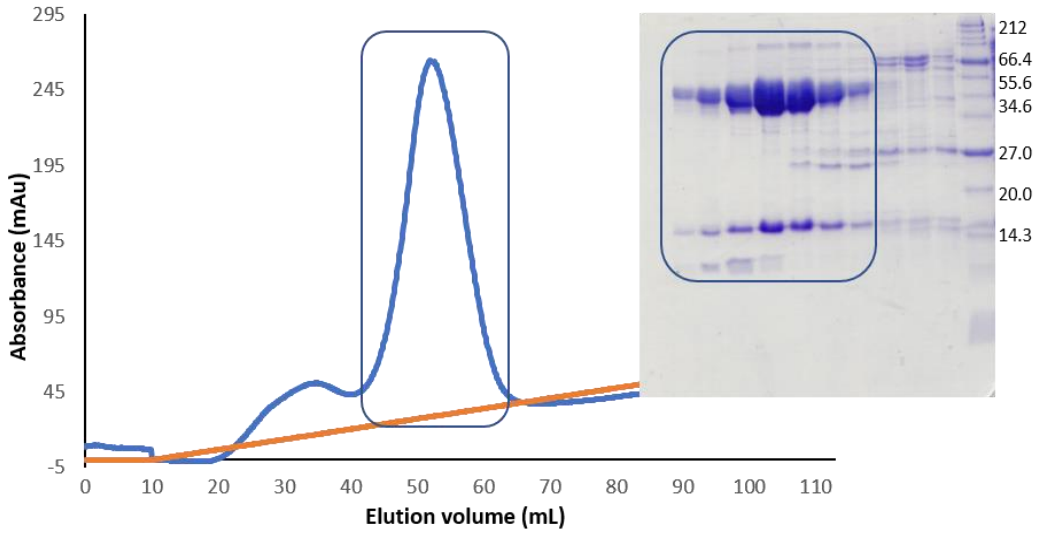


Figure 6.13. SDS-PAGE gel and chromatogram of Nickle column of USP4htt-H106N with linear diubiquitin.
The graph showed a single peak of complex USPFhtt-H106N with linear diubiquitin. USP4htt was expected on 50.5 kDa with linear diubiquitin on 17kDa. The SDS-PAGE gel indicates both products, USP4htt-H106N, and diubiquitin.

The chromatogram of gel filtration showed an elution volume of protein complex at 67.04 mL with a single high peak after void volume. This elution volume correspondent to 158.6 kDa based on the standard curve. This size showed that complex USP4htt-H106N was consist of dimer USP4htt-H106N and trimer diubiquitin (Figure 6.14).

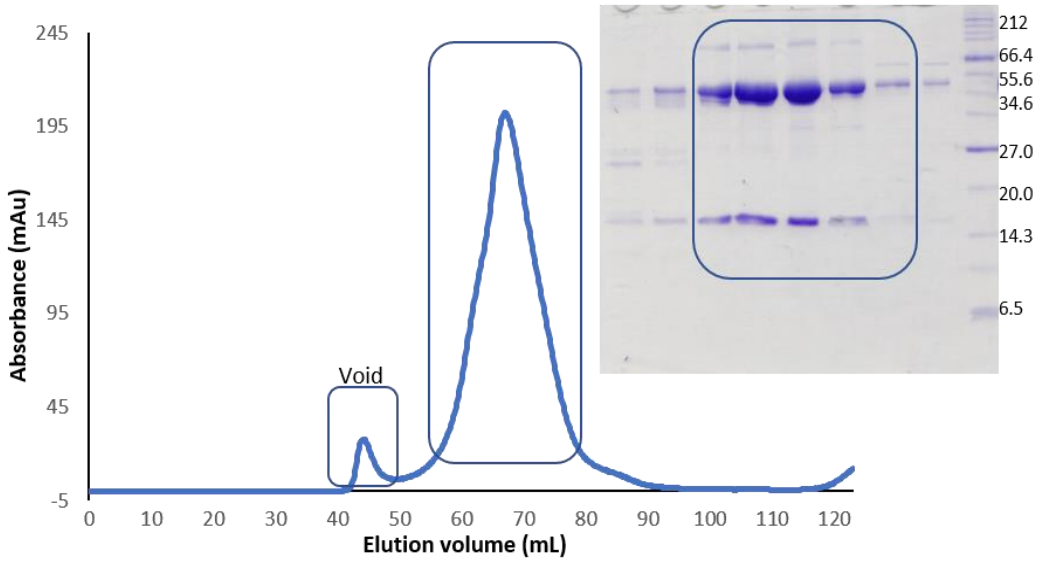


Figure 6.14. SDS-PAGE gel and chromatogram after gel filtration from USP4htt-H106N coexpression and purification with linear diubiquitin.
The graph showed that complex USPFhtt-H106N with linear diubiquitin elute at 67.04 mL which equal to 158,6 kDa on the standard curve. Indicates dimer USP4htt-H106N binds with trimer diubiquitin. USP4htt-H106N was expected on 50.5 kDa with linear diubiquitin on 17 kDa as shown on the SDS-PAGE gel.

The SDS-PAGE gel also showed the complexity of the active site mutant USP4htt-H106N with diubiquitin were in correct molecular weight as indicated by protein marker at about 55.6 kDa and 14.3 kDa. The total protein from this purification was 0.98 mg from 5 L cultures or in average was 0.2 mg/L culture. The active site mutant of USP4htt-H106N also combined with UbGGG to form a protein complex. The expression using co-expression method and the

purification by affinity chromatography and gel filtration purification. Nickel column purification (Figure 6.15) showed the protein target elutes from 21% buffer B, after a low peak, with a single peak. SDS-PAGE gel showing the protein complex expressed well. Both of USP4htt-H106N and UbGGG was on the gel. Some degradation still appears on the top of the gel.

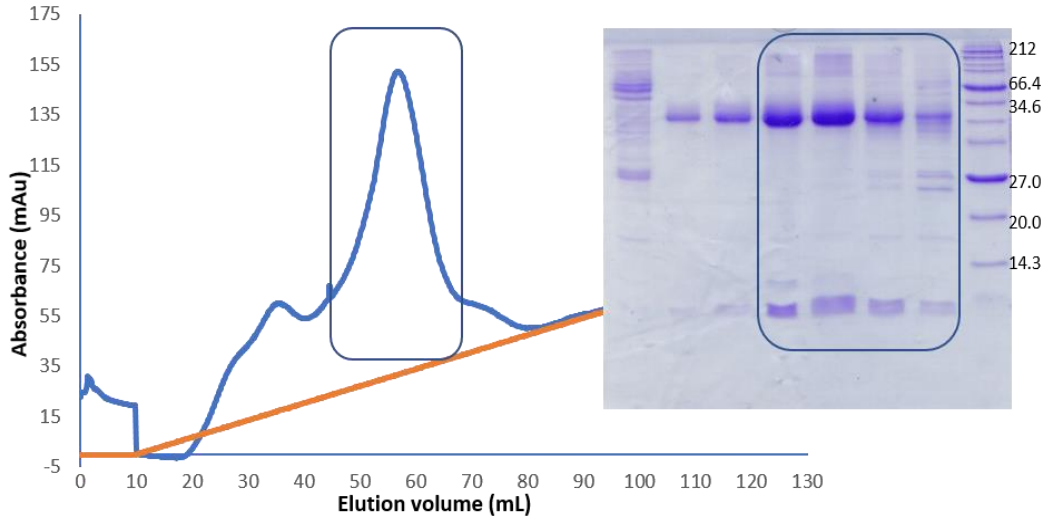


Figure 6.15. Nickel column chromatogram and SDS PAGE gel of active site mutant USP4htt-H106N with Ub-GGG

The purification of gel filtration (Figure 6.16) showed a single peak which elutes at 47.85 mL. This volume was void volume. The protein complex eluted as void volume and detected by SDS-PAGE analysis. It showed a clear band at the USP4htt-H106N size and a very low at the UbGGG band. A protein of UbGGG was not clear. The protein concentration was 1.47 mg/mL at 200 µL volume, and the total was 0.3 mg. The yield was 0.06 mg/L culture.

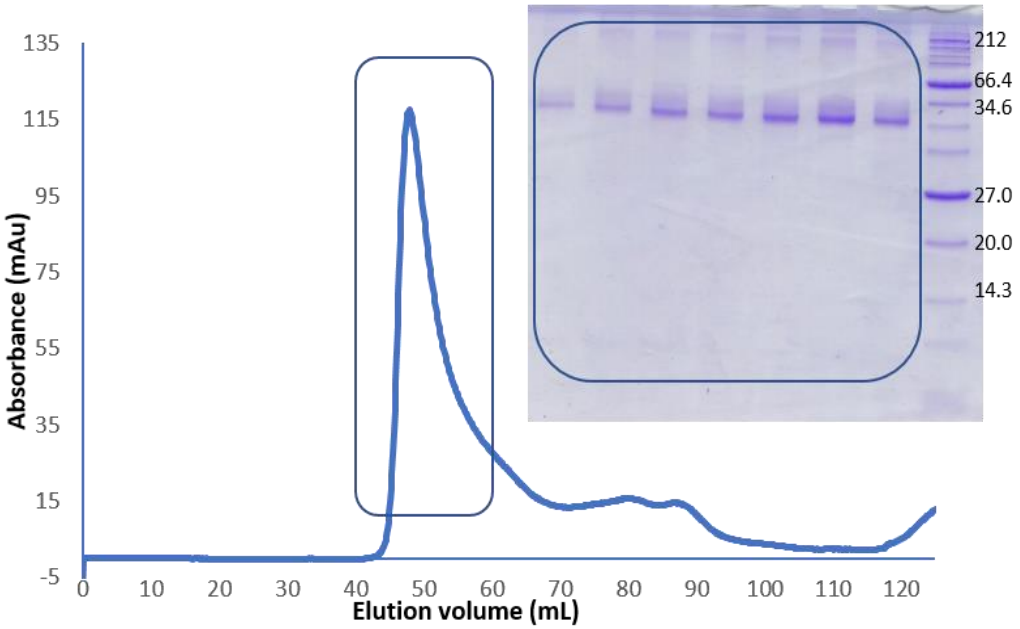


Figure 6.16. Gel filtration chromatography of USPhtt-H106N with Ub-GGG
Chromatogram using Superdex 200 HiLoad 16/60 showed a single peak at the void volume, around 47.85 mL, but on the SDS-PAGE gel, the result showed one band clear which is active site mutant of USP4htt (50.5 kDa) and an unclear band of UbGGG on the bottom gel (8.5 kDa). The yield of protein complex was 0.06 mg/L culture.

The active site mutant USP4htt-H106N also forms a protein complex with diubiquitin L73X. The purification was done by affinity chromatography (Ni column) and gel filtration. In the Ni column purification, the column did not work properly, so the peak of protein sample did not show on the chromatogram (Figure 6.17). By SDS-PAGE gel detection, the protein complex was expressed and showed on the gel. Upon gel filtration purification, the chromatogram showed an irregular pattern which indicates very low protein complex. The protein yield after gel filtration purification of the protein complex of an active-site mutant USP4htt-H106N-diUbL73X was low, only 0.007 mg/L culture. Some degradation might have happened which results in

the low concentration observed. This purification used a low salt concentration in the gel filtration buffer, 150 mM NaCl, 50 mM Tris-Cl pH 7.5 and 1 % Glycerol.

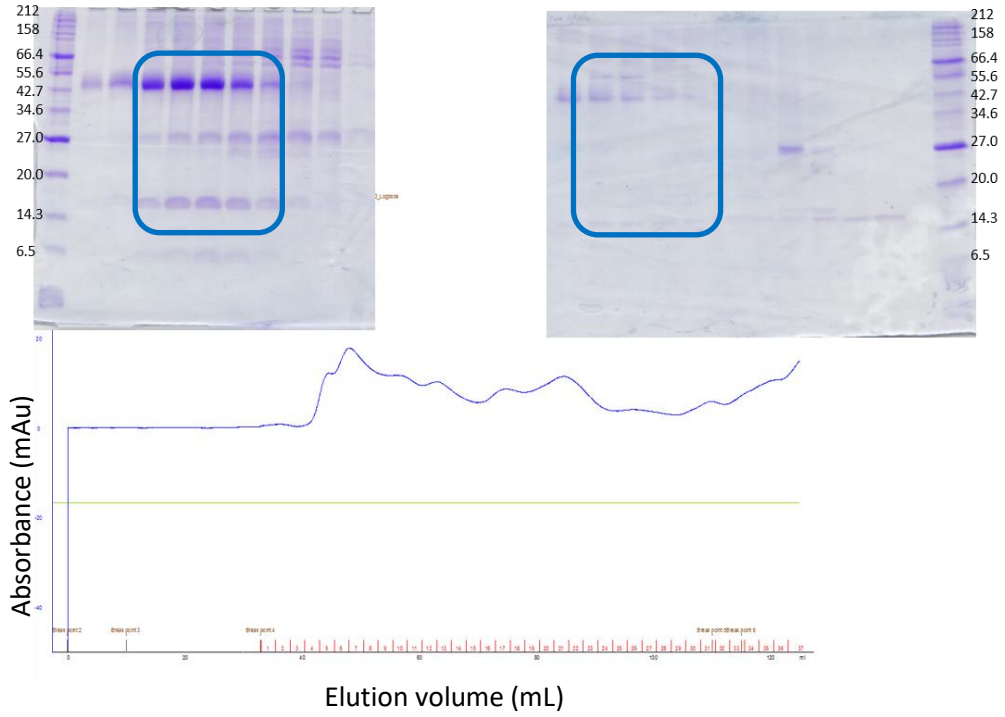


Figure 6.17. Chromatogram of Ni column and gel filtration of USP4htt-H106N with diUbL73X.
Chromatogram of Ni column cannot show because of air disruption. SDS-PAGE of Ni column result (left) showed bands that correspondent to active site mutant of USP4htt-H106N (50.5 kDa) and a band of diubiquitin-L73X on the bottom gel (17 kDa).
The chromatogram of size exclusion of USP4htt-H106N with diUbL73X using Superdex 200 HiLoad 16/60 showed a low peak which contains the protein target. SDS-PAGE gel result (right) showed bands that correspondent to active site mutant of USP4htt-H106N (50.5 kDa) and a band of diubiquitin-L73X on the bottom gel (17 kDa). The diubiquitin was less than Ni column result.

6.2.2. Expression and Purification of Active Site Mutant USP4htt-C311S with Ubiquitin Variants

Another mutation to develop protein complex in USP4htt construct was to create the C311S mutant. The protein complex with diubiquitin and UbGGG

performed. Both complexes, USP4httC311S with diubiquitin and USP4httC311S with UbGGG successfully expressed and purified.

Ni column purification of USP4httC311S with diubiquitin showed several peaks on the beginning elution of buffer B, but after 24% the single peak appeared which contain the protein complex. SDS-PAGE gel showed relatively pure protein of USP4httC311S and diubiquitin (Figure 6.18). The mutation of Cysteine to Serine on USP4htt construct to increase the purity of the protein.

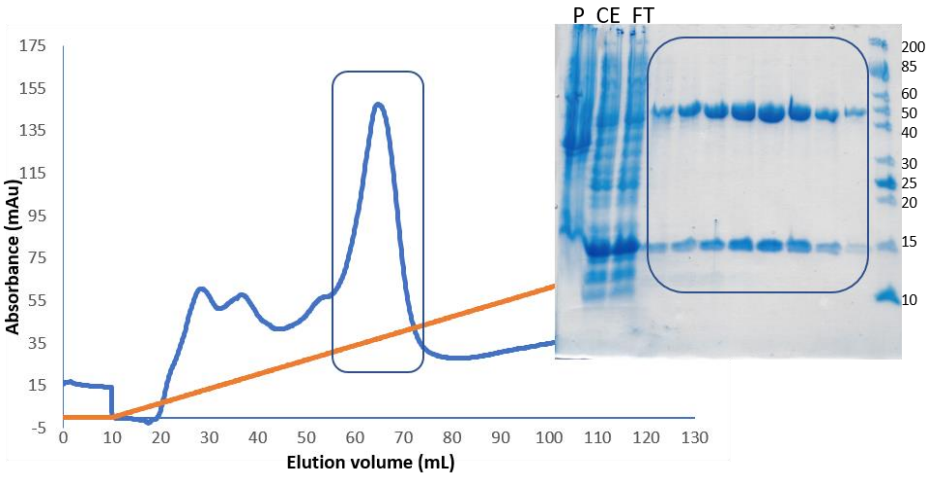


Figure 6.18. Ni column chromatogram from co-expression of USP4htt-C311S with diubiquitin
Chromatogram showed a low peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4htt-C311S (50.5 kDa) and a band of diubiquitin on the bottom gel (17 kDa).

Gel filtration (S200 HiLoad 16/60) purification of USP4httC311S with diubiquitin showed less void volume (Figure 6.19). Although there was a single low peak after elution of void volume, then a single high peak elutes the protein target at 75.52 mL. The elution volume correspondent to 77.5 kDa

based on the standard curve. This value indicates that the complex formation was monomer USP4htt-C311S and monomer diubiquitin.

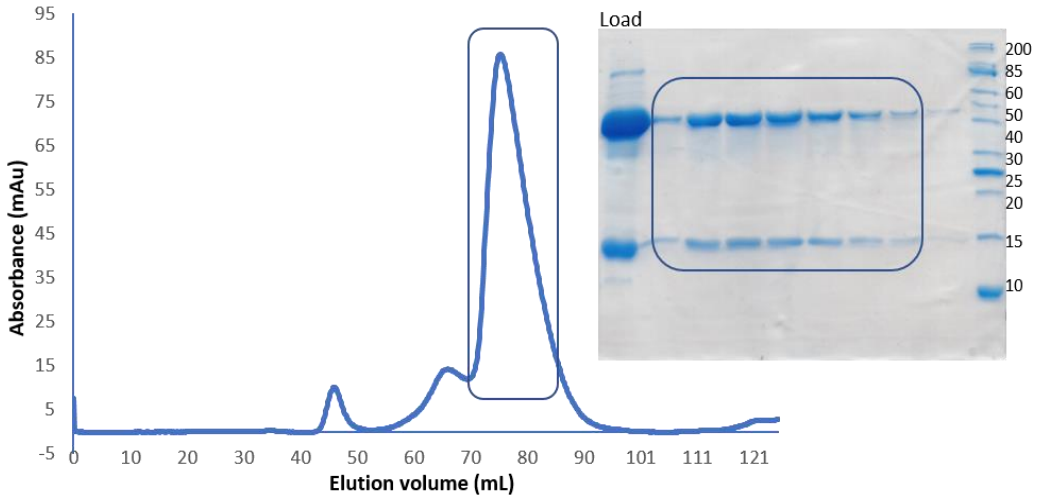


Figure 6.19. Chromatogram of gel filtration of USP4htt-C311S with diubiquitin
Chromatogram using Superdex 200 HiLoad 16/60 showed a single peak which contains the protein target elutes at 75.52 mL correlate with 77.5 kDa. Indicates the monomer USP4htt-C311S bind to monomer diubiquitin. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4htt-C311S (50.5 kDa) and a band of diubiquitin on the bottom gel (17 kDa).

The SDS PAGE gel confirms that the protein complex purified well and the band of USP4httC311S and diubiquitin were pure. Based on the performance of the SDS-PAGE gel, the mutation C311S showed clearer band. The final concentration of protein complex was 7.7 mg/mL on about 300 µL volume. The protein complex was 2.3 mg in total or 0.46 mg/L culture.

The complex of USP4httC311S with UbGGG also expressed well. The chromatogram showed a single peak after 23.5 % buffer B. SDS-PAGE gel indicates less degradation and the band was relatively pure for USP4httC311S, but for UbGGG was less expression, the band appears thin (Figure 6.20).

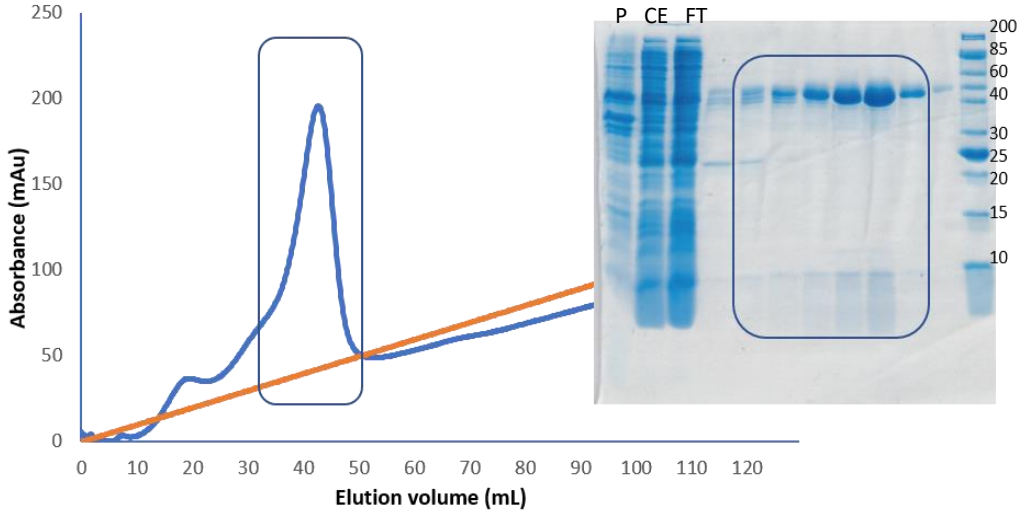


Figure 6.20. Chromatogram of Ni column of USP4htt-C311S with UbGGG
Chromatogram showed a low peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4htt-C311S (50.5 kDa) and a band of UbGGG on the bottom gel (8.5 kDa). (P=pellet; CE= crude extract; FT=flow trough)

On the gel filtration purification, the chromatogram showed the complex elutes at 73.6 mL. The volume correspondent with the molecular weight 83.8 kDa standard curve. This size showed that the protein complex USP4htt-C311S with Ub-GGG consists of monomer USP4htt-C311S and tetramer Ub-GGG.

The pattern on the SDS-PAGE gel was equal with Ni column purification, well expressed for USP4httC311S but less expression for UbGGG (Figure 6.21). The final concentration of the protein complex was 5.29 mg/mL at 300 µL volume or 1.6 mg. The yield was 0.3 mg/L culture. This protein complex then continues to Crystallisation trial.

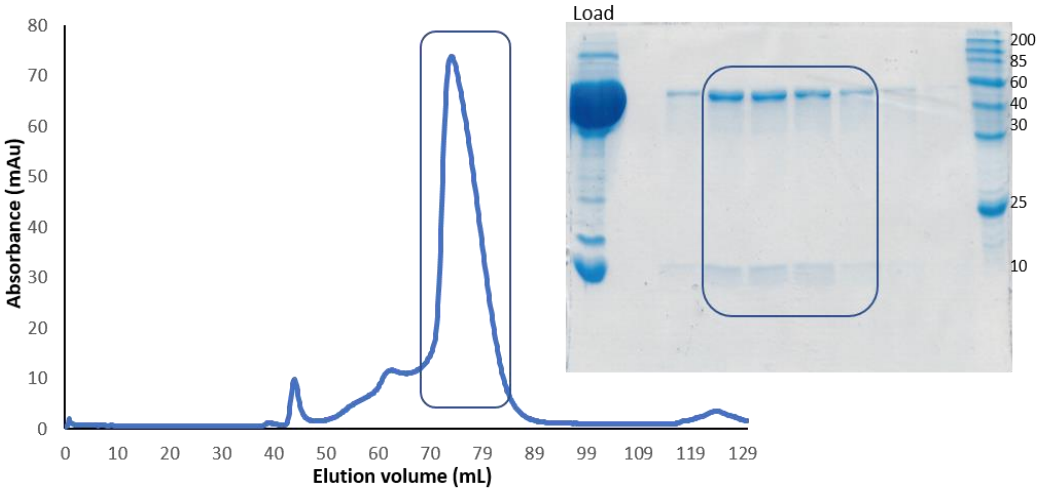


Figure 6.21. Chromatogram of gel filtration of USP4htt-C311S with UbGGG
Chromatogram using Superdex 200 HiLoad 16/60 showed a single peak which contains the protein target. The elution volume was 73.6 mL which refer to 83.8 kDa of the standard curve. This size correspondent with monomer USP4htt-C311S with a tetramer of Ub-GGG. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4htt-C311S (50.5 kDa) and a band of UbGGG on the bottom gel (8.5 kDa).

The construct of USP4htt showed an increase of yield in the C311S mutation rather than H106N (Table 6.2). There was an increase yield ~2-fold in the complex with diubiquitin. The H106N mutation in complex with Ub-GGG and diubiquitin-L73X produce a void volume. In the solution, mutation change from H106N to C311S in USP4htt made the dimer formation to the monomer. The monomer of USP4htt-C311S in complex possibly more stable refers to the yield was obtained.

Table 6.2. Comparison of yield, Elution volume and molecular weight prediction of protein complex USP4htt-H881N and C311S with Ubiquitin variant

Constructs	Yield (mg/L culture)	Elution volume (mL)	Mass Prediction (kDa)	Actual size (kDa)	Interpretation
USP4htt-H106N with diubiquitin	0.2	67.04	158.6	50.5 + 17	Dimer USP4htt-H106N bind with trimer diubiquitin
USP4htt-H106N with UbGGG	0.06	47.85 [Void volume]	-	50.5 + 8.5	-
USP4htt-H106N with diubiquitin-L73X	0.007	Irregular peak	-	50.5 + 17	-
USP4htt-C311S with diubiquitin	0.46	75.52	77.5	50.5 + 17	Monomer USP4htt C311S + monomer diubiquitin
USP4htt-C311S-with UbGGG	0.3	73.6	83.8	50.5 + 8.5	Monomer USP4htt- C311S with tetramer UbGGG

6.3. Protein expression and purification of USP4httΔUbl2-H106N active site mutant

The active site mutant of USP4httΔUbl2-H106N was co-expressed and purified with diubiquitin. The method for purification also uses the Nickel column and gel filtration purification. The affinity chromatography of USP4httΔUbl2-H106N with diubiquitin showed the target protein elutes after 24.5% of buffer B. SDS-PAGE gel showed both of active site mutant protein and the ligand were on the gel on the correct position refer to the marker protein (Figure 6.22).

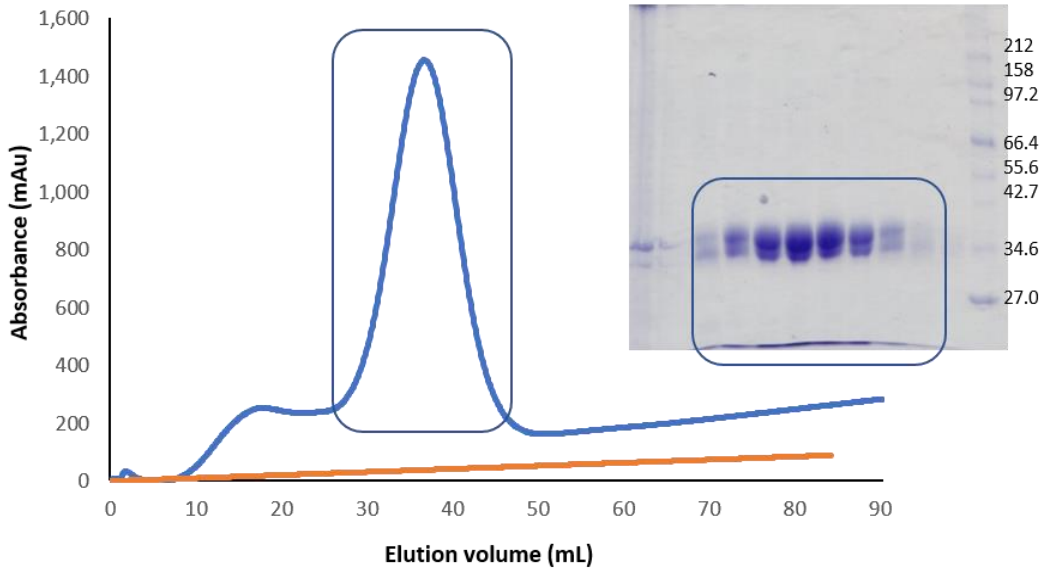


Figure 6.22. Ni column protein purification of USP4httΔUbl2-H106N with linear diubiquitin Chromatogram from Ni column purification with a single high peak showed the protein product. SDS-PAGE gel from Ni column of the product USP4 httΔUbl2-H106N at 40.7 kDa. Some degradation product of USP4httΔUbl2-H106N seen. Diubiquitin is shown at the bottom (17 kDa).

The gel filtration purification showed almost have no void volume and a single peak at elution volume 78.01 mL was indicated to a protein complex of monomer USP4httΔUbl2 and monomer diubiquitin because of correspondent to 62.8 kDa on the standard curve (Figure 6.23). This result was equal to SDS PAGE gel data which showed a less degradation of the protein complex, USP4htt ΔUbl2 was on the top of the gel, with some smear at above 34.6 kDa, and diubiquitin was in the bottom at above 14.3 kDa. The final concentration of the complex was 4.76 mg/mL at about 400 µL. The total protein was 1.9 mg or approximately 0.4 mg/L culture (Table 6.3). This protein complex continues for crystallisation trial.

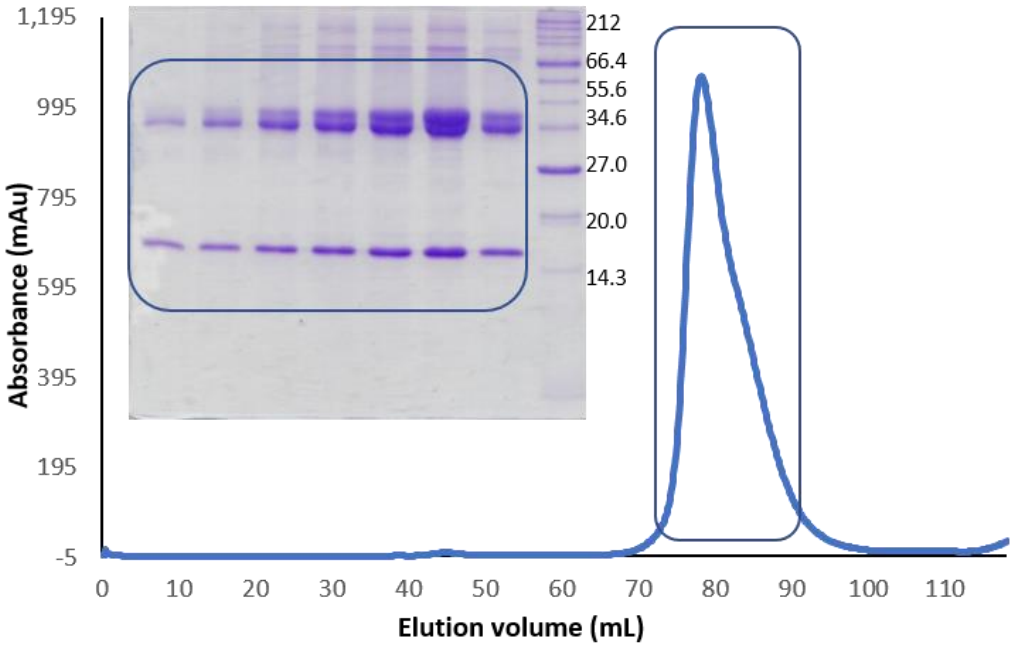


Figure 6.23. Gel filtration purification of USP4httΔUbl2-H106N with diubiquitin
A chromatogram from gel filtration in Superdex 200 HiLoad 16/60 with a single high peak at 78.01 mL elution volume. This volume equal with 62.8 kDa of the standard curve, which indicates the monomer of a USP4httΔUbl2 bind with monomer diubiquitin to form a complex. SDS-PAGE gel showed the protein was on 40.74 kDa with diubiquitin on 17 kDa

Table 6.3. Yield, Elution volume and molecular weight prediction of protein complex USP4httΔUbl2-H106N with diubiquitin

Constructs	Yield (mg/L culture)	Elution volume (mL)	Mass Prediction (kDa)	Actual size (kDa)	Interpretation
USP4httΔUbl2-H106N with diubiquitin	0.4	78.01	62.8	40.7	Monomer USP4httΔUbl2-H106N bind with monomer diubiquitin

6.4. Protein expression and purification active site mutant of USP4C1C2 H881N and C311S

The active site mutant of USP4C1C2 was mutated both on Histidine residue (H881N) and cysteine residue (C311S). The active site mutant of USP4C1C2 co-expressed with ubiquitin variants, such as linear diubiquitin and UbGGG.

6.4.1. Protein expression and purification active site mutant of USP4C1C2-H881N with Ubiquitin Variant

The expression method of complex USP4C1C2-H881N with diubiquitin was copurification, and the complex with UbGGG was co-expression. All the complex of USP4C1C2-C311S with diubiquitin and Ub-GGG were co-expression.

The first protein complex was USP4C1C2-H881N with diubiquitin. The result of nickel column purification (Figure 6.24) and gel filtration purification (Figure 6.25) USP4C1C2-H881N with diubiquitin showed the protein product and the diubiquitin bound. Ni column purification showed that the complex elutes at 27.5 buffer B. SDS PAGE gel confirmed that the complex expressed with the

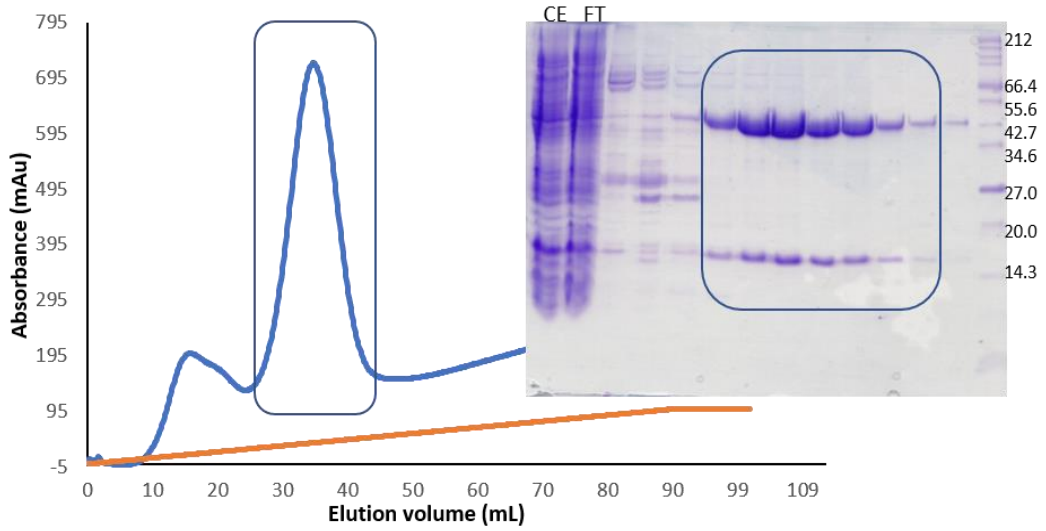


Figure 6.24. Ni column of USP4C1C2-H881N with linear diubiquitin.
SDS-PAGE gel of USP4C1C2-H881N showed on 46 kDa with linear diubiquitin on 17 kDa. The graph showed that complex USP4C1C2-H881N with linear diubiquitin elutes at 27.5% buffer B. (CE=crude extract; FT=flow trough).

clean band. The USP4C1C2-H881N was on the top at around 42.7 kDa, and the diubiquitin was on the bottom of gel about 14.3 kDa of a protein marker.

The gel filtration purification has less void volume and only one single peak which elutes at 74.01 mL (Figure 6.25). The elution volume equal with 88.05 kDa of the standard curve. The molecular weight corresponds to the monomer of USP4C1C2-H881N with dimer diubiquitin. This peak contains protein complex as shown at SDS PAGE gel. The band on the gel has the same pattern with Ni column purification, with more clear and thick band with no degradation product. The USP4C1C2-H881N with linear diubiquitin were purified well. The concentration of the complex was 8.2 mg/mL at 400 µL. The yield of this complex was 0.6 mg/L culture. This complex continued to crystallisation trial.

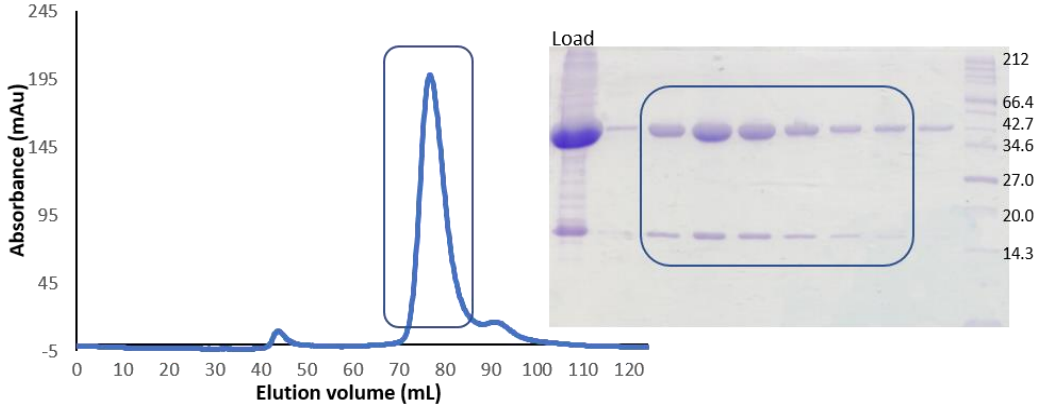


Figure 6.25. Chromatogram gel filtration of USP4C1C2-H881N with linear diubiquitin
SDS PAGE result of USP4C1C2-H881N showed on 46 kDa with linear diubiquitin on 17 kDa. The graph using SuperDex 200 HiLoad 16/60 showed that complex USP4C1C2-H881N with linear diubiquitin elutes at 74.01 mL which equal to 88.05 kDa of the standard curve. This data showed that monomer USP4C1C2-H881N bind with dimer diubiquitin.

Active site mutant of USP4C1C2-H881N continues to combine with Ub-GGG to form a protein complex. USP4C1C2-H881N and Ub-GGG coexpressed in *E. coli* BL21 strain. The purification method used affinity chromatography and gel filtration Superdex 200 HiLoad 16/60.

The Nickel column purification of USP4C1C2-H881N with Ub-GGG showed that the protein complex eluted after 22.5% of buffer B (Figure 6.26). SDS PAGE gel explained that protein complex expressed with some degradation above at the USP4C1C2-H881N band. The Ubiquitin band was relatively pure at around 6.5 kDa of a protein marker.

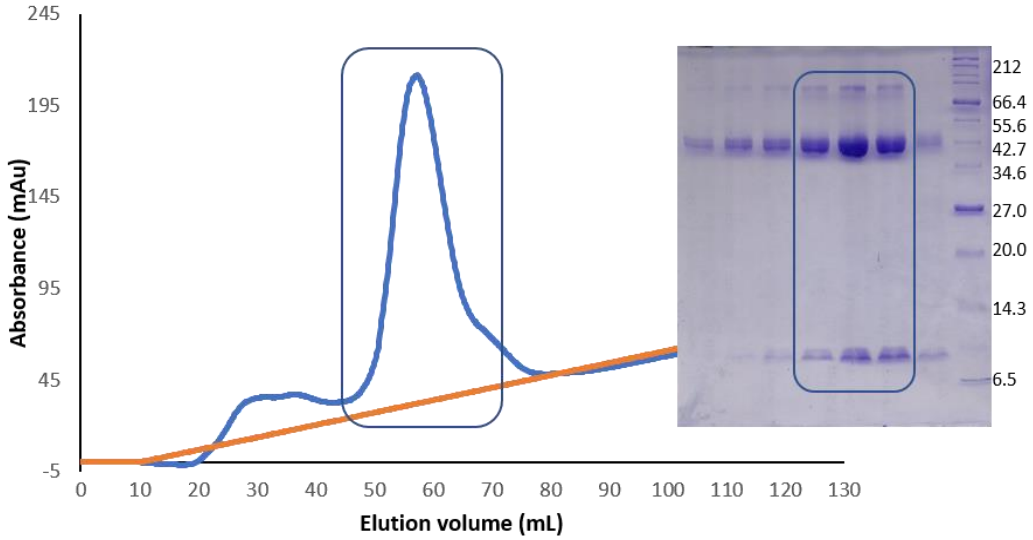


Figure 6.26. Affinity chromatography of USP4C1C2-H881N with Ub-GGG
Chromatogram showed single peak correspondent to the protein product. SDS-PAGE gel result showed two bands which are: UbGGG and active site mutant of USP4C1C2.

Gel filtration purification also indicates that protein complex has purified. The chromatogram showed no void volume and the protein complex was on the single peak eluting at 86.23 mL (Figure 6.27). Based on the standard curve, the

elution volume equal with 31.4 kDa of molecular weight, indicates a very compact of protein complex USP4C1C2-H881N bind with Ub-GGG.

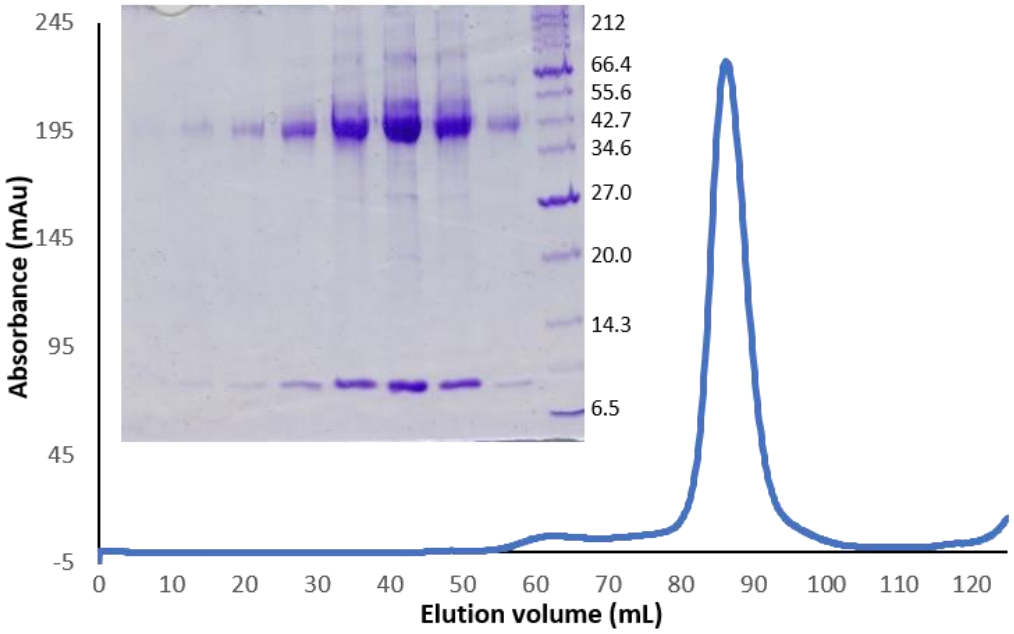


Figure 6.27. Gel filtration chromatography of USP4C1C2-H881N with Ub-GGG.
Chromatogram using Superdex 200 HiLoad 16/60 showed a single peak, and the elution volume was 86.23 mL equal with 31.4 kDa on the standard curve. SDS-PAGE gel result showed two bands which are: UbGGG and active site mutant of USP4C1C2-H881N. The total protein was 19.1 mg from 5 L cultures (3.8 mg/L culture)

SDS PAGE gel of USP4C1C2-H881N with Ub-GGG showed the same pattern with Nickel column gel with some degradation protein near the USP4C1C2-H881N band and pure band at the UbGGG band. The protein complex after concentrated was 38.3 mg/mL at 300 µL. The total protein was 11.49 mg. The yield was 2.3 mg/L culture. This protein complex continues to Crystallisation trial.

Another protein complex of USP4C1C2-H881N combined with diubiquitin-L73X. The purification of an active site mutant USP4C1C2-H881N with

diubiquitin-L73X done by Nickel column and gel filtration Superdex 200 HiLoad 16/60. Nickel column purification showed a single peak after 31% of buffer B (Figure 6.28). The peak was containing protein complex as shown on the SDS PAGE gel. The USP4C1C2-H881N was a thick band with some degradation protein on the top of the gel. Also, the diubiquitin-L73X was at the bottom refer to 14.3 kDa marker protein.

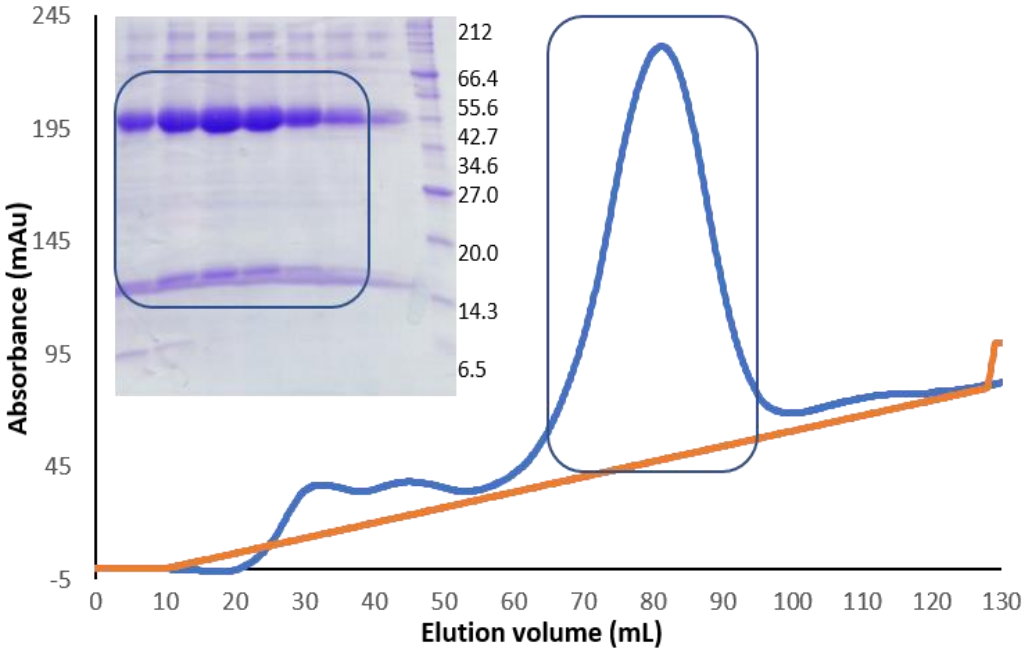


Figure 6.28. Chromatogram of Ni column of USP4C1C2-H881N-diUbL73X
Chromatogram showed a big high peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-H881N (46.4 kDa) and a band of diubiquitin-L73X on the bottom gel (17 KDa).

The contaminant on the top of the SDS PAGE gel which closes to the USP4C1C2-H881N band decreased after gel filtration purification. Also, the band for ubiquitin was clearer. The elution volume of the protein complex was 81.1 mL (Figure 6.29). The elution volume refers to 48.4 kDa compared to the standard curve. The mass of protein complex showed less than actual size

indicates a tight and compact binding and folding between USP4C1C2-H881N with diubiquitin-L73X.

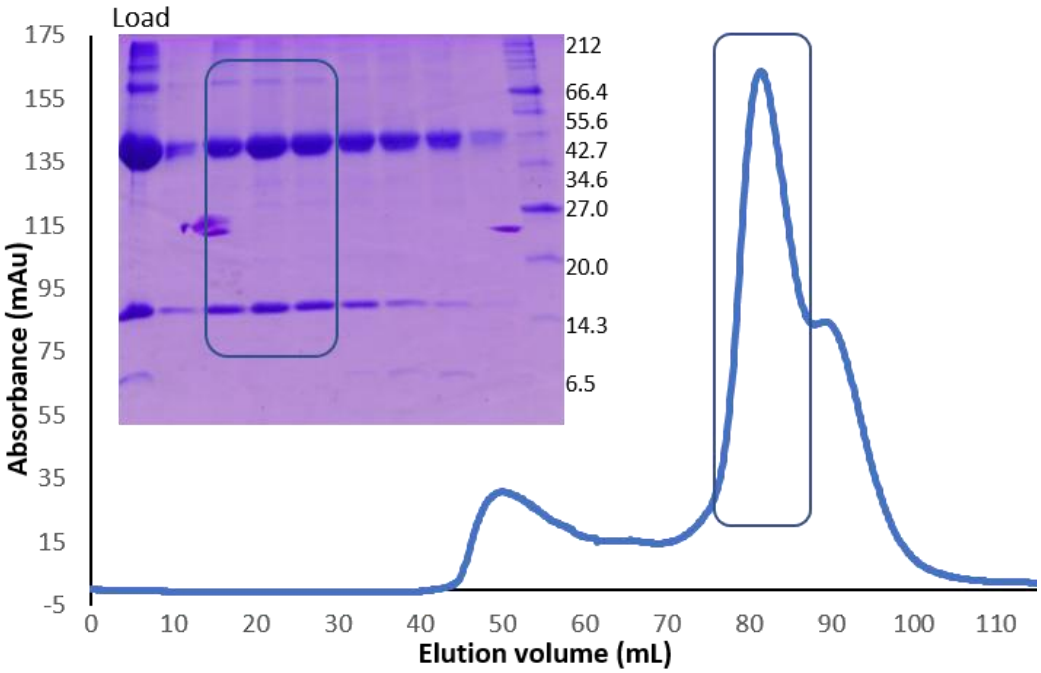


Figure 6.29. Gel filtration of active site mutant USP4C1C2-H881N-diUbL73X Chromatogram using Superdex 200 HiLoad 16/60. showed a big high peak which contains the protein target. The elution volume was 81.1 mL was correspondent to 48.4 kDa of the standard curve. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-H881N (46.4 kDa) and a band of diubiquitin-L73X on the bottom gel (17 kDa). The protein concentration was 23.9 mg/mL in about 500 µL volume (Yield was 11.9 mg).

This protein complex purification showed relatively pure protein and high concentration, such as 23.9 mg/mL after gel filtration in a volume 500 µL. The total protein was 11.9 mg from 5 L cultures, and the yield was 2.38 mg/L culture. The complex then continues for Crystallisation trial. For further analysis by mass spectrometry, the protein complex was dialysis using Ammonium acetate 150 mM pH 6.8 for overnight at 4°C.

6.4.2. Protein expression and purification of active site mutant USP4C1C2-C311S with Ubiquitin Variant

The mutation in the active site USP4C1C2 continues to mutate Cysteine to Serine in the catalytic 1. The active site mutant of USP4C1C2-C311S was co-expressed with diubiquitin and UbGGG for Crystallisation trial. The purification of those construct using Ni column and gel filtration method.

Expression and purification of complex USP4C1C2-C311S with diubiquitin has increased the purities of the protein complex. Based on the Ni column chromatogram, the complex eluted after 25.5 % buffer B. The gel showed a clear band both of USP4-C1C2-C311S and the diubiquitin, on the SDS PAGE gel (Figure 6.30).

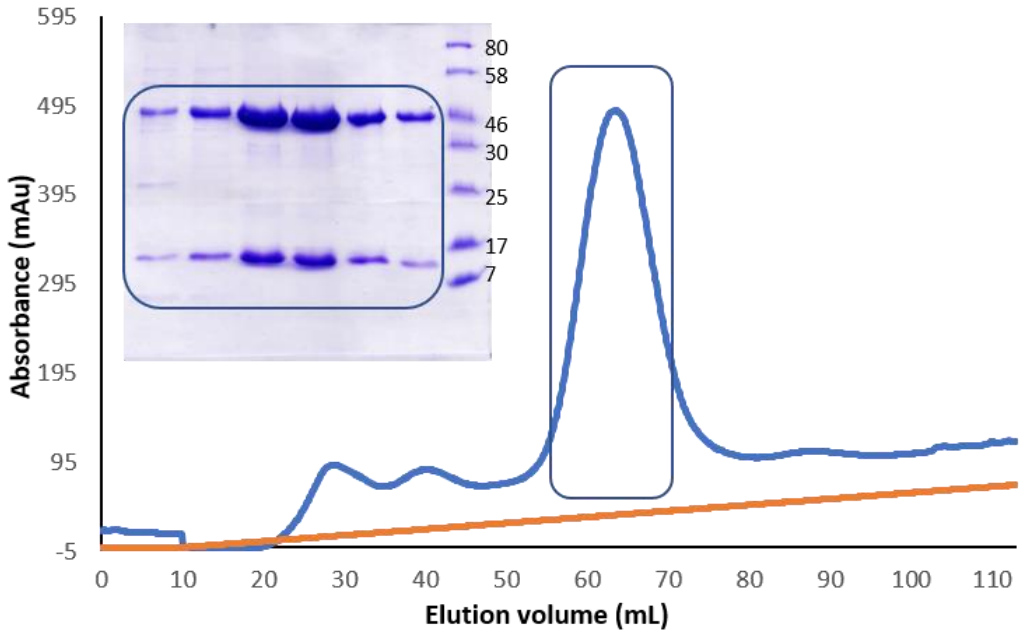


Figure 6.30. Chromatogram of Ni column of USP4C1C2-C311S with diubiquitin Chromatogram showed a single peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S (46.4 kDa) and a band of diubiquitin on the bottom gel (17 kDa).

The gel filtration purification of complex USP4C1C2-C311S with diubiquitin showed the complex eluted at 77.6 mL, equal with 59.8 kDa on the standard curve. This size indicates a protein complex of monomer USP4C1C2-C311S bind with monomer diubiquitin. SDS PAGE gel showed the thick band on both constructs, the USP4C1C2-CS about 46 kDa and diubiquitin at around 17 kDa (Figure 6.31). It concluded that the complex of USP4C1C2-C311S and diubiquitin purified well. The concentration of the protein complex was 18.4 mg/mL in 500 µL. The yield of this complex was 1.84 mg/L culture. The sample continued for Crystallisation trial.

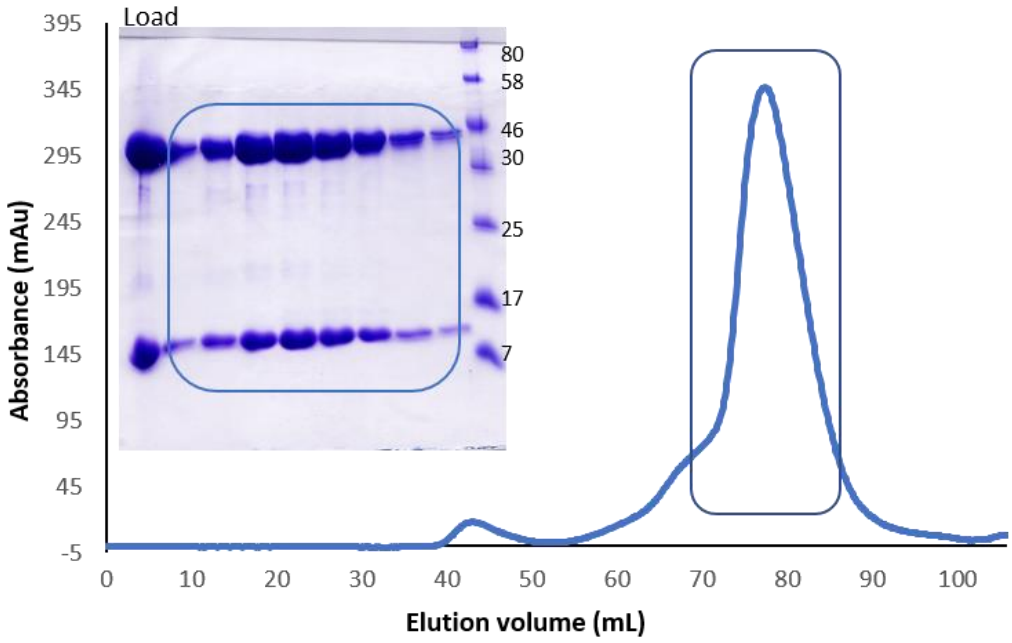


Figure 6.31. Chromatogram of gel filtration of USP4C1C2-C311S with diubiquitin. Chromatogram using Superdex 200 HiLoad 16/60 showed a single peak which contains the protein target in the elution volume 77.6 mL. The volume refers to 59.8 kDa of the standard curve which equals to monomer USP4C1C2-C311S and monomer diubiquitin. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S (46.4 kDa) and a band of diubiquitin on the bottom gel (17 kDa).

The second protein complex of USP4C1C2-C311S combined with UbGGG. The complex also purified with Ni column and gel filtration method. The protein complex was express well, and the purification using Nickle column showed that the protein complex elutes at 25% buffer B. SDS-PAGE gel showed that the USP4C1C2-C311S has more quantity of the protein and Ub-GGG showed less quantity (Figure 6.32). Compare to the complex with diubiquitin; the Ub-GGG has a thin band. Both on the same lane with the correct molecular weight refer to the protein marker.

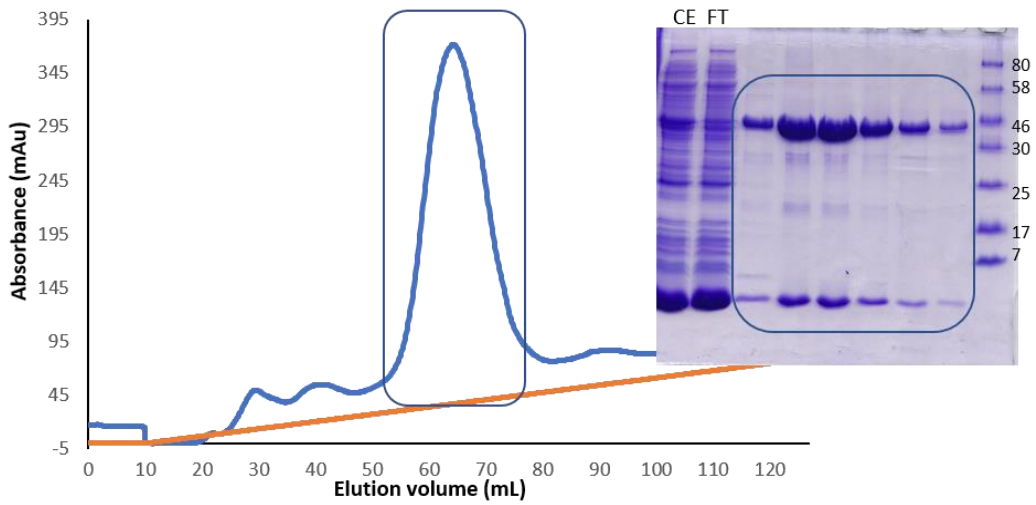


Figure 6.32. Chromatogram of Ni column USP4C1C2-C311S with UbGGG
Chromatogram showed a peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S (46.4 kDa) and a band of UbGGG on the bottom gel (8.5 kDa). (CE=crude extract; FT=flow trough).

The result of gel filtration chromatogram showed no void volume and the protein complex elutes at 80.21 mL. The elution volume equal with 52.2 kDa on the standard curve, indicates the complex of USP4C1C2-C311S with UbGGG consists of both monomer bind each other. SDS-PAGE gel showed

increasing the quantity of UbGGG on the protein complex (Figure 6.33). The concentration of the complex was 9.8 mg/mL in 500 µL volume. The yield was 0.98 mg/L culture. This complex continues for Crystallisation trial. The complex USP4C1C2-C311S with diubiquitin has higher yield compared to complex with UbGGG.

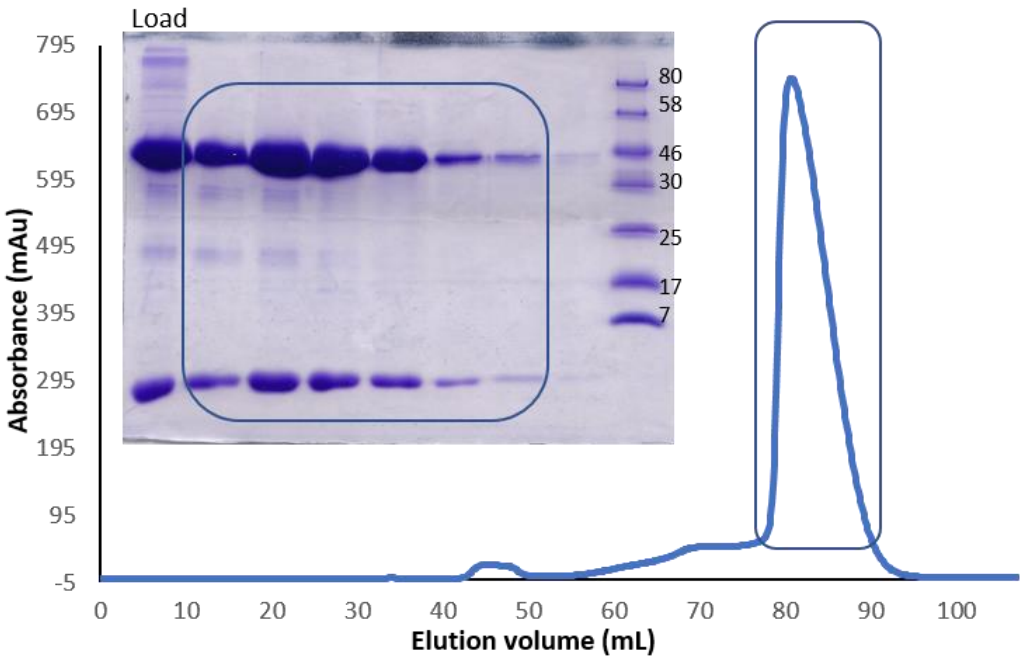


Figure 6.33. Chromatogram of gel filtration of USP4C1C2-C311S with UbGGG. Chromatogram with Superdex 200 HiLoad 16/60showed a peak which contains the protein target at the elution volume 80.21 mL. Based on the standard curve equal with 52.2 kDa. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S (46.4 kDa) and a band of UbGGG on the bottom gel (8.5 kDa).

The construct of USP4C1C2-C311S has mutated again on the Y928 to stop codon (Y928X). The protein complex obtained with co-expressed and purified with diubiquitin and UbGGG. The purification of protein complex USP4C1C2-C311S-Y928X with diubiquitin using Ni column showed at the 24 % buffer B the protein eluted on the single peak. Both proteins expressed well and on the

correct molecular weight, refer to the protein marker, but the band of diubiquitin showed decreasing quantity. The band of USP4C1C2-C311S-Y928X seems thicker than diubiquitin as shown on the SDS PAGE gel (Figure 6.34).

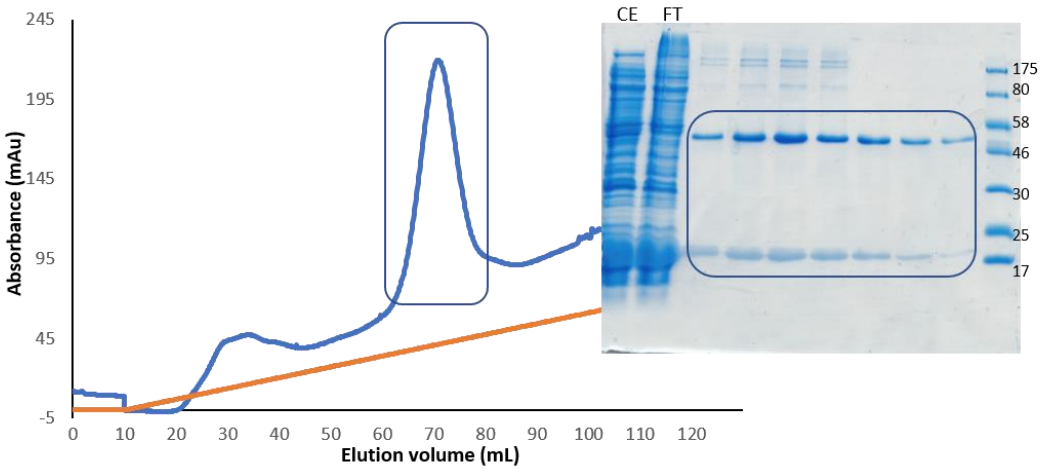


Figure 6.34. Chromatogram of Ni column of USP4C1C2-C311S-Y928X with diubiquitin Chromatogram showed a peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S-Y928X (46.4 kDa) and a band of diubiquitin on the bottom gel (17 kDa). (CE=crude extract; FT=flow trough).

The gel filtration of complex USP4C1C2-C311S-Y928X with diubiqutin showed an elution volume at 85.64 mL. A single asymmetric peak showed that the protein complex was purified. The elution volume, compared to the standard curve, equal with 33 kDa. This molecular weight indicates that the protein complex of USP4C1C2-C311S-Y928X increase the binding affinity with the diubiquitin and folding compactly.

SDS PAGE gel of USP4C1C2-C311S-Y928X with diubiqutin indicates no degradation protein and the protein complex only consisted USP4C1C2-C311S-Y928X and diubiquitin (Figure 6.35). Because the peak was not an equal distribution, the total protein separated into two, the left side was 24.4 mg/mL,

and the right side was 33.05 mg/mL at 500 µL volume. The yield was 2.44 mg/L culture and 3.3 mg/L culture. Both protein concentrate continues for Crystallisation trial.

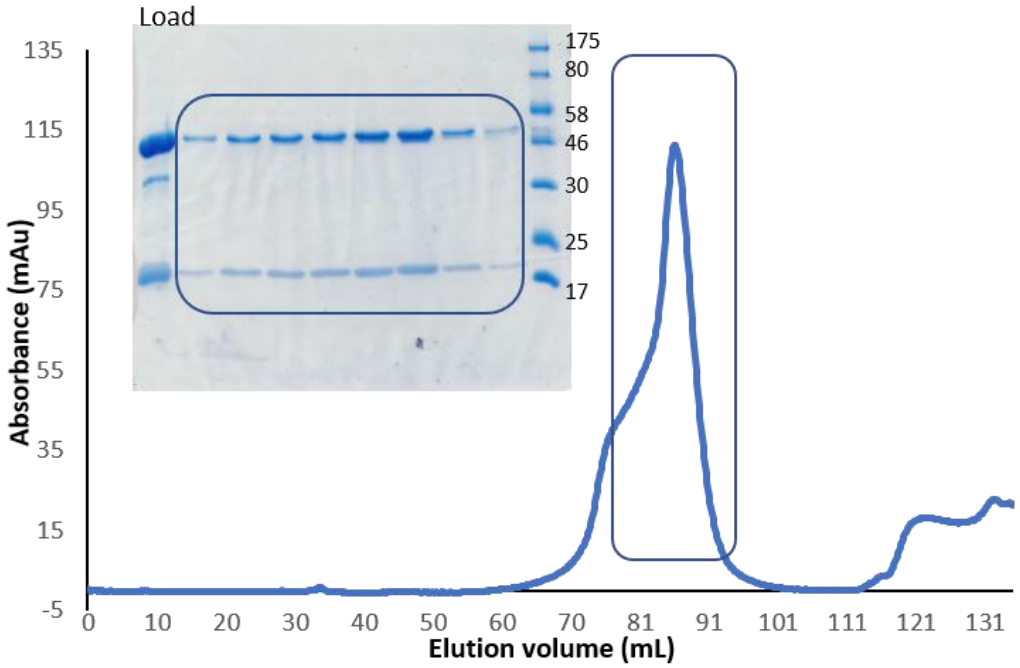


Figure 6.35. Chromatogram of gel filtration of USP4C1C2-C311S-Y928X with diubiquitin Chromatogram using superdex 200 HiLoad 16/60 showed a low peak which contains the protein target at the elution volume 85.64 mL. The elution volume refers to 33 kDa on the standard curve. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S-Y928X (46.4 kDa) and a band of diubiquitin on the bottom gel (8.5 kDa).

Another complex of USP4C1C2-C311S-Y928X combined with UbGGG. Purification using Ni column have an air problem, so the chromatogram cannot show. SDS PAGE gels confirmed the availability of protein complex. The gel showed that USP4C1C2-C311S-Y928X expressed, but for Ub-GGG seems very low band.

Using gel filtration purification Superdex 200 HiLoad 16/60, the complex of USP4C1C2-C311S and UbGGG has elution volume at 89.58 mL. The elution

volume refers to 23.6 kDa on the standard curve, less almost a half than the combination of both actual sizes. The elution volume indicates that complex of USP4C1C2-C311S-Y928X with UbGGG were consists of monomer from each protein and bound very tight each other to form a compact folding structure. SDS PAGE gel showed less quantity of UbGGG indicates by a thin band, but for the USP4C1C2-C311S-Y928X still showed the thick band with the high band (Figure 6.36). The concentration of the protein complex was 5.87 mg/mL in 500 µ volume. The yield was 2.93 mg/L culture.

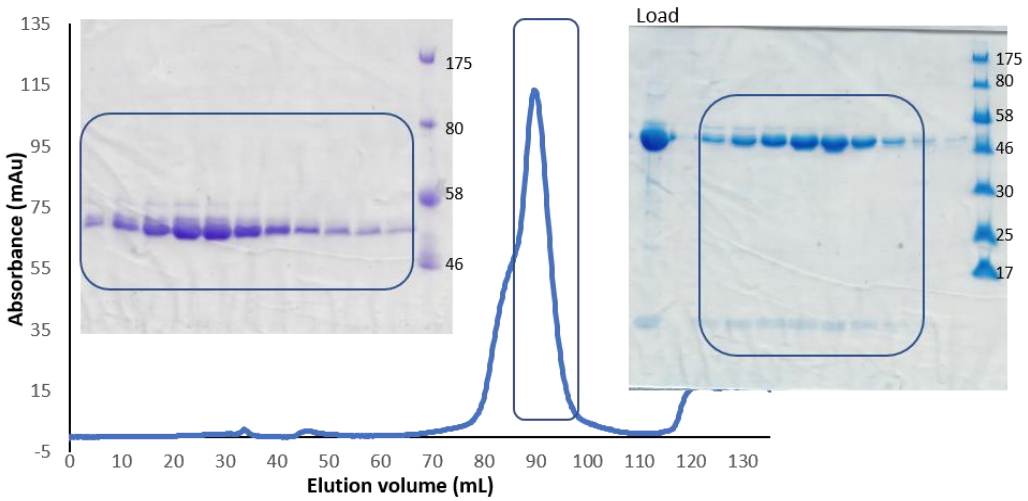


Figure 6.36. Chromatogram of gel filtration of USP4C1C2-C311S-Y928X with UbGGG
Chromatogram of Superdex 200 HiLoad 16/60 showed a peak which contains the protein target at the elution volume 89.58 mL. Based on the standard curve equal with 23.6 kDa. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S-Y928X (46.2 kDa) and a band of UbGGG on the bottom gel (8.5 kDa).
SDS-PAGE on the left was Ni column purification without chromatogram.

The result indicates the yield of both complexes has the same pattern with the C311S mutation only. The protein complex mutant USP4C1C2-C311S-Y928X with diubiquitin has a higher yield, 24.4 (33.05) mg/mL compared to the

UbGGG complex was only 5.87 mg/mL. This complex continues for crystallisation trial.

The last purification for USP4C1C2 was active site mutant USP4C1C2-C311S only. This purification prepared for ITC assay with Ubiquitin variant. Ni column purification showed a single peak which contains protein target (Figure 6.37). It was indicating from the single band on the SDS PAGE gel which correlates molecular weight with protein target, 46.4 kDa, at the marker protein around 40 kDa. A little shadow of the band appears below the protein target USP4C1C2-C311S, probably too many samples when loading to the gel.

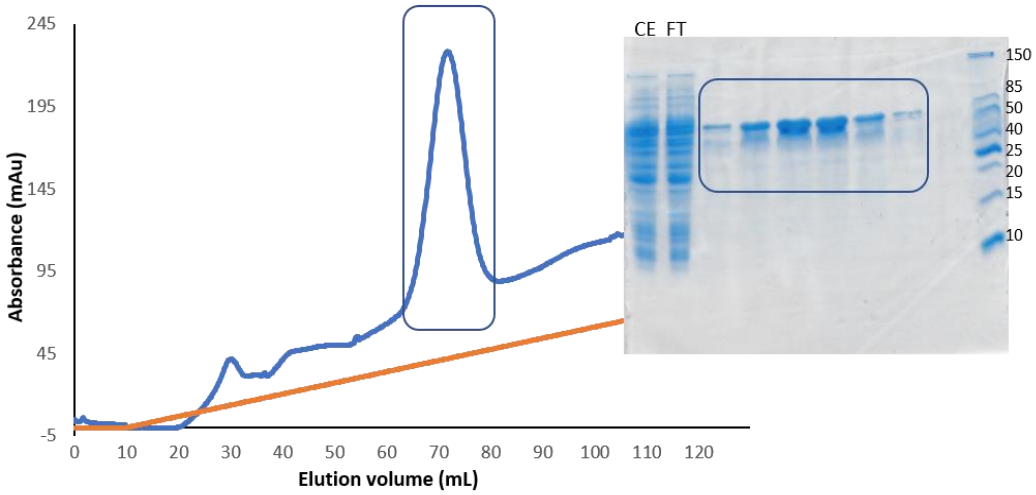


Figure 6.37. Chromatogram of Ni column USP4C1C2-C311S
Chromatogram showed a peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S (46.4 kDa)

The gel filtration superdex 75 HiLoad 16/60 result showed a pure band of an active-site mutant USP4C1C2-C311S (Figure 6.38). The protein elutes at volume 51.77 mL. This volume correlates with 117.8 kDa on the standard

curve. The molecular weight indicates the protein elutes as a dimer of USP4-C311S. The purification gave the yield 0.7 mg/L culture. The sample USP4C1C2-C311S continue for ITC experiment.

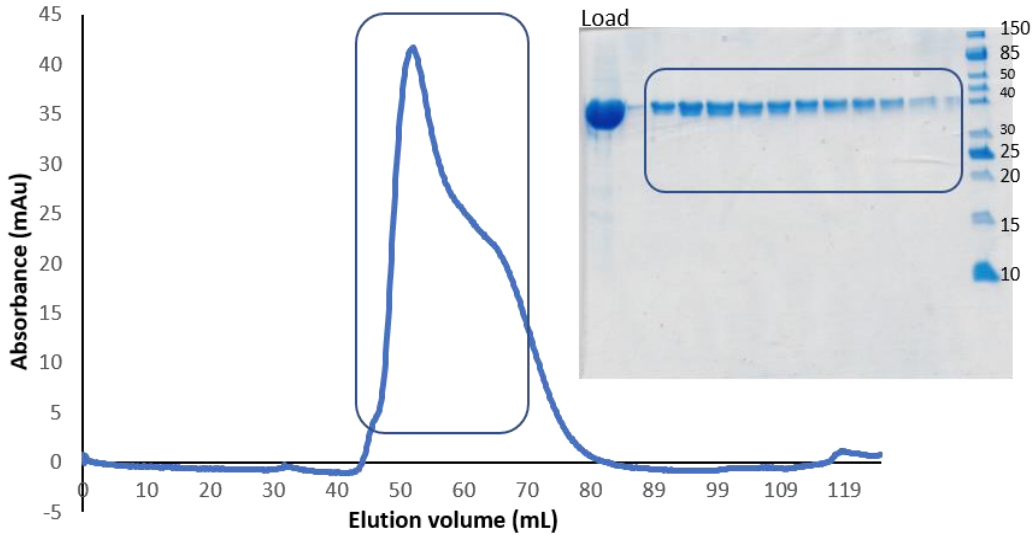


Figure 6.38. Chromatogram of gel filtration of USP4C1C2-C311S
Chromatogram of superdex 75 HiLoad 16/60showed a peak which contains the protein target at the elution volume 51.77 mL. Based on the standard curve, equal with 117.8 kDa. This size indicates a dimer formation of USP4C1C2-C311S. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4C1C2-C311S (46.4 kDa)

The construct of USP4C1C2-H881N showed the highest yield in complex with Ub-GGG and diubiquitin-L73X. There was a decrease ~4-fold in diubiquitin complex. Interestingly, in the USP4C1C2-C311S, there was a reverse comparison, the yield was increased ~2-fold in complex with diubiquitin compare to Ub-GGG. The mutation C311S and Y928X in complex with diubiquitin and Ub-GGG showed an equal yield. The formation of active site mutant USP4C1C2 in complex with Ubiquitin variant was always monomer in every different mutation (Table 6.4).

Table 6.4. Comparison of yield, Elution volume and molecular weight prediction of protein complex USP4C1C2-H881N, C311S and C311S with Ubiquitin variant

Constructs	Yield (mg/L culture)	Elution volume (mL)	Mass Prediction (kDa)	Actual size (kDa)	Interpretation
USP4C1C2-H881N with diubiquitin	0.6	74.01	88.05	46.4 + 17	Monomer USP4C1C2-H881N bind with dimer diubiquitin
USP4C1C2-H881N with UbGGG	2.3	86.23	31.4	46.4 + 8.5	Monomer USP4C1C2-H881N bind with monomer UbGGG
USP4C1C2-H881N with diubiquitin-L73X	2.38	81.1	48.4	46.4 + 17	Monomer USP4C1C2- H881N bind with monomer diubiquitin-L73X
USP4C1C2-C311S with diubiquitin	1.84	77.6	59.8	46.4 + 17	Monomer USP4C1C2- C311S- bind with monomer diubiquitin
USP4C1C2-C311S with UbGGG	0.98	80.21	52.2	46.4 + 8.5	Monomer USP4C1C2- C311S bind with monomer Ub-GGG
USP4C1C2-C311S-Y928X with diubiquitin	2.44 (3.3)	85.64	33	46.4 + 17	Monomer USP4C1C2- C311S-Y928X bind with monomer diubiquitin
USP4C1C2-C311S-Y928X with UbGGG	2.93	89.58	23.6	46.4 + 8.5	Monomer USP4C1C2- C311S-Y928X bind with monomer Ub-GGG
USP4C1C2-C311S	0.7	51.77	117.8	46.4	Dimer USP4C1C2-C311S

6.5. Protein expression and purification active site mutant of USP4ΔDU H881N and C311S

The active site mutant of USP4ΔDU was mutated with Histidine to Asparagine (H881N) and Cysteine to Serine (C311S). On the mutation H881N, the complex formation of active site mutant USP4ΔDU was only with ubiquitin. For mutation C311S, active site mutant USP4ΔDU was purified without complex formation and prepare for ITC experiment.

6.5.1. Protein expression and purification of active site mutant USP4ΔDU-H881N with Ubiquitin variant

The complex formation USP4ΔDU-H881N was co-expressed with ubiquitin and purified using Ni column and gel filtration method. The chromatogram of Nickle column showed a single peak which starts to elute at 17.5% buffer B. SDS-PAGE gel showed some protein degradation (Figure 6.39).

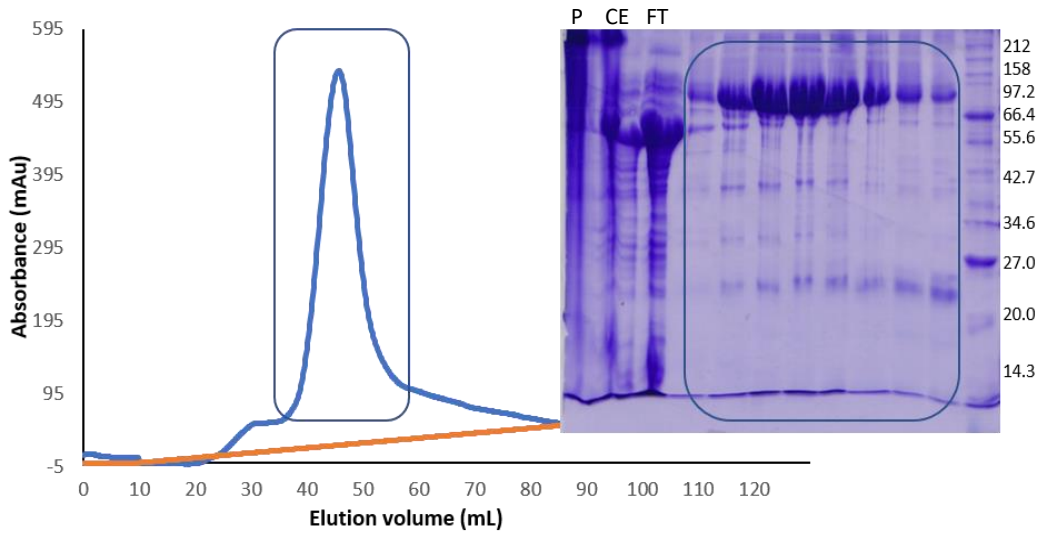


Figure 6.39. Chromatogram of Ni column of USP4ΔDU-H881N with Ubiquitin
Chromatogram showed a peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4ΔDU-H881N (79.5 kDa) and ubiquitin (8.5 kDa). (P=Pellet; CE=crude extract; FT=flow trough).

A thick band showed the USP4ΔDU-H881N protein on the top around 97.2 kDa protein marker, and ubiquitin was on the bottom of the gel. The degradation product after Ni column seems decreases after gel filtration purification (Figure 6.40).

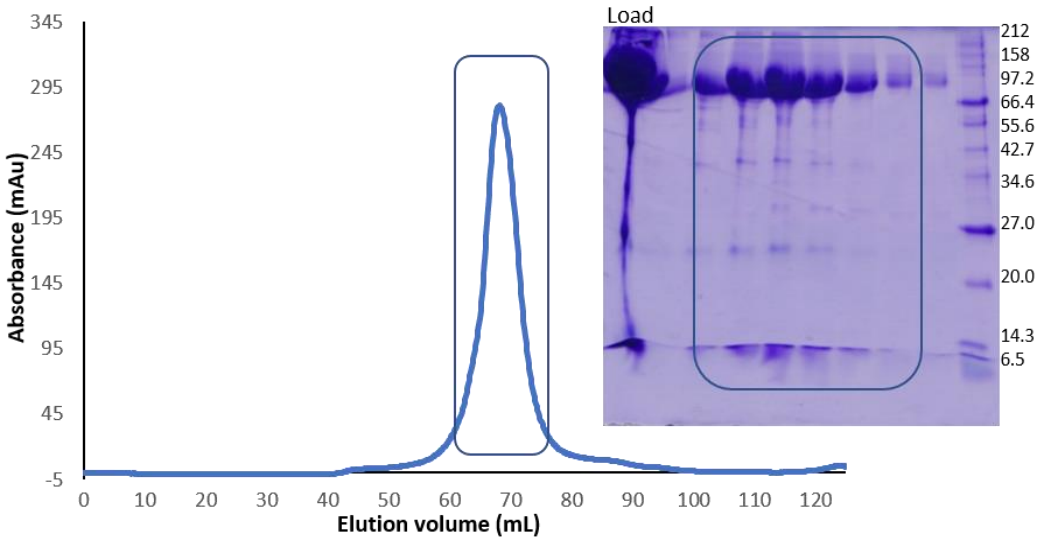


Figure 6.40. Chromatogram of gel filtration of USP4ΔDU-H881N with Ubiquitin
Chromatogram of superdex 200 HiLoad 16/60 showed a peak which contains the protein target elutes at 68.24 mL equal with 143 kDa based on the standard curve. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4ΔDU-H881N (79.5 kDa) and ubiquitin (8.5 kDa)

The gel filtration purification showed one peak which elutes at 68.24 mL. The elution volume refers to 143 kDa of the standard curve. The molecular weight was more than the actual size of the protein complex, indicates the complex probably consists of monomer USP4ΔDU-C311S bind with hexamer Ubiquitin. A polyubiquitin chain in the complex formation.

Less degradation protein showed on the SDS PAGE gel with the clearer band for USP4ΔDU-H881N and ubiquitin. The total protein was 49 mg/mL in 200 µL

final volume. The yield was 1.96 mg/L culture. This protein complex was continued with Crystallisation trial and mass spectrometry analysis.

6.5.2. Protein expression and purification of active site mutant USP4ΔDU-C311S

The other mutation for USP4ΔDU was C311S, and there was no complex formation with this mutation. The purification of active site mutant USP4ΔDU-C311S only was prepared for ITC assay. The method was Ni column and gel filtration purification.

Ni column purification result of USP4 ΔDU-C311S indicates some degradation (Figure 6.41). The chromatograph showed that there was another peak before the main peak eluate. The USP4ΔDU-C311S eluted from 24% buffer B. SDS PAGE gel showed some degradation protein in the protein target which was 79.5 kDa.

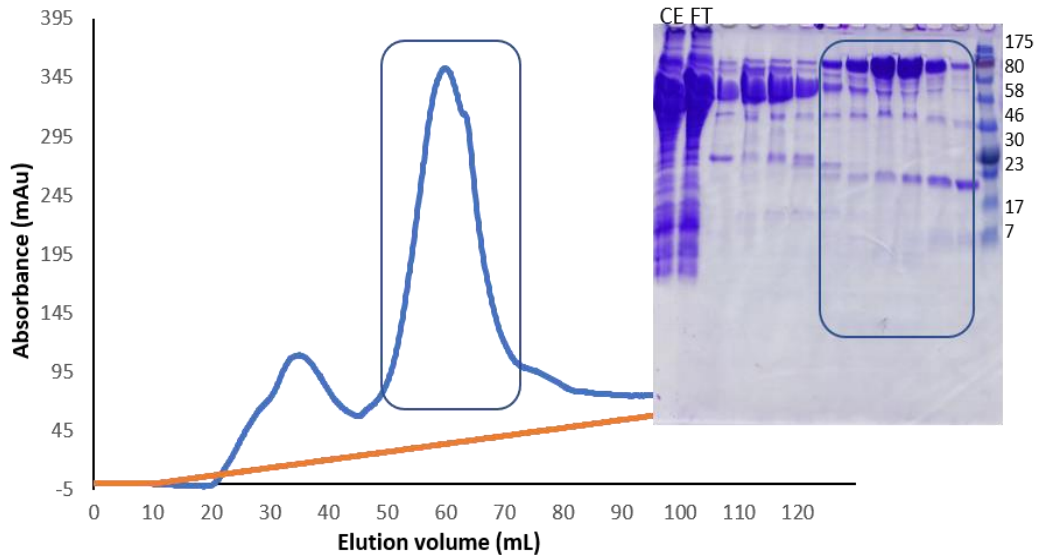


Figure 6.41. Chromatogram of Ni column USP4ΔDU-C311S
Chromatogram showed a peak which contains the protein target. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4ΔDU-C311S (79.5 kDa). (CE=crude extract; FT=flow trough).

Gel filtration purification Superdex200 HiLoad 16/60 can decrease the contaminant because appear on the chromatogram only one single peak which elutes at 65.59 mL (Figure 6.42). Based on the standard curve, the volume refers to mass 179 kDa. This size was a dimer formation for USP4ΔDU.

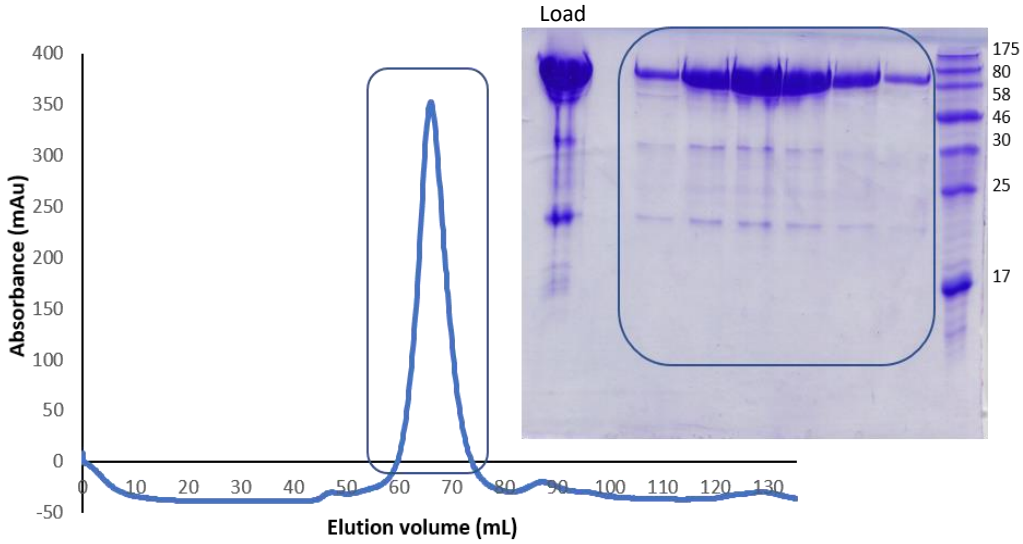


Figure 6.42. Chromatogram of gel filtration of USP4ΔDU-C311S
Chromatogram of S200 HiLoad 16/60 showed a peak elutes at 65.59 mL which contains the protein target. Based on the standard curve, the elution volume refers to 179 kDa, indicates the USP4ΔDU was in the dimer formation. SDS-PAGE gel result showed bands that correspondent to active site mutant of USP4ΔDU-C311S (79.5 kDa)

SDS PAGE gel, there was USP4ΔDU band with high intensity and less contamination. The band at around 80 kDa marker protein which closes to actual size 79.5 kDa. The final concentration of the protein was 8.7 mg/mL in 1 mL volume protein, so the total protein was 8.7 mg from 5 L cultures. The yield was 1.74 mg/L culture.

There yield between complex USP4ΔDU-H881N with Ubiquitin and USP4ΔDU-C311S relatively the same (Table 6.5). More purification sample of mutant

USP4ΔDU was needed to gain the proportional comparison. The formation in solution is also difficult to compare due to the limited of the sample.

Table 6.5. Comparison of yield, elution volume and molecular weight prediction of protein complex USP4ΔDU-H881N and C311S with Ubiquitin

Constructs	Yield (mg/L cultures)	Elution volume (mL)	Mass Prediction (kDa)	Actual size (kDa)	Interpretation
USP4ΔDU-H881N with ubiquitin	1.96	68.24	143	79.5 + 8.5	Monomer USP4ΔDU H881N bind with hexamer ubiquitin
USP4ΔDU-C311S	1.74	65.59	179	79.5	Dimer USP4ΔDU-C311S

6.6. Crystallisation Trials

Crystallisation was used as the method to try and solve protein structures of USP4. Crystallisation needed a pure and relatively large amount of protein sample and crystallisations were set up after purification. To increase the possibility of getting crystals, a complex formation of USP4 active site mutants with ligand ubiquitin variants were performed.

The protein complexes of USP4 active site mutants (USP4FL-H881N, USP4FL-C311S-Y928X, USP4htt-H106N, USP4htt-C311S, USP4C1C2-H881N, USP4C1C2-C311S, USP4C1C2-C311S-Y928X, USP4httΔUbl2-H106N, USP4httΔUbl2-C311S, and USP4ΔDU-H881N) with Ubiquitin variants: ubiquitin, linear diubiquitin, UbGGG, and diubiquitin-L73X were set up for crystallisation.

6.6.1. Crystallisation trials of protein complexes of USP4FL

The complex of USP4FL-H881N with linear diubiquitin resulted in crystal formation. The crystal from a complex of USP4FL-H881N with linear diubiquitin was obtained after three months incubation at 10˚C. Crystallisation conditions were obtained from Morpheus, PACT, and JCSG plus screens (Figure 6.43).

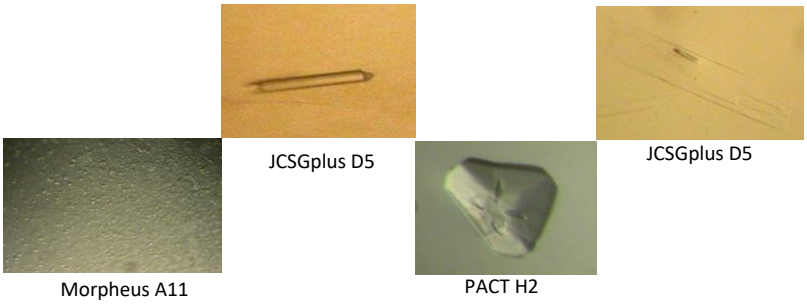


Figure 6.43. Crystals with various solutions from a sample of complex USP4FL-H881N with linear diubiquitin.

- Morpheus A11 consists of 0.06 M divalents (0.3 M Magnesium chloride hexahydrate; 0.3 M Calcium chloride dehydrate), 0.1 M Tris (base) and BICINE pH 8.5, and 50% precipitant (25% MPD, 25% PEG 1000, 25% PEG 3350)
- JCSG plus D5 consists of 0.1 M HEPES pH 7.5 and 70% MPD
- PACT H2 consists of 0.2 M Sodium bromide, 0.1 M Bis-Tris propane pH 8.5 20% PEG 3350

Some of the crystals were sent to synchrotron to test the X-ray diffraction. Sample crystals from Morpheus solutions F5 and F9 diffracted at a 2.7 Å, and 3 Å and datasets were collected. Unfortunately, the crystals figure of F5 and F9 were not available. The data collection statistics are shown in Table 6.6. Morpheus F5 conditions consist of 0.12 monosaccharides, 0.1 M buffer system 2 pH 7.5 (Sodium HEPES and MOPS) and 50% precipitant mix (40% PEG 500 and 20% PEG 20000). Morpheus F9 consists of 0.12 monosaccharides, 0.1 M buffer system 3 pH 8.5 (Tris Base and BICINE) and 50% precipitant mix (40%

PEG 500 and 20% PEG 20000). An image of the crystal diffraction of F5 is shown in Figure 6.44.

Table 6.6. Cell parameter of sample protein complex USP4FL-H881N with diubiquitin

Cell Parameter			
Space group	P 1 2_1 1		
a, b, c	87.34Å	135.28 Å	157.37 Å
α, β, γ	90	102.69	90
Resolution	2.99 Å		
Rmerge	0.133 (0.96)*		
Rpim	0.082 (0.589)		
$I / \sigma I$	13.0 (2.3)		
CC 1/2	0.995 (0.799)		
Completeness (%)	99.7 (98.5)		
Redundancy	7.0 (7.0)		

*value in parentheses are for highest-resolution shell

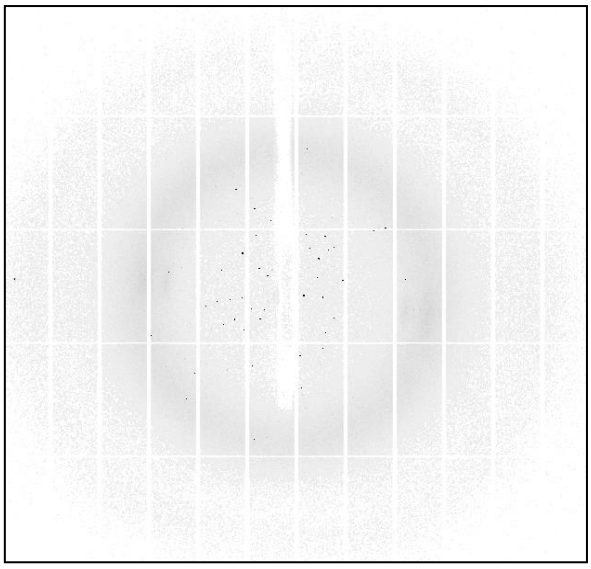
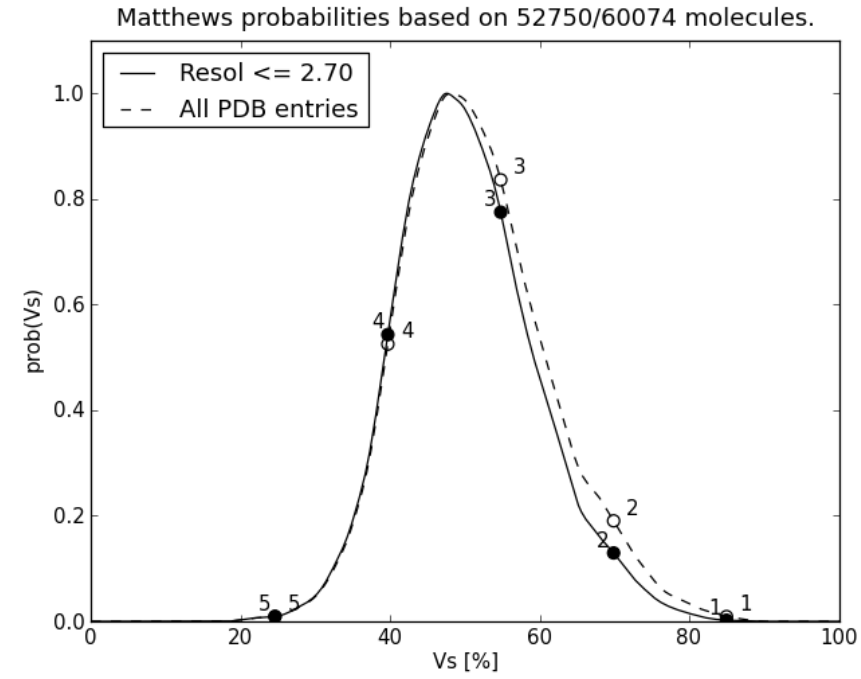
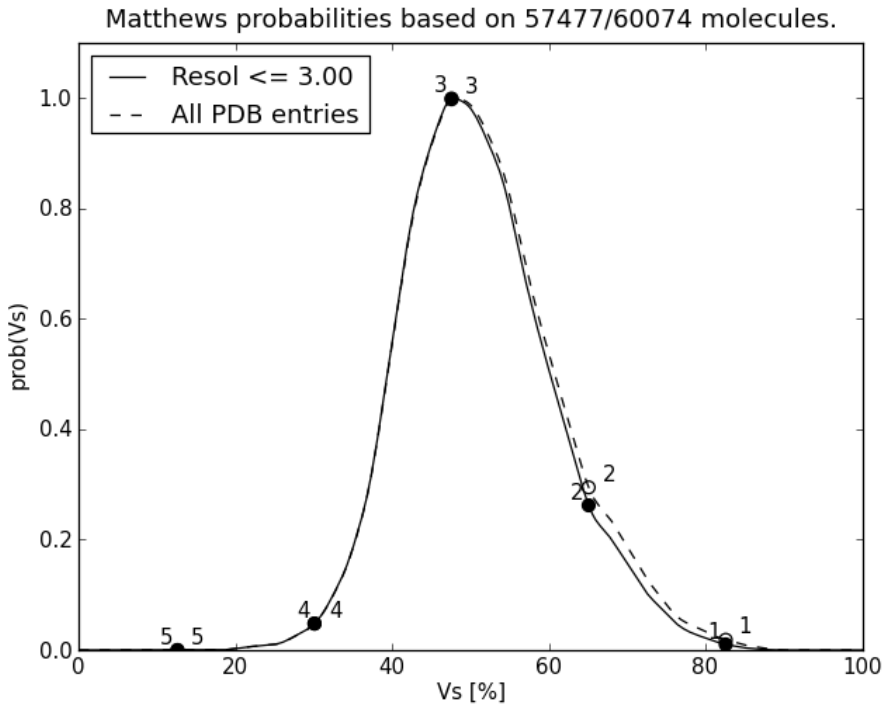


Figure 6.44. Crystal diffraction of sample F5-USP4FL-H881N with diubiquitin in resolution 3 Å

Prediction of the crystal USP4FL-H881N was calculated by Matthew Probability Calculator (Kantardjieff and Rupp, 2003). The calculations predict the crystal USP4FL-H881N on its own and complex USP4FL-H881N with diubiquitin. Figure 6.45 and Table 6.7 showed the prediction of the crystal USP4FL has a



A



B

Figure 6.45. Matthews coefficient of crystal USP4FL-H881N and USP4FL-H881N-diubiquitin. The crystal USP4FL was predicted has 3 or 4 in the asymmetric unit (A), and in complex with diubiquitin, the crystal USP4FL was predicted has 3 or 2 in the asymmetric unit (B).

higher probability (0.5374 and 0.3483) with 3 or 4 molecules in the asymmetric unit and the complex USP4FL-H881N with diubiquitin has a high probability (0.1987 and 0.7571) with 2 or 3 molecules.

Table 6.7. Calculation of Matthews Probabilities and Solvent content of USP4FL-H881N and USP4FL-H881N with diubiquitin

N (mol)	Prob (N) For resolution	Prob (N) overall	Vm A^3/Da	Vs % solvent	Mw Da
USP4FL-H881N					
1	0.0027	0.0058	8.16	84.92	111121.00
2	0.1055	0.1213	4.08	69.84	222242.00
3	0.5374	0.5328	2.72	54.76	333363.00
4	0.3483	0.3342	2.04	39.68	444484.00
5	0.0060	0.0058	1.63	24.60	555605.00
USP4FL-H881N-diubiquitin					
1	0.0078	0.0137	7.04	82.52	128784.00
2	0.1987	0.2168	3.52	65.05	257568.00
3	0.7571	0.7340	2.35	47.57	386352.00
4	0.0364	0.0354	1.76	30.09	515136.00
5	0.0000	0.0000	1.41	12.62	643920.00

In order to prove that the crystal contained USP4, the crystallisation drop was further analysed. The dissolved crystal and crystallisation solutions (the drops that did not contain the crystal used for data collection but contained same protein solution) were also run on the SDS-PAGE gel. SDS-PAGE gel result showed some degradation protein and needed to be identified. Three bands from the result (66, 55 and 43 kDa (green arrow)) were sent for mass spectrometry analyses to investigate the protein's identity (Figure 6.46).

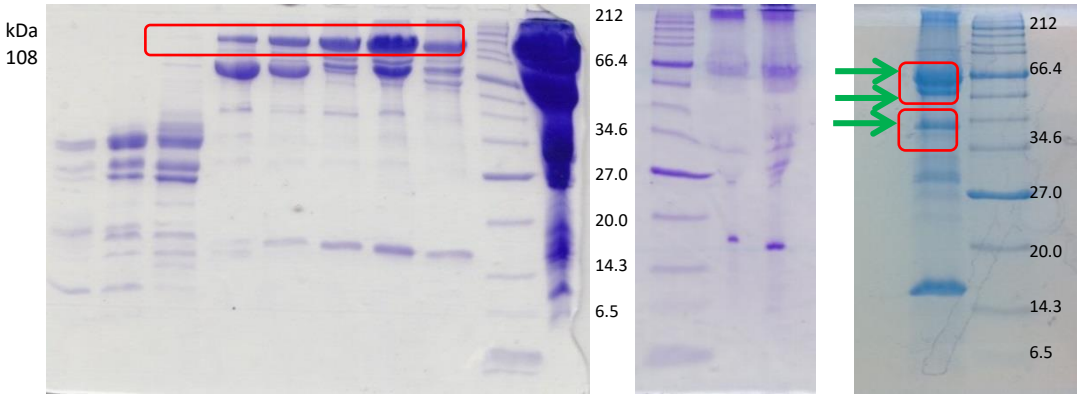


Figure 6.46. The comparison of the SDS-PAGE gel of USP4FL-H881N with linear diubiquitin. From gel filtration purification result (A), from crystal formation (Morpheus) (B), and protein precipitant in the drop (Morpheus) (C). The red circle was a sample that sent to mass spec for protein identification.

The alignment data from mass spectrometry result and USP4 construct showed that the three samples (66, 55 and 43 kDa), contain a part of USP4, such as DUSP-Ubl1, Ubl1 and insert. For example, from the band of 66 kDa, the mass spec results are LYYDEQESEAYEK (part of the insert), LLNWYGCVEGQQPIVR (part of DUSP) and LDNTVQDAGLYQGQVLVIEPQ NEDGTWPR (part of Ubl1).

The results from 55 kDa are VEVYLLELK (part of Ubl1), YMSNTYEQLSK (part of DUSP-Ubl1), LYYDEQESEAYEK (part of Ubl1), LLNWYGCVEGQQPIVR (part of the insert) and LDNTVQDAGLYQGQVLVIEPQNED GTW PR (part of Ubl1). Moreover, the results from 43 kDa are YMSNTYEQLSK (Ubl1), LYYDEQESEAYEK (insert) and ERPDAETQKSELGPLMR (part of DUSP). The other main component identified in mass spec were Chain A, Glucosamine 6-phosphate synthase with Glucose 6-phosphate, alcohol dehydrogenase, and chaperonin GroEL. All of them are *Escherichia coli* proteins, so the identity of the protein in the crystal

still needs to be resolved. Glucosamine 6-phosphate synthase is a common crystallisation artefact (Niedziakolwska et al., 2015).

The structure of DUSP-Ubl (PDB code: 3JYU, the catalytic core domain: PDB code 2Y6E (Clerici et al., 2014), and diubiquitin: PDB code 2W9N (Komander et al., 2009)) were used to solve the phasing problem, but no solution was obtained. Programs that were tried were Phaser, Molrep, *MrBUMP*, and Balbes.

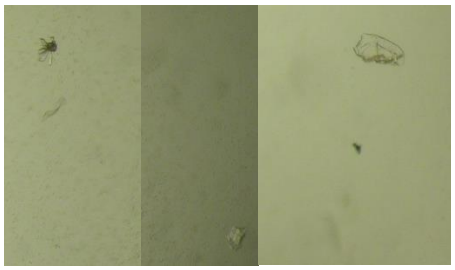
6.6.2. Crystallisation trial of protein complex USP4htt

The Crystallisation of protein complex USP4htt with Ubiquitin variants did not yield any crystals. Crystallisation trials (as seen in Table 6.1.) were conducted with active mutant USP4htt-H106N in complex with diubiquitin, Ubiquitin, or diubiquitin-L73X. Five different screens were performed, such as Morpheus, PACT Premier, Structure, PEG Suite II, and JCSG plus but there was no significant crystal growth. Some objects looked like crystals no diffraction was obtained (Figure 6.47).

Figure 6.47. Crystals from preparation complex USP4htt-H106N with UbGGG
Active site mutant USP4htt-H106N-UbGGG in JCSGplus D10 (left) and Active site mutant USP4htt-H106N-UbGGG in JCSGplus E3 (right).

6.6.3. Crystallisation trials of protein complexes of USP4C1C2 mutants

Although the structure of USP4C1C2 was already solved, how the protein interacts with Ubiquitin and its variants was interesting to study. Crystallisation trials were performed by using USP4C1C2-H881N active site mutant complexes with diubiquitin and UbGGG, and protein complexes of USP4C1C2-C311S-Y928X with Ub-GGG and diubiquitin. Several screens were conducted with the protein complexes, such as PACT premier, Morpheus, PEG II Suite, Structure, MIDAS, PGA, MemGold, Proplex, and JCSG plus. Some crystals grew, but no diffracting crystal was obtained (Figure 6.48).



A B C

Figure 6.48. The crystal from protein complex active site mutant USP4C1C2-H881N combine with Ub-GGG

A. Active site mutant USP4C1C2-H881N-UbGGG in Morpheus C5
B. Active site mutant USP4C1C2-H881N-UbGGG in Morpheus C5
C. Active site mutant USP4C1C2-H881N-UbGGG in Morpheus C7

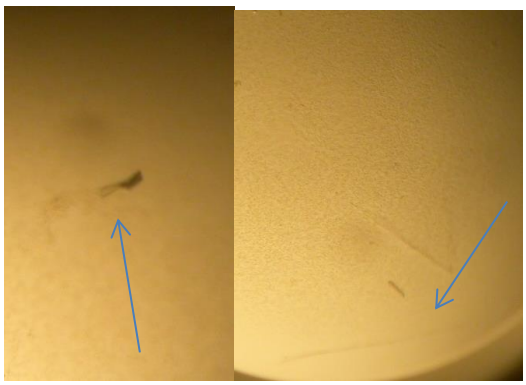
Based on crystallisation trials, the complex active site mutant USP4C1C2-H881N with UbGGG has produced some crystals. Other than that, the protein complex active site mutant with diubiquitin-L73X has not yet produced a crystal.

The crystal was from the complex of an active-site mutant USP4C1C2-H881N–UbGGG growth after about 5-month incubation at 10°C. The condition of the

crystal was Morpheus screen C5 and C7. C5 consists of 0.3 M Sodium Nitrate, 0.3 Sodium Phosphate dibasic, 0.3 M ammonium sulfate; 1 M Sodium Hepes and MOPS pH 7.5; and the precipitants 40% v/v PEG 500 MME, 20% w/v PEG 20000. Moreover, for Morpheus C7 was the same, but has a difference in precipitant namely 40% v/v Glycerol and 20% w/v PEG 4000. The sample was sent to synchrotron but not diffracted.

6.6.4. Crystallisation trials of protein complexes USP4C1C2-mitoxantrone

Another interaction of USP4C1C2 with inhibitor mitoxantrone was studied in this research. The protein of USP4C1C2 was mixed with mitoxantrone solution and prepared in the PACT and JCSG plus screen. The incubation was at 20°C. After several months there was a very small crystal growth, a single small long rectangular crystal (Figure 6.49). The condition was PACT primer A6: 0.1 M SPG pH 9.0 (Succinic Acid, sodium phosphate monobasic monohydrate, Glycine), 25 % w/v PEG 1500. The other condition was JCSG plus D8 (0.1 M Tris pH 8.0 and 40 % v/v MPD).



A B

Figure 6.49. The crystal from protein complex active site USP4C1C2-mitoxantrone
A. Active site USP4C1C2 in PACT premier A6; B. Active site USP4C1C2 in JCSG plus D8

6.6.5. Crystallisation trials of protein complexes with USP4ΔDU

Crystallisation trials of protein complexes of USP4ΔDU were performed only with active site mutant H881N and in complex with a monoubiquitin. The protein complex was put on the PGA and PACT Premier screen. After more than three months incubation at 10°C, there was still no crystal. The drops from the PGA screen were mostly clear, and in the PACT premier screen, most drops had precipitation. The concentration was 5 and 2.5 mg/mL of the protein complex.

6.6.6. Crystallisation trial of protein complex USP4httΔUbl2

The complexes of USP4httΔUbl2 H106N with diubiquitin was used for crystallisation. The protein complex of USP4httΔUbl2-H106N with diubiquitin was screened with PACT, MIDAS, and the Structure screen. The complex was incubated at 10°C and prepared at 2 mg/mL and 1 mg/mL concentrations in buffer 150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5. There was no crystal obtained.

6.6.7. Crystallisation trials of protein complexes of USP4DU with the SART3

USP4DU was crystallised with SART3. This complex was attempted to crystallise on some screens: PACT, JCSG plus, dan Morpheus. The concentration of this complex was 2 and 1 mg/mL. Another preparation was at a higher concentration. The protein complex was incubated at 10°C, but there was no crystal forming. All the crystallisation trial of USP4 constructs was in Table 6.8.

Table 6.8 Crystallisation trials of USP4 constructs with various conditions

Constructs	Complexes and buffer	Temp.	Screen	Concentration
USP4FL - H881N	**USP4FL-H881N with diubiquitin (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)**	**10°C**	**Morpheus JCSG plus** PGA	**32.1 mg/mL & 16.05 mg/mL**
	USP4FL-H881N with diubiquitin (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	PGA JCSG plus Morpheus	5.7 mg/mL & 2.8 mg/mL
	USP4FL-H881N with diubiquitin (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	**10°C**	Morpheus JCSG plus **PACT** premier	**6 mg/mL & 3 mg/mL**
	USP4FL-H881N with diubiquitin (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	JCSG plus PACT premier Morpheus	7 mg/mL & 3 mg/mL
	USP4FL-H881N with diubiquitin (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Classic suite	4 mg/mL & 2 mg/mL
	USP4FL-H881N with diubiquitin-L73X (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	JCSG plus PACT Premier Morpheus	0.65 mg/mL
	USP4FL-H881N with diubiquitin-L73X (Ammonium acetate 150 mM pH 6.8)	10°C	JCSG plus	0.39 mg/mL
	USP4FL-H881N with diubiquitin-L73X (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C (Lobster cell)	PACT Premier Morpheus PGA	6 mg/mL & 3 mg/mL
USP4FL - C311S-Y928X	USP4FL -C311S-Y928X with Ub-GGG (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5, 1 mM DTT)	10°C	MIDAS JCSG plus Morpheus PGA	4 mg/mL & 2 mg/mL
	USP4FL -C311S- Y928X with diubiquitin (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	MIDAS JCSG plus Morpheus PGA	4 mg/mL & 2 mg/mL
USP4C1C2 -H881N	USP4C1C2-H881N with diubiquitin (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Structure MIDAS PGA	9.3 mg/mL & 4.6 mg/mL
	USP4C1C2-H881N (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	20°C	PACT Structure MIDAS	8.2 mg/mL & 4.1 mg/mL
	USP4C1C2-H881N with diubiquitin (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Morpheus JCSG plus	27.6 mg/mL 13.8 mg/mL

Table 6.8 (Continued) Crystallisation trials of USP4 constructs with various conditions

Constructs	Complexes and buffer	Temp.	Screen	Concentration
	USP4C1C2-H881N with Ub-GGG (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	**10°C**	JCSG plus Structure **Morpheus**	**26.4 mg/mL & 13.2 mg/mL**
	USP4C1C2-H881N with diubiquitin-L73X (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	JCSG plus PACT premier Morpheus	23.9 mg/mL & 11,5 mg/mL
USP4C1C2 -C311S	USP4C1C2-C311S with diubiquitin (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	JCSG plus Structure Clear Start	5 mg/mL & 3 mg/mL
	USP4C1C2-C311S with UbGGG (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	MIDAS JCSG plus Structure PGA	5 mg/mL & 3 mg/mL
USP4C1C2	USP4C1C2 with Vialinin (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	20°C	JCSG plus Morpheus PACT Premier	5.1 mg/mL & 50 µM
	USP4C1C2 with Vialinin (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	20°C	Morpheus	110 µM - 100 µM
	USP4C1C2 with Vialinin (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	20°C	Morpheus	100 µM - 1000 µM
	USP4C1C2 with Mitoxantrone (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	**20°C**	**JCSG plus**	**5.1 mg/mL – 500 µM**
	USP4C1C2 with Mitoxantrone (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	**20°C**	**PACT Premier**	**5.1 mg/mL – 500 µM**
USP4C1C2 -C311S-Y928X	USP4C1C2-C311S-Y928X with UbGGG (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Structure JCSG plus MemGold	6 mg/mL & 3 mg/mL
	USP4C1C2-C311S-Y928X with diubiquitin (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C (Elad Cell)	PACT Premier Morpheus	2 mg/mL & 1 mg/mL
	USP4C1C2-C311S-Y928X with diubiquitin (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Structure PGA MemGold	4 mg/mL & 2 mg/mL
	USP4C1C2-C311S-Y928X with diubiquitin: (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Proplex PGA Structure	4 mg/mL & 2 mg/mL

Table 6.8 (Continued) Crystallisation trials of USP4 constructs with various conditions

Constructs	Complexes and buffer	Temp.	Screen	Concentration
USP4htt-H106N	USP4htt-H106N with diubiquitin (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	20°C	Morpheus PACT Premier Structure	1.9 mg/mL & 0.9 mg/mL
	USP4htt-H106N with diubiquitin (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Morpheus JCSG plus PACT	5.8 mg/mL & 2.9 mg/mL
	USP4htt-H106N with Ub-GGG (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	**10°C**	PACT Premier **JCSG plus**	**1.47 mg/mL**
	USP4htt-H106N with diubiquitin-L73X (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Morpheus	0.35 mg/mL
USP4htt-C311S	USP4htt-C311S with UbGGG (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Morpheus PACT premier JCSG plus	5 mg/mL & 2.5 mg/mL
	USP4htt-C311S with diubiquitin (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	JCSG plus PACT premier Morpheus	6 mg/mL & 3 mg/mL
USP4httΔUbl2-H106N	USP4httΔUbl2-H106N with diubiquitin (150 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	PACT MIDAS Structure	4.7 mg/mL & 2.3 mg/mL
USP4ΔDU-H881N	USP4ΔDU-H881N with Ubiquitin (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	PGA PACT Premier	5 mg/mL 2.5 mg/mL
USP4DU	USP4DU with SART3 (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	PACT premier JCSG plus Morpheus	8.7 mg/mL & 4.4 mg/mL
	USP4DU with SART3 (300 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Morpheus JCSG plus PACT premier	9.5 mg/mL & 4.8 mg/mL
	USP4DU with SART3 (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Morpheus PACT	5.4 mg/mL 2.7 mg/mL
	USP4DU with SART3 (100 mM NaCl, 1% Glycerol, 50 mM Tris pH 7.5)	10°C	Morpheus PACT Premier	4.3 mg/mL 2.2 mg/mL

*bold letter indicates the complex of protein forming a crystal

6.7. Discussion

Protein complex active site mutants USP4 construct with Ubiquitin variants reveal an interesting interaction in the protein complex purification. How the active site mutant and the Ubiquitin variant interact can be interpreted by comparing the elution volume and mass prediction. In the USP4FL-H881N complex, Ubiquitin was on the polyubiquitin formation. More than one ubiquitin binds to the monomer USP4FL-H881N while in diubiquitin interaction, both proteins were in dimer formation. Truncation diubiquitin on the C terminus to be diubiquitin-L73X changing the binding interaction with USP4FL-H881N. The interaction USP4FL-H881N with diubiquitin-L73X became 1:1 binding. Another mutation C311S and then Y928X in USP4FL decrease the ability of diubiquitin and Ub-GGG to bind. Only monomer diubiquitin and monomer Ub-GGG can bind with dimer USP4FL-C311S.

In the USP4htt construct, a mutation in H106N (originally H881N) showed to increase the binding interaction compare to C311S. The dimer formation in USP4htt-H106N with trimer diubiquitin decrease to be monomer USP4htt-C311S with monomer diubiquitin. Ub-GGG binding cannot compare because the complex USP4htt-H106N was eluted in the void volume. The mutation in L73X was also caused the complex USP4htt-H106N with diubiquitin-L73X elutes in the void volume.

The constructs of USP4httΔUbl2-H106N only have one complex formation with diubiquitin. Compare with USP4htt-H106N, USP4httΔUbl2-H106N has less binding interaction with diubiquitin.

USP4C1C2 construct always showed in monomer formation in all mutation (H881N, C311S and C311S-Y928X) when to form a complex with Ubiquitin variant. Ub-GGG showed consistently in monomer formation in every binding formation. Diubiquitin also in monomer formation, except interact with USP4C1C2-H881N mutation, was in the dimer. Moreover, the USP4C1C2-C311S itself was in dimer formation. The polyubiquitin chain also obtains in the complex of USP4ΔDU-H881N with Ubiquitin. Based on the mass prediction by elution volume, there was around a hexamer Ubiquitin bind to monomer USP4ΔDU.

In summary, the Ubiquitin can build a polyubiquitin chain in their original construct with variant active site mutant of USP4, while UbGGG constructs only always in monomer formation when binding with variant active site mutant of USP4 constructs.

Linear diubiquitin can bind in dimer or more formation when interacting in the mutation of H881N (or H106N) USP4 variant, however, in the mutation of C311S and C311S-Y928, the formation only in the monomer. The last, a diubiquitin-L73X construct need more binding interaction experiments to conclude.

The crystallisation trial variant active site mutant USP4 with Ubiquitin variant seems difficult to form a big crystal and then get a dataset after diffracted in the synchrotron. The best data was the crystal from protein complex USP4FL-H881N with diubiquitin (Table 6.8). Some optimation with more detail conditions does not reproduce the crystal. The interpretation of the dataset was difficult, and a molecular replacement has not yet provided a solution to solve the structure. Possibly there was a contaminant in the crystal based on the protein identification, a Glucosamine 6-phosphate synthase (GlmS).

Another protein complex with the ligand Ub-GGG looks could increase the probability for protein crystal production. The other forming crystal was from the complex of an active site mutant USP4C1C2-H881N–UbGGG and active site mutant USP4htt-H106N with ubiquitin GGG (Figure 6.47-6.48). The crystal growth after about 5-month incubation at 10°C. The condition of the crystal was Morpheus screen for active site mutant USP4C1C2 and JCSG plus for USP4htt-H106N. Although all of the crystal was sent to the synchrotron, the sample did not diffract, and there was no data set collected.

An interaction of catalytic core USP4-C1C2 with inhibitor mitoxantrone also need to optimise to form a crystal then analysed. This data will be interesting to compare with another similar DUB for example USP15-C1C2 which already known bind with mitoxantrone.

CHAPTER 7

Result and Discussion: Binding Interaction between USP4FL, USP4C1C2, and USP4ΔDU with Ubiquitin, diUbiquitin-L73X, and Ub-GGG

USP4 is a deubiquitinating enzyme which can cleave the isopeptide bond between the protein target and Ubiquitin. The binding interaction between USP4 and ubiquitin variants is interesting to study because not much is known about the details of this interaction. In this chapter, we investigate the binding interaction between USP4 active site mutant at H881N and C311S residues in the constructs: USP4FL, USP4C1C2, and USP4ΔDU, with Ubiquitin variants, including Ubiquitin, diubiquitin, diubiquitin L73X, and Ub-GGG. The complex active site mutants USP4 H881N and C311S constructs with Ubiquitin variant analyse using native mass spectrometry. Characterization of the binding interactions were performed using isothermal titration calorimetry (ITC).

The sample for mass spectrometry analysis were: for USP4FL: complex active site USP4FL with Ubiquitin, and active site mutant USP4FL-H881N with diubiquitin-L73X. While for ITC assay was USP4FL-C311S with Ubiquitin variant (Ubiquitin, Ubiquitin-GGG, and diubiquitin). The sample of USP4C1C2 for mass spectrometry was USP4C1C2-H881N with diubiquitin-L73X and Ub-GGG. For ITC was USP4C1C2-C311S with Ubiquitin variant (Ubiquitin, Ubiquitin-GGG, and diubiquitin). The sample for USP4ΔDU was USP4ΔDU and active site mutant USP4ΔDU-H881N with Ubiquitin for mass spectrometry analysis. In addition, active site mutant USP4C1C2-C311S with Ubiquitin variant was used for ITC test. Although there was a difference in the active site mutant used, the results are not expected to differ of a great extent.

7.1. USP4FL interaction with Ubiquitin variant

7.1.1. Mass spectrometry analysis

Electrospray ionisation-mass spectrometry is a valuable technique for analysing protein complexes. Using this type mass spectrometry, the formation of the complexes and their stoichiometry can be confirmed and whether USP4 contains additional ubiquitin binding sites investigated. Because of the limited sample preparation, the investigation only for the complex USP4FL with Ubiquitin and active site mutant USP4FL-H881N with diubiquitin and diubiquitin-L73X.

Figure 7.1 showed the spectrum of mass spectrometry from protein USP4FL with Ubiquitin. The sample for mass spectrometry was prepared by desalting using ammonium acetate after gel filtration purification to remove unwanted buffer components from the purification. The spectra showed there was free Ubiquitin, unbound USP4FL and complex of USP4FL with Ubiquitin present. The spectra revealed that a monomer of USP4FL bind with monomer Ubiquitin. The stoichiometry was 1:1 to form a complex. The unbound Ubiquitin and unbound USP4FL itself also found from the complex solution. The complex of USP4FL and unbound protein USP4FL and Ubiquitin were showed by mass from the spectra which were 120130.00 ± 84.33, 111467.70 ± 57.45 and 8558.81 ± 8.82 respectively. Another spectrum showed a peak at around 134 kDa (green letter) that possibly indicated a bigger mass which contains a

complex of monomer USP4FL with dimer Ubiquitin, but further investigation needed. The presence of free USP4FL suggests that USP4FL has lower affinity to Ubiquitin.

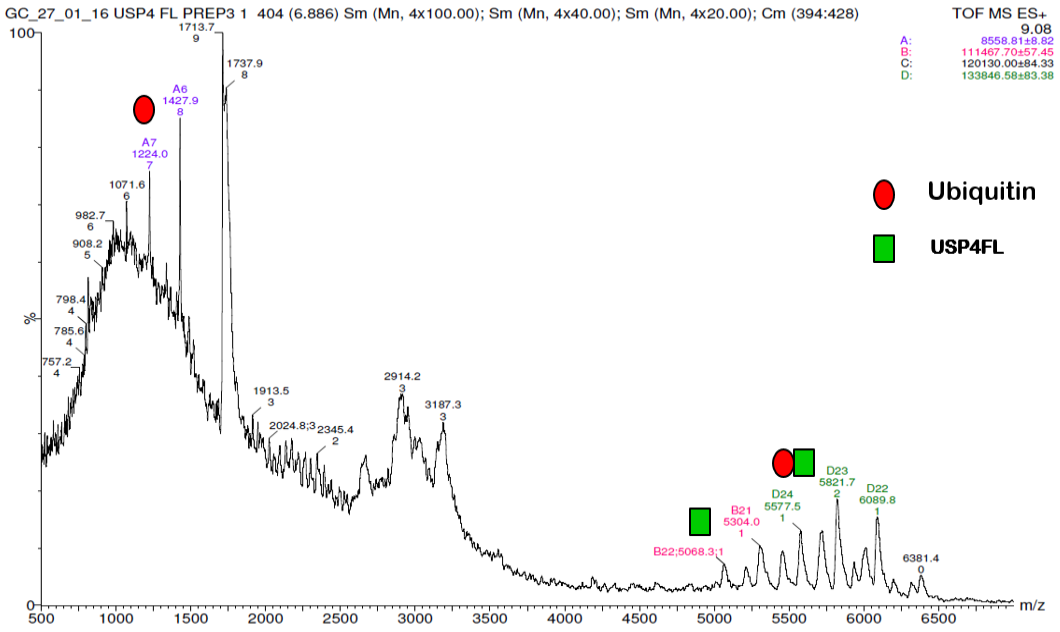


Figure 7.1. Spectrum of complex USP4FL with Ubiquitin
The spectra show three types of peaks that correspondent to unbound Ubiquitin, unbound USP4FL and complex USP4FL with Ubiquitin. Not all USP4FL bind to Ubiquitin. The presence of free Ubiquitin suggests the weak interaction between USP4FL with Ubiquitin.

The complex of active site mutant USP4FL-H881N with diubiquitin only showed the spectra of unbound diubiquitin in 17114.98 ± 28.91. There were no spectra of USP4FL (Figure 7.2). Possibly some degradation or aggregation developed during sample preparation. Complex active site mutant USP4FL-H881N with diubiquitin was not stable in the ammonium acetate buffer.

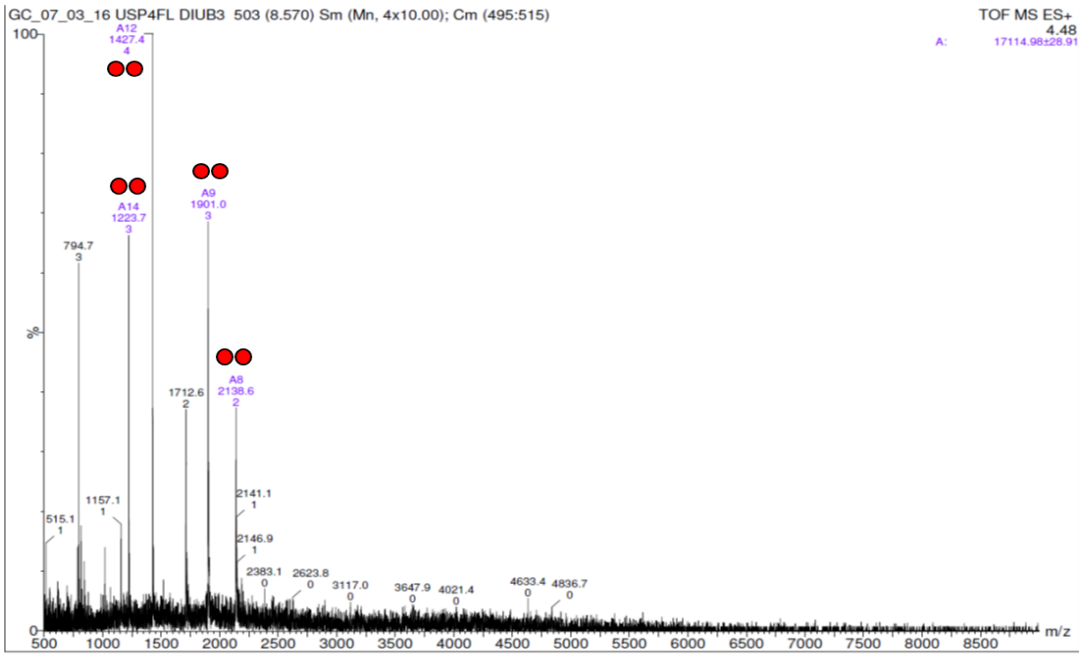


Figure 7.2. Spectrum of complex USP4FL-H881N with diubiquitin.
The spectra only contain diubiquitin (double red circle) and there was no USP4FL-H881N spectrum. USP4FL-H881N still not stable on the ammonium acetate as the buffer for mass spectrometry assay.

Another protein complex was USP4FL-H881N with diubiquitin-L73X to investigate how the diubiquitin bind to the USP4FL-H881N and to study the effect of truncation Leucine to stop codon. However, unfortunately, the spectrum of protein complex USP4FL-H881N with diubiquitin-L73X revealed that only contain ubiquitin and diubiquitin (Figure 7.3 and Table 7.1). Possibly some degradation or aggregation developed during sample preparation. The solution only had the unbound diubiquitin and monomer Ubiquitin. The presence of ubiquitin maybe come from the degradation of diubiquitin.

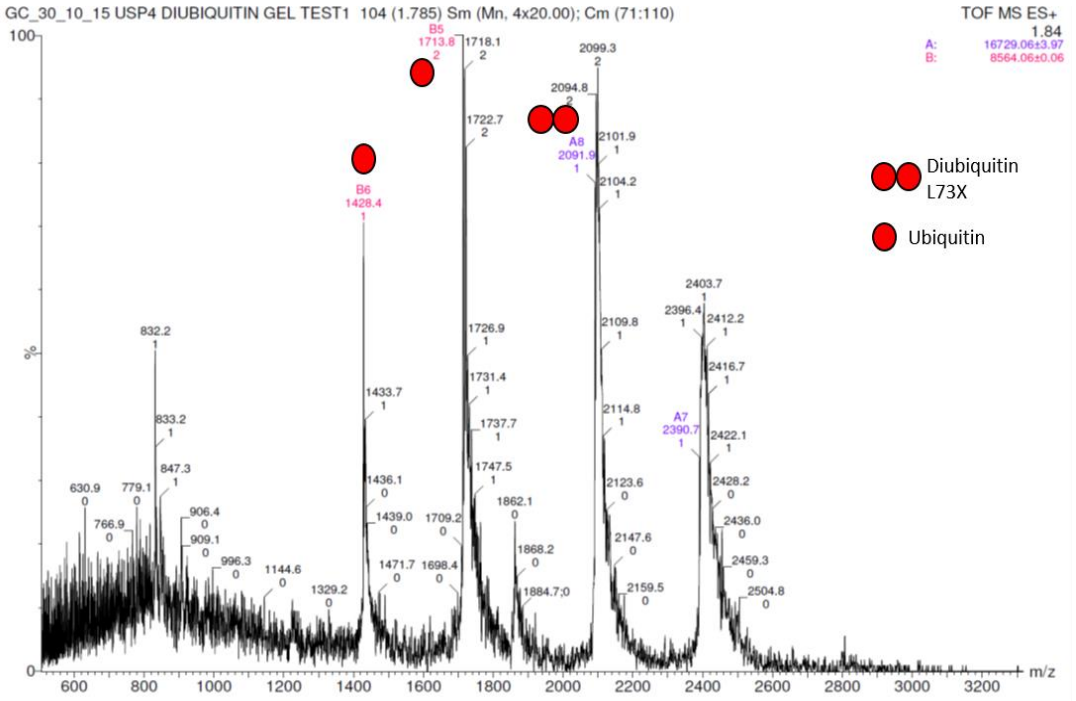


Figure 7.3. Spectrum of complex USP4FL-H881N with diubiquitin-L73X.
The spectra seemed to contain what looks like diubiquitin (double red circle) and monomer ubiquitin (single red circle), and there was no USP4FL-H881N spectrum. USP4FL-H881N, as a large protein, could be not stable on the ammonium acetate as the buffer for mass spectrometry assay.

7.1.2. Isothermal Titration Calorimetry Analysis

The affinity of the interaction between USP4FL and Ubiquitin variants is not fully described. Clerici, et al., (2014) explain that the N terminus DUSP-Ubl domain is required for the full USP4 enzymatic activity *in vitro*. The Ubiquitin release was promoted by the DUSP domain using catalytic turnover.

The investigation of the binding interaction between USP4FL-C311S and Ubiquitin variants (Ubiquitin, diubiquitin and Ubiquitin-GGG) used isothermal titration calorimetry. In this experiment, the concentration of protein USP4FL-C311S was prepared at 10 and 20 μM. At the higher concentration, the protein

Table 7.1. The protein complex of USP4FL-H881N with diubiquitin result by ESI-MS analysis

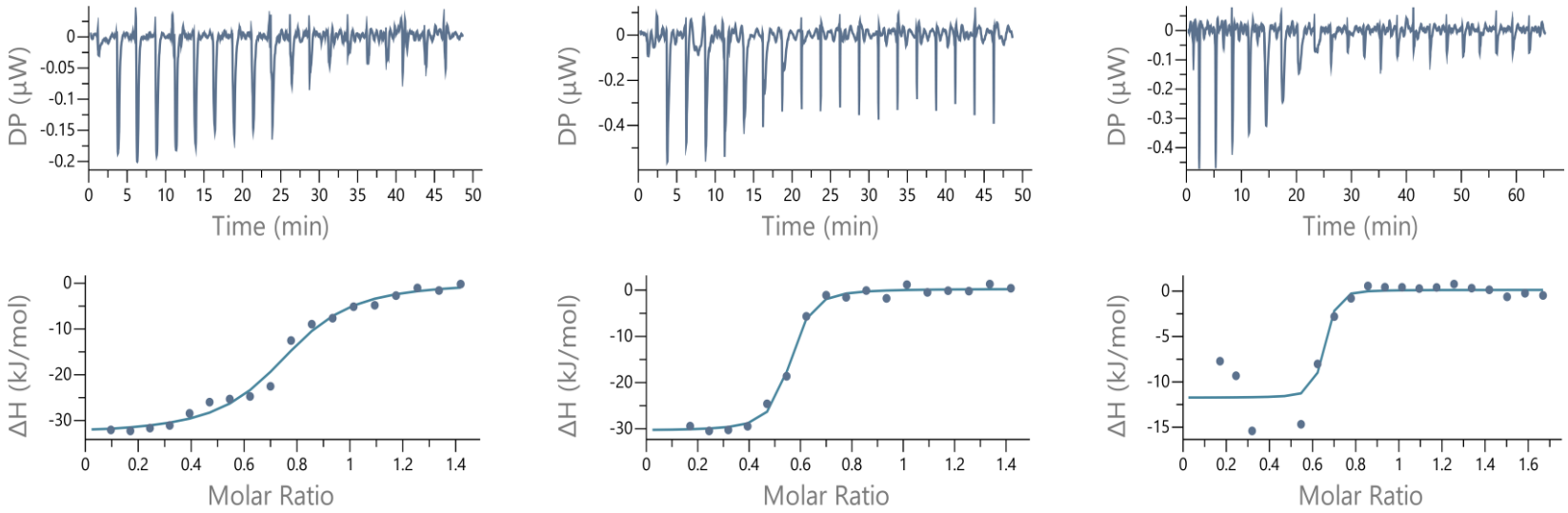
Protein or Complex Protein	MW based on Protparam (Da)	MW based on MS result (Da)	Prediction of protein and interaction	Purification / Sample preparation
USP4FL-H881N + diUb-L73X	111121.9 (USP4FL-H881N) 16728.2 (diUb-L73X) 8564.8 (Ubiquitin)	- 16729.06 ±3.97 8564.06 ± 0.06	- Diubiquitin-L73X Ubiquitin	GF: 300 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
USP4FL + ubiquitin	111121.9 (USP4FL) 8564.8 (Ubiquitin)	8558.81 ± 8.82 111467.70 ± 57.45 120130.00 ± 84.33 133846.58 ± 83.38	Ubiquitin USP4FL USP4FL - Ubiquitin (1:1) USP4FL – dimer Ubiquitin	GF: 150 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
USP4FL-H881N + diubiquitin	111121.9 (USP4FL-H881N) 17111.6 (diubiquitin)	- 17114.98 ± 28.91	- Diubiquitin	GF: 100 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8

tends to be difficult to inject to the ITC machine due to bubble formation. The ligand (Ubiquitin variant) was prepared at a ten times higher concentration.

Binding interaction data between USP4FL-C311S with Ubiquitin variants showed that USP4FL-C311S has the tightest binding with diubiquitin, then Ubiquitin-GGG and Ubiquitin, based on the K_D value. The K_D value for diubiquitin was the lowest compared to Ubiquitin-GGG and ubiquitin. K_D value for USP4FL-C311S with diubiquitin was $5.83e^{-9} \pm 15.6e^{-9}$ M. The K_D value for Ubiquitin-GGG was about 10 times higher and for Ubiquitin was more than 40 times higher. The additional Glycine residue on the Ubiquitin structure increases the binding interaction with active site mutant USP4FL-C311S.

Although there was no mass spectrometry data for interaction between USP4FL-C311S with Ubiquitin-GGG and diubiquitin, mass spectrometry data of USP4FL-H881N with ubiquitin showed there was more unbound ubiquitin compare to the complex USP4FL with ubiquitin. This data confirmed the ITC data that suggest a weaker binding interaction of USP4FL with ubiquitin.

The binding interaction between USP4FL-C311S with all Ubiquitin variants is comprised of hydrogen bonding and hydrophobic interactions. These indicated by the negative or favorable binding enthalpy (ΔH) and entropy factor (TΔS) (Figure 7.4).



USP4FL-C311S (µM)	: 10	USP4FL-C311S (µM)	: 20	USP4FL-C311S (µM)	: 20
Ubiquitin (µM)	: 100	Ubiquitin-GGG (µM)	: 200	diubiquitin (µM)	: 200
N (sites)	: 0.736 ± 1.5 e^{-2}	N (sites)	: 0.522 ± 5.2 e^{-3}	N (sites)	: 0.623 ± 2.6 e^{-2}
KD (M)	: 240 e^{-9} ± 50.7 e^{-9}	KD (M)	: 59.1 e^{-9} ± 14.5 e^{-9}	KD (M)	: 5.83 e-9 ± 15.6 e^{-9}
ΔH (kJ/mol)	: -33.3 ± 0.916	ΔH (kJ/mol)	: -30.7 ± 0.561	ΔH (kJ/mol)	: -9.74 ± 0.956
ΔG (kJ/mol)	: -37.8	ΔG (kJ/mol)	: -41.3	ΔG (kJ/mol)	: -47.0
-T ΔS (kJ/mol)	: -4.53	-T ΔS (kJ/mol)	: -10.6	-T ΔS (kJ/mol)	: -37.3

Figure 7.4. ITC graph of Ubiquitin 100 µM titrated into USP4FL-C311S 10 µM (left), and Ubiquitin-GGG 200 µM into USP4FL-C311S 20 µM (center), and diUbiquitin 200 µM titrated into USP4FL-C311S 20 µM (right)

7.2. USP4C1C2 interaction with Ubiquitin variant

7.2.1. Mass Spectrometry analysis

Mass spectrometry analysis of protein complex USP4C1C2-H881N with diubiquitin-L73X showed the form of the complex, unbound USP4C1C2 and free diubiquitin (Figure 7.5). The spectra for USP4C1C2 were A12 and A13 (purple), and protein complex was B13 and B14 (pink) and diubiquitin itself (D6, D7, green).

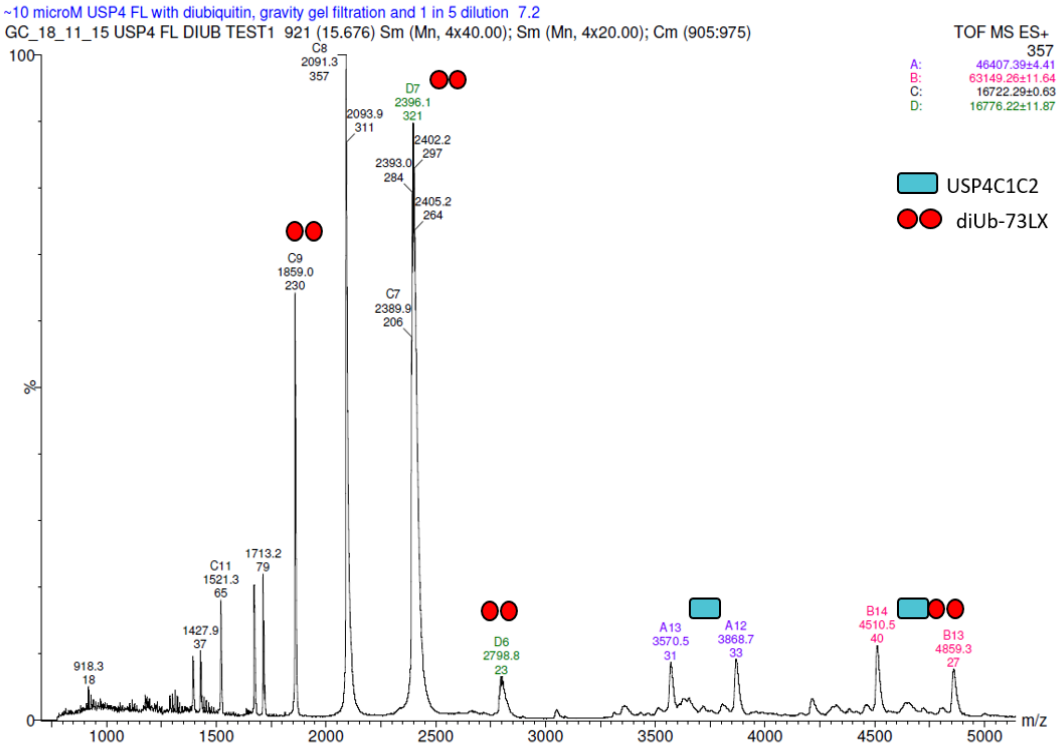


Figure 7.5. Spectrum of complex USP4C1C2-H881N with diubiquitin-L73X.
The spectra contain USP4C1C2 (A12, A13, purple) and protein complex (B13, B14, pink) and diubiquitin (D6, D7, green).

The ratio of unbound USP4C1C2 to bound complex USP4C1C2-H881N with diubiquitin-L73X was 64 : 67. About 51% of diubiquitin-L73X was interacting

with USP4C1C2-H881N, and the other is unbound. The result indicates that there was a binding affinity between USP4C1C2-H881N with diubiquitin-L73X. The investigation of binding interaction continued to a complex of USP4C1C2-H881N with Ub-GGG. USP4C1C2-H881N has a complex formation with Ub-GGG. Data from spectra showed that the formation of the complex was around 70%. Some free USP4C1C2-H881N monomer unbound and unbound Ubiquitin-GGG were available. The degradation product of USP4C1C2 H881N was on the spectra (green) (Figure 7.6; Table 7.2).

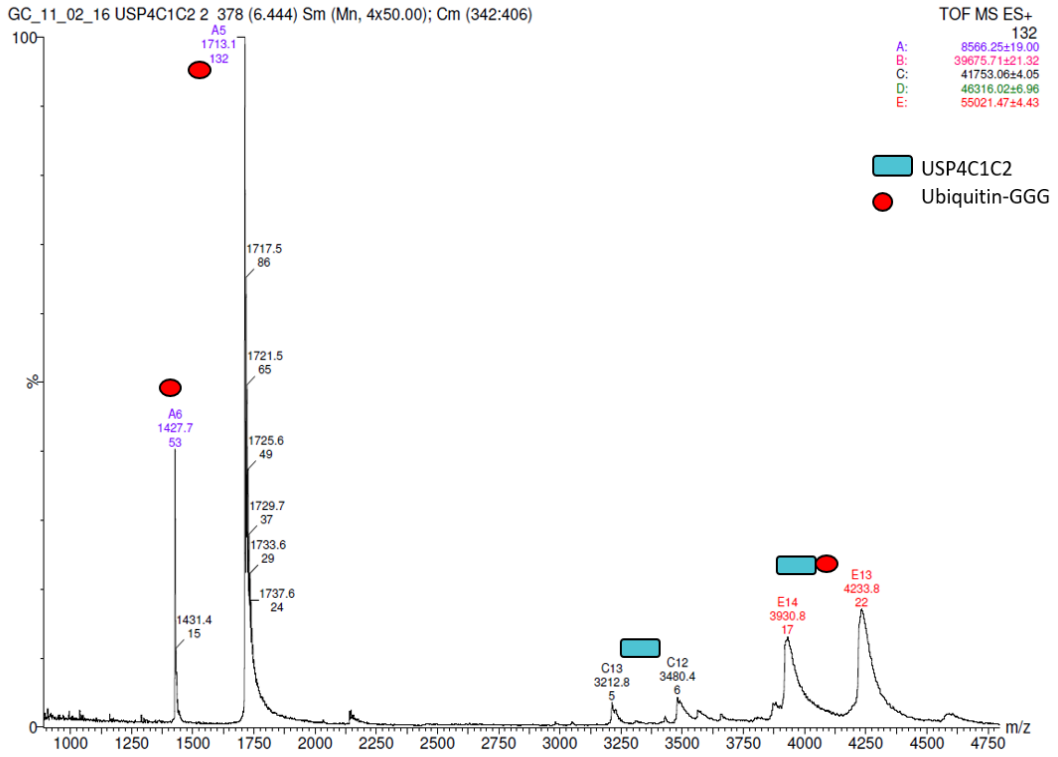


Figure 7.6. Spectrum of complex USP4C1C2-H881N with ubiquitin-GGG.
The spectra contain USP4C1C2 itself, protein complex USP4C1C2-H881N with Ub-GGG and Ubiquitin-GGG. Some degradation of USP4C1C2-H881N also showed on the spectra

Table 7.2. The protein complex of USP4C1C2-H881N with diubiquitin-L73X and Ub-GGG result by ESI-MS analysis

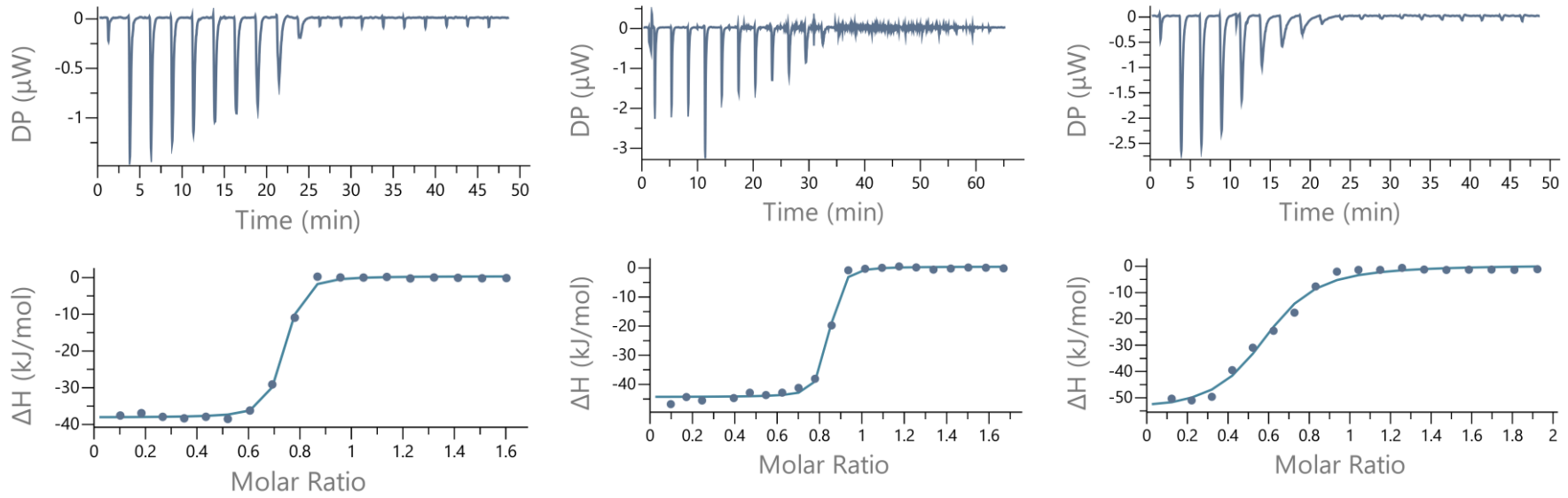
Protein Complex	MW based on Protparam (Da)	MW based on MS result (Da)	Prediction mass	Prediction of protein and/or interaction	Purification / Sample preparation
USP4C1C2-H881N + diUb-L73X	46411.2 (USP4C1C2) 16728.2 (diUb-L73X)	16722.29 ± 0.63		Diubiquitin-L73X	GF: 300 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		16776.22 ± 11.87		Diubiquitin-L73X	
		46407.39 ± 4.41		USP4C1C2	
		63149.26 ± 11.64	63139.4	USP4C1C2- Diubiquitin-L73X (1:1)	
USP4C1C2-H881N + Ub-GGG	46411.2 (USP4C1C2) 8621.9 (Ub-GGG)	8566.25 ± 19		Ub-GGG	GF: 100 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		39675.71± 21.32	43126.6	USP4C1C2 without His-Tag	
		41753.06 ± 4.05	43126.6	USP4C1C2 without His-Tag	
		46316.02 ± 6.96		USP4C1C2	
		55021.47 ± 4.43	55033.1	USPC1C2 : Ub-GGG (1:1)	

7.2.2. Isothermal Titration Calorimetry analysis

Active site mutant of USP4C1C2-H881N already known interacts with diubiquitin-L73X and Ubiquitin-GGG from mass spectrometry analysis. Using ITC method, the difference binding interaction was continued to be investigated. In this investigation, active site mutant of USP4C1C2-C311S was titrated with Ubiquitin, Ubiquitin-GGG, and diubiquitin. The concentration of the protein sample USP4C1C2-C311S was about ten-fold lower than the ligand Ubiquitin variant. The protein concentration of USP4C1C2-C311S was 30, and 50 µM and the concentration of the ligand ubiquitin variant was 250 – 500 µM.

ITC graph results (Figure 7.7) showed that all interactions between USP4C1C2-C311S with ubiquitin, Ubiquitin-GGG, and diubiquitin were binding well. Data K_D showed that USP4C1C2-C311S with ubiquitin-GGG has the lowest value compared to Ubiquitin and diubiquitin. The K_D value was about 23.6 nM, 37.2 nM, and 1.5 µM respectively. This value concludes that USP4C1C2-C311S has possibly the highest affinity for ubiquitin-GGG, but is very similar to Ubiquitin. Diubiquitin has a lower affinity for USP4C1C2-C311S.

The binding interaction between USP4C1C2-C311S with Ubiquitin contributed hydrogen bonding and hydrophobic interactions because the ΔH and -TΔS was a negative value. However, the binding affinity between USP4C1C2-C311s with



USP4C1C2-C311S (µM)	: 30	USP4C1C2-C311S (µM)	: 30	USP4C1C2-C311S (µM)	: 50
Ubiquitin (µM)	: 250	Ubiquitin-GGG (µM)	: 300	diubiquitin (µM)	: 500
N (sites)	: 0.696 ± 3.23 e-3	N (sites)	: 0.804 ± 3.3e-3	N (sites)	: 0.558 ± 1.5 e-2
KD (M)	: 37.2 e^{-9} ± 6.70 e^{-9}	KD (M)	: 23.6 e^{-9} ± 6.30e^{-9}	KD (M)	: 1.55e^{-6} ± 322e^{-9}
ΔH (kJ/mol)	: -38.4 ± 0.328	ΔH (kJ/mol)	: -44.7 ± 0.430	ΔH (kJ/mol)	: -55.9 ± 1.97
ΔG (kJ/mol)	: -42.4	ΔG (kJ/mol)	: -43.6	ΔG (kJ/mol)	: -33.2
-T ΔS (kJ/mol)	: -4.03	-T ΔS (kJ/mol)	: 1.12	-T ΔS (kJ/mol)	: 22.7

Figure 7.7. The ITC graph of Ubiquitin 250 µM titrated into USP4C1C2-C311S 30 µM (left) and ubiquitin-GGG 300 µM titrated into USP4C1C2-C311S 30 µM (center), diubiquitin 500 µM titrated into USP4C1C2-C311S 50 µM (right)

Ubiquitin-GGG and diubiquitin involves more conformational changes as indicated by the positive value of -TΔS.

7.3. USP4ΔDU interaction with Ubiquitin variant

7.3.1. Mass spectrometry analysis

The construct of USP4ΔDU was the USP4FL without DUSP Ubl1 domain. Based on the mass spectrometry result (Figure 7.8), USP4ΔDU construct can be monomer and dimer formation, which means can be binding each other to form a complex or in unbound formation as a monomer. The spectra at the 154.5 and 160.2 indicated the presence of a dimer USP4ΔDU, and the spectra of 79 refer to monomer USP4ΔDU.

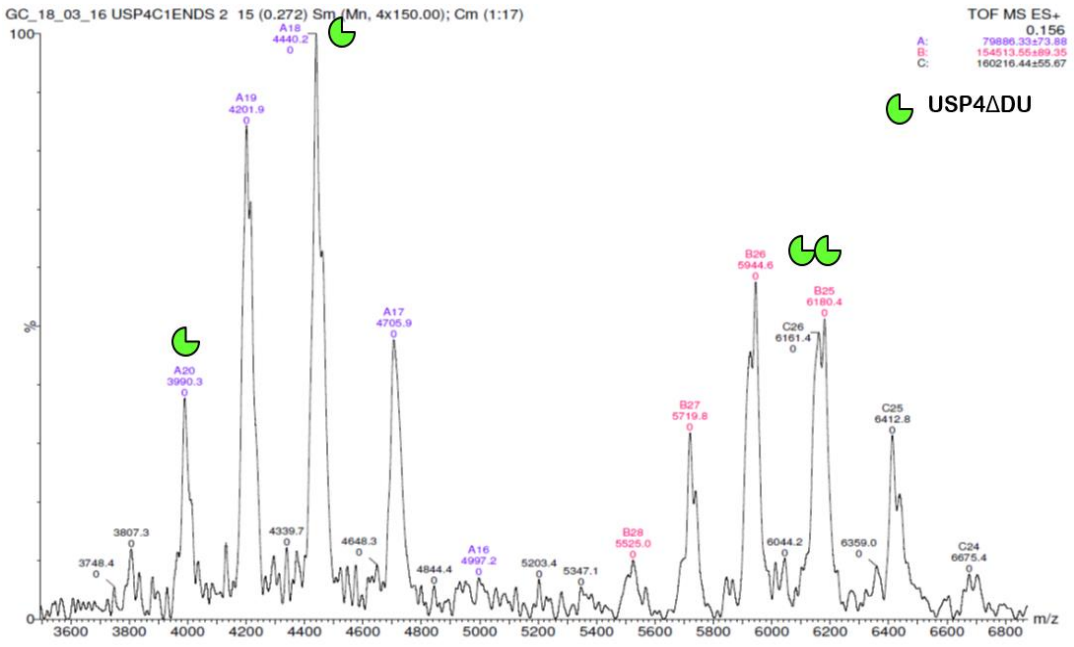


Figure 7.8. Mass spectrometry result of the active site USP4ΔDU
The spectra showed that USP4ΔDU could be monomer or dimer structure. The active protein of USP4ΔDU has no DUSP Ubl domain.

The further investigation was an interaction between USP4ΔDU-H881N with Ubiquitin. The spectra data showed there were: monomer ubiquitin and complex USP4ΔDU with ubiquitin. There was no dimer of USP4ΔDU. The only monomer of USP4ΔDU bind to form a complex with monomer Ubiquitin (Figure 7.9). The mass spectrometry data also showed that there was no unbound of USP4ΔDU in the complex with Ubiquitin (Table 7.3). Possibly, the Ubiquitin prevents USP4ΔDU from forming a dimer. All USP4ΔDU made a complex with Ubiquitin with binding stoichiometry 1:1. Otherwise, free Ubiquitin was visible.

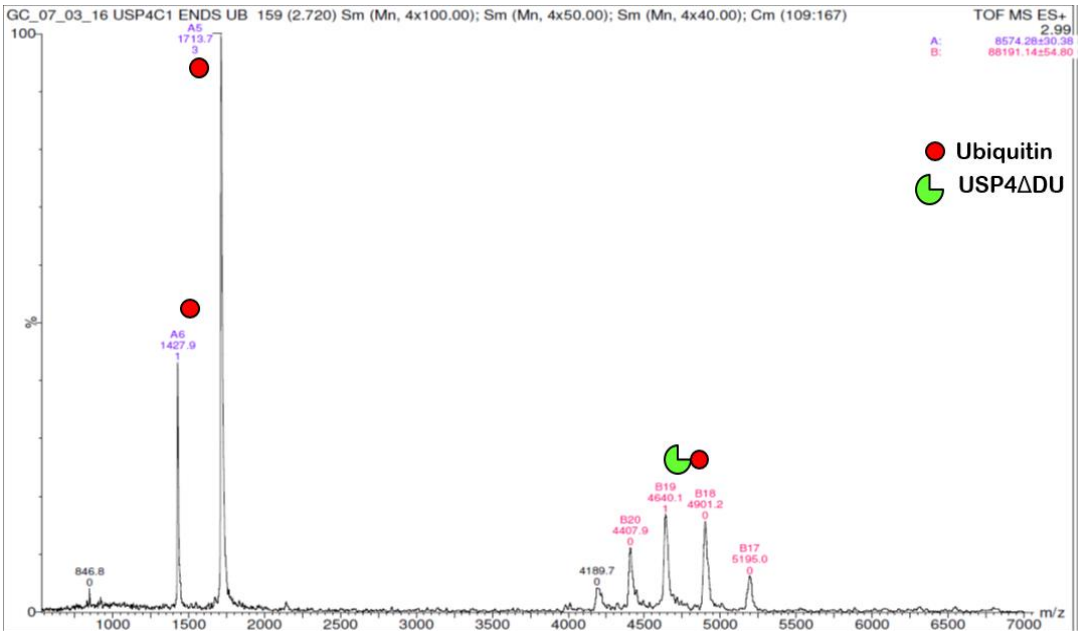


Figure 7.9. Mass spectrometry result of protein active site USP4ΔDU-H881N and Ubiquitin The spectra showed unbound Ubiquitin and complex of USP4ΔDU with Ubiquitin. All the USP4ΔDU was bind with Ubiquitin. There were no spectra unbound USP4ΔDU

Table 7.3. Protein complex of USP4ΔDU-H881N with ubiquitin result by ESI-MS analysis

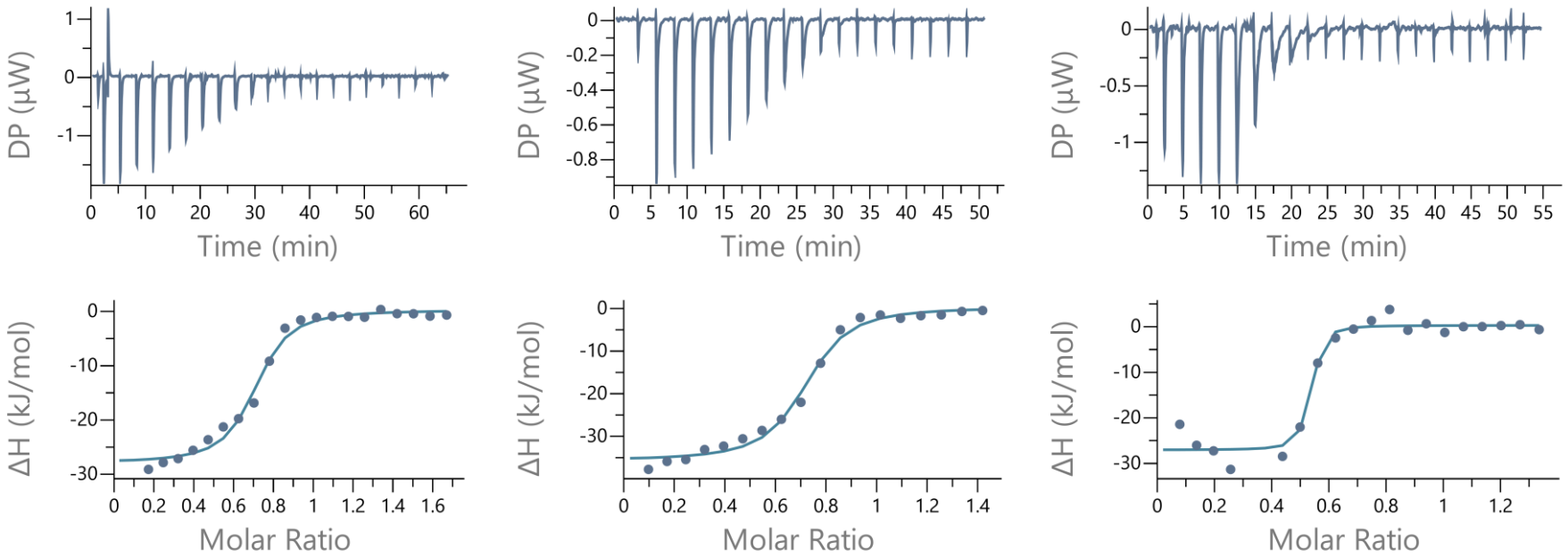
Protein or Protein Complex	MW based on Protparam (Da)	MW based on MS result (Da)	Mass prediction of complex	Prediction of protein and/or interaction	Purification / Sample preparation
USP4ΔDU	79467.3	79886.33 ± 73.88		USP4ΔDU	From Ni column directly GF S200 with Ammonium acetate 150 mM pH 6.8 as buffer
		154513.55 ± 89.35	152191.4 (monomer: 76095.7)	USP4ΔDU without His tag (dimer)	
		160216.44 ± 55.67	158934.6	USP4ΔDU-USP4ΔDU (dimer)	
USP4ΔDU-H881N + Ubiquitin	79467.3 (USP4ΔDU) 8564.8 (Ubiquitin)	8555.29 ± 7.05		Ubiquitin	GF: 100 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		88028.70 ± 39.99	88032.1	USP4ΔDU-Ubiquitin (1:1)	
		83210.90 ± 59.61	84660.5	USP4ΔDU without His Tag-Ubiquitin (1:1)	
		88168.74 ± 18.53	88032.1	USP4ΔDU-Ubiquitin (1:1)	

7.3.2. Isothermal Titration Calorimetry Analysis

Further investigation of binding interaction performed to USP4ΔDU-C311S with Ubiquitin variant. The USP4ΔDU was prepared at the concentration 20 and 50 μM, and the Ubiquitin variant was 200, 400 and 500 μM. The concentration of Ubiquitin variant as the ligand was ten-fold higher than the protein sample of USP4ΔDU.

The ITC graph showed that USP4ΔDU has the strongest interaction with diubiquitin, then followed to Ubiquitin-GGG and Ubiquitin. The K_D values were 34.7 nM, 213 nM, and 527 nM respectively. This result correlates with the interaction between USP4FL-C311S with Ubiquitin variant. The diubiquitin has the strongest interaction. Compare to the binding interaction result of complex USP4C1C2-C311S with Ubiquitin variant; it can be suggested that on the larger protein construct with insert region (USP4FL and USP4ΔDU) the interaction with diubiquitin will increase. This suggestion also for the interaction with Ubiquitin-GGG, which additional Glycine residue is increasing the affinity.

The ITC data (Figure 7.10) also explain that the binding interaction of complex USP4ΔDU with Ubiquitin variant supported by hydrogen bonding and hydrophobic interactions as the negative value for binding enthalpy (ΔH) and -TΔS.



USP4ΔDU-C311S (µM)	: 50	USP4ΔDU-C311S (µM)	: 20	USP4ΔDU-C311S (µM)	: 50
Ubiquitin (µM)	: 500	Ubiquitin-GGG (µM)	: 200	diubiquitin (µM)	: 400
N (sites)	: 0.680 ± 1.1e-2	N (sites)	: 0.691 ± 1.0e^{-2}	N (sites)	: 0.506 ± 8.6e^{-3}
KD (M)	: 527e^{-9} ± 123e^{-9}	KD (M)	: 213e-9 ± 46.8 e^{-9}	KD (M)	: 34.7e^{-9} ± 27.6 e^{-9}
ΔH (kJ/mol)	: -28.2 ± 0.705	ΔH	: -36.0 ± 0.775	ΔH	: -27.3 ± 1.06
ΔG (kJ/mol)	: -35.9	ΔG	: -38.1	ΔG	: -42.6
-T ΔS (kJ/mol)	: -7.70	-T ΔS	: -2.08	-T ΔS	: -15.3

Figure 7.10. The graph of binding interaction using ITC of USP4ΔDU-C311S with diubiquitin (left) and USP4ΔDU-C311S with Ub-GGG (center), and USP4ΔDU-C311S with diubiquitin (right)

7.4. Comparison binding affinity of various USP4-C311S with Ubiquitin variant

The binding affinity between various USP4-C311S with Ubiquitin variant showed an interesting data. Ubiquitin, the smallest ligand, has the tightest affinity with USP4C1C2-C311S, followed by USP4FL-C311S and USP4ΔDU-C311S. Removing DUSP-Ubl domain decreasing the affinity of Ubiquitin with USP4 (Table 7.4).

Table 7.4. Comparison binding affinity between various USP4-C311S with Ubiquitin

Parameter	USP4FL-C311S (μM): 10 Ubiquitin (μM): 100	USP4C1C2-C311S (μM): 30 Ubiquitin (μM): 250	USP4ΔDU-C311S (μM): 50 Ubiquitin (μM): 500
N (sites)	0.736 ± 1.5 e^{-2}	0.696 ± 3.23 e^{-3}	0.680 ± 1.1e^{-2}
KD (M)	240 e^{-9} ± 50.7 e^{-9}	37.2 e^{-9} ± 6.70 e^{-9}	527e-9 ± 123e^{-9}
ΔH (kJ/mol)	-33.3 ± 0.916	-38.4 ± 0.328	-28.2 ± 0.705
ΔG (kJ/mol)	-37.8	-42.4	-35.9
-T ΔS (kJ/mol)	-4.53	-4.03	-7.70

When there was an additional Glycine residue on the Ubiquitin, Ub-GGG, USP4C1C2-C311S also has the highest affinity (Table 7.5). But, USP4FL-C311S was the lowest affinity with Ub-GGG. Additional Glycine residue increase the binding affinity with construct without DUSP-Ubl. The reversible data with Ub-GGG binding interaction was showed by diubiquitin binding. The construct USP4C1C2-C311S has the lowest affinity, followed by USP4ΔDU-C311S and USP4FL-C311S. USP4FL-C311S binding well with diubiquitin (Table 7.6).

Table 7.5. Comparison binding affinity between various USP4-C311S with Ub-GGG

Parameter	USP4FL-C311S (μM): 20 Ubiquitin-GGG (μM): 200	USP4C1C2-C311S (μM): 30 Ubiquitin-GGG (μM): 300	USP4ΔDU-C311S(μM): 20 Ubiquitin-GGG (μM): 200
N (sites)	0.522 ± 5.2 e^{-3}	0.804 ± 3.3e^{-3}	0.691 ± 1.0e^{-2}
KD (M)	59.1 e^{-9} ± 14.5 e^{-9}	23.6 e^{-9} ± 6.30e^{-9}	213e-9 ± 46.8 e^{-9}
ΔH (kJ/mol)	-30.7 ± 0.561	-44.7 ± 0.430	-36.0 ± 0.775
ΔG (kJ/mol)	-41.3	-43.6	-38.1
-T ΔS (kJ/mol)	-10.6	1.12	-2.08

Table 7.6. Comparison binding affinity between various USP4-C311S with diubiquitin

Parameter	USP4FL-C311S (μM): 20 diubiquitin (μM): 200	USP4C1C2-C311S (μM): 50 diubiquitin (μM): 500	USP4ΔDU-C311S (μM): 50 diubiquitin (μM): 400
N (sites)	0.623 ± 2.6 e^{-2}	0.558 ± 1.5 e^{-2}	0.506 ± 8.6e^{-3}
KD (M)	5.83 e-9 ± 15.6 e^{-9}	1.55e^{-6} ± 322e^{-9}	34.7e^{-9} ± 27.6 e^{-9}
ΔH (kJ/mol)	-9.74 ± 0.956	-55.9 ± 1.97	-27.3 ± 1.06
ΔG (kJ/mol)	-47.0	-33.2	-42.6
-T ΔS (kJ/mol)	-37.3	22.7	-15.3

7.5. The binding interaction between USP4DU and SART3

SART3 was already known to bind with USP15DU. A dimer of the SART3 HAT domains can bind two molecules of USP15DU (Grazette, 2015). This was also later confirmed in a study by Park, et al., (2016). This research wants to study the binding interaction between SART3 and USP4DU using the ESI-mass spectrometry.

The mass spectrometry analysis was done by three different sample preparations at the protein purification process. The USP4DU and SART3 were expressed separately and purified by the co-purification method. The first purification used high salt 300 mM NaCl in the buffer followed by desalting of

the sample into 150 mM ammonium acetate pH 6.8. From this purification, the mass spectrum showed a heterogenous mixture in the USP4DU-SART3 sample. The spectrum showed there was: unbound USP4DU, free monomer SART3, a complex of monomer USP4DU with monomer SART3, complex dimer SART3 with monomer USP4DU, and dimer SART3 with dimer USP4DU. The presence of a monomer SART3 is unusual because SART3 was already known as a dimer (Figure 7.11). An interaction between monomer SART3 and monomer USP4DU was also interesting to observe.

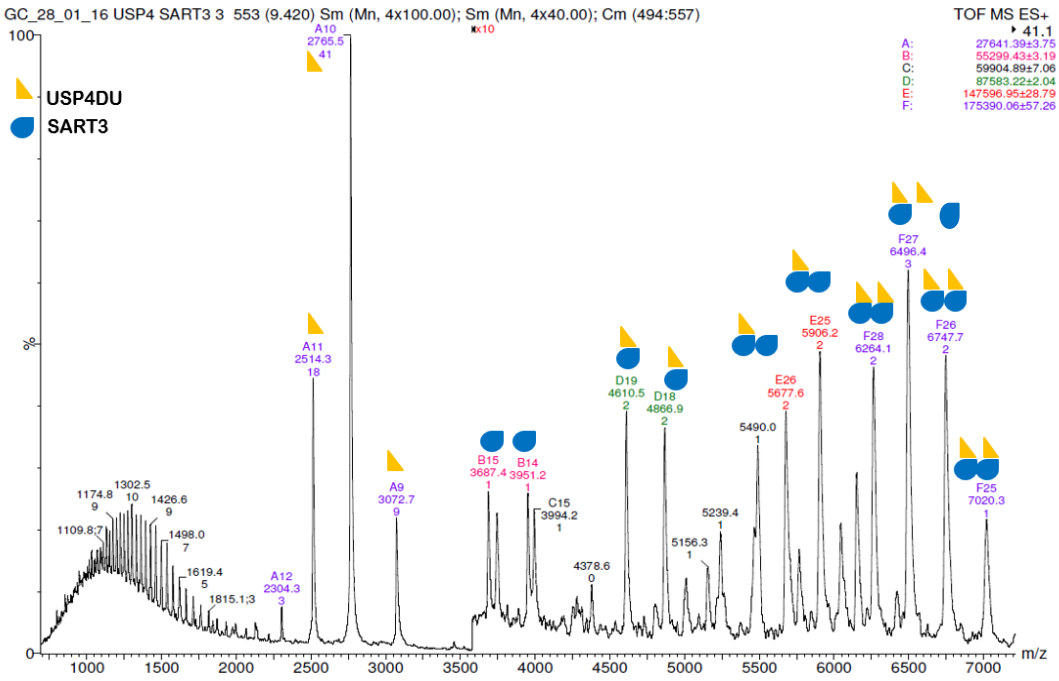


Figure 7.11. Mass spectrometry result of protein active site USP4DU-SART3 (1st sample) The spectra showed the presence of dimer SART3, dimer SART3 binds with monomer USP4DU, and dimer SART3 binds with dimer USP4DU.

Analysis continues with the second purification using a lower concentration of salt, namely 100 mM NaCl, then followed by dialysis into 150 mM ammonium acetate pH 6.8 for desalting the sample. Based on this purification, the spectra

were different; there was: free USP4DU, unbound SART3, dimer SART3, dimer SART3 bind with monomer USP4DU and dimer SART3 bind with dimer USP4DU. There was no dimer SART3 on the first purification, and there was no binding monomer USP4DU and monomer SART3 on the second purification (Figure 7.12). The change of the salt concentration might affect the electrostatic interactions in solution. The SART3 dimer becomes more stable, and there was an increase number of a dimer SART3.

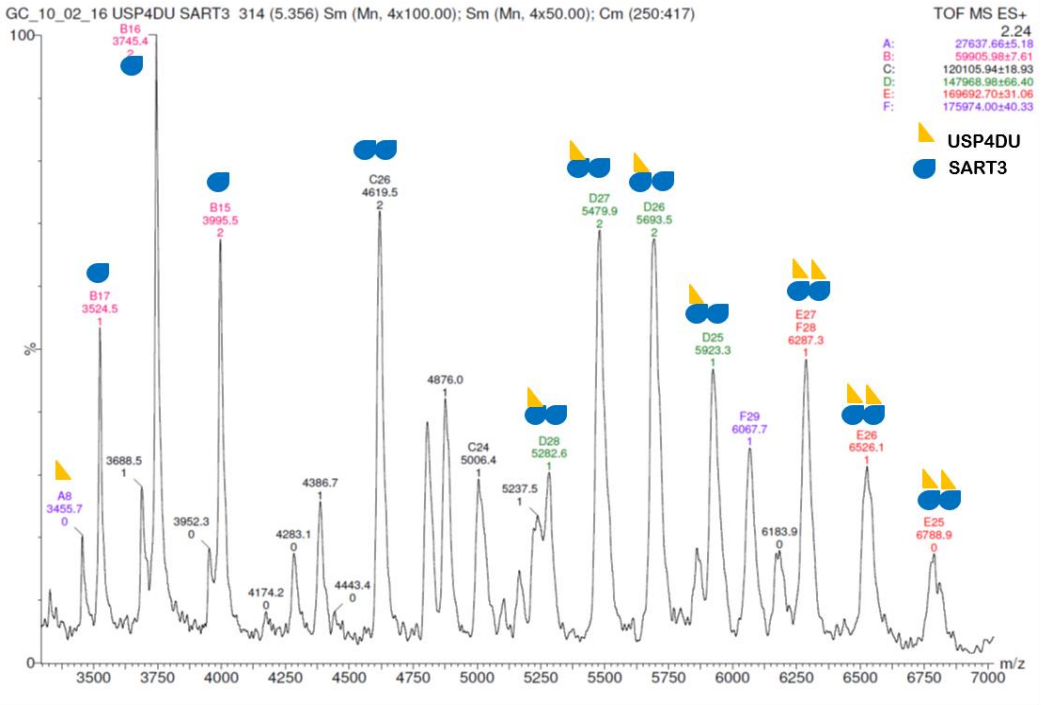


Figure 7.12. Mass spectrometry result of protein active site USP4DU-SART3 (the 2nd sample) The spectra showed the presence of dimer SART3, dimer SART3 binds with monomer USP4DU, and dimer SART3 binds with dimer USP4DU.

The third purification was done based on the Nickel column data, which resulted in two peaks. Each peak was then subjected to a separate gel filtration purification. In total, there were four samples collected. The first sample gave

the spectra contains; monomer USP4DU, monomer SART3, dimer SART3, dimer SART3 bind to monomer USP4DU, and dimer SART3 bind to dimer USP4DU (Figure 7.13).

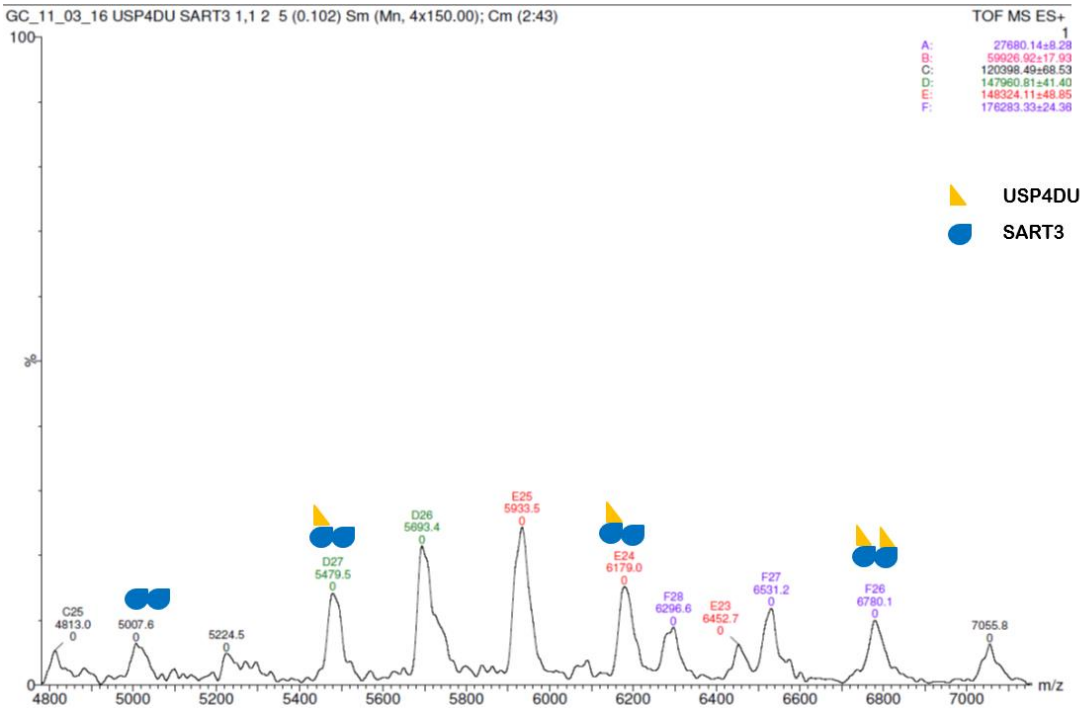


Figure 7.13. Mass spectrometry result of protein active site USP4DU-SART3 (3rd sample-1) The spectra showed the presence of dimer SART3, dimer SART3 binds with monomer USP4DU, and dimer SART3 binds with dimer USP4DU.

The second sample only contains monomer SART3 and USPDU (Figure 7.14). The third sample contains monomer USP4DU, dimer SART3, dimer SART3 binds to dimer USP4DU (Figure 7.15). The spectra also detect monomer SART3, complex USP4DU-SART3, and dimer SART3 binds to monomer USP4DU.

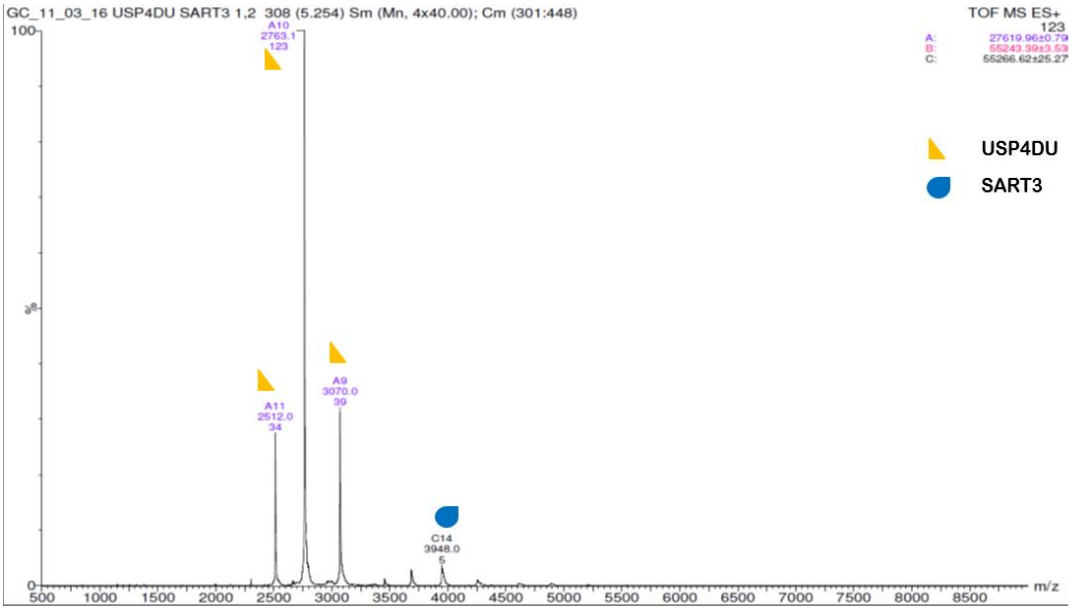


Figure 7.14. Mass spectrometry result of protein active site USP4DU-SART3 (3rd sample-2) The spectra showed the presence of monomer SART3 and monomer USP4DU only.

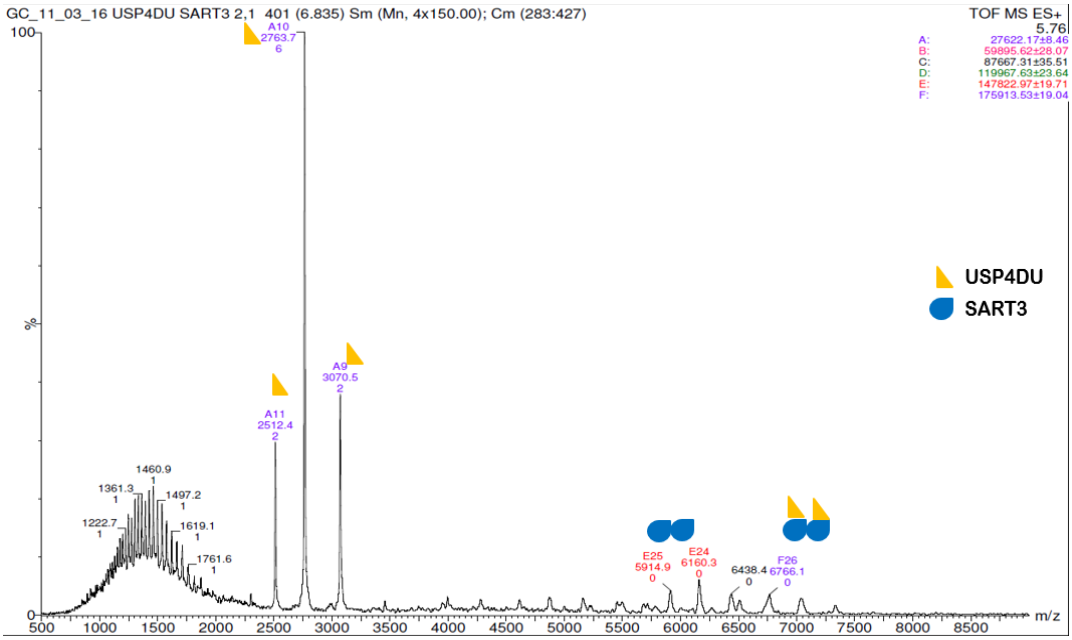


Figure 7.15. Mass spectrometry result of protein active site USP4DU-SART3 (3rd sample-3) The spectra showed the presence of dimer SART3 bind to dimer USP4DU, dimer SART3, and monomer USP4DU.

The fourth sample showed the spectra of monomer USPDU, monomer SART3 and complex monomer USP4DU with monomer SART3 (Figure 7.16; Table 7.7).

The interaction molecules of USP4DU with SART3 in this sample was already found on the first purification.

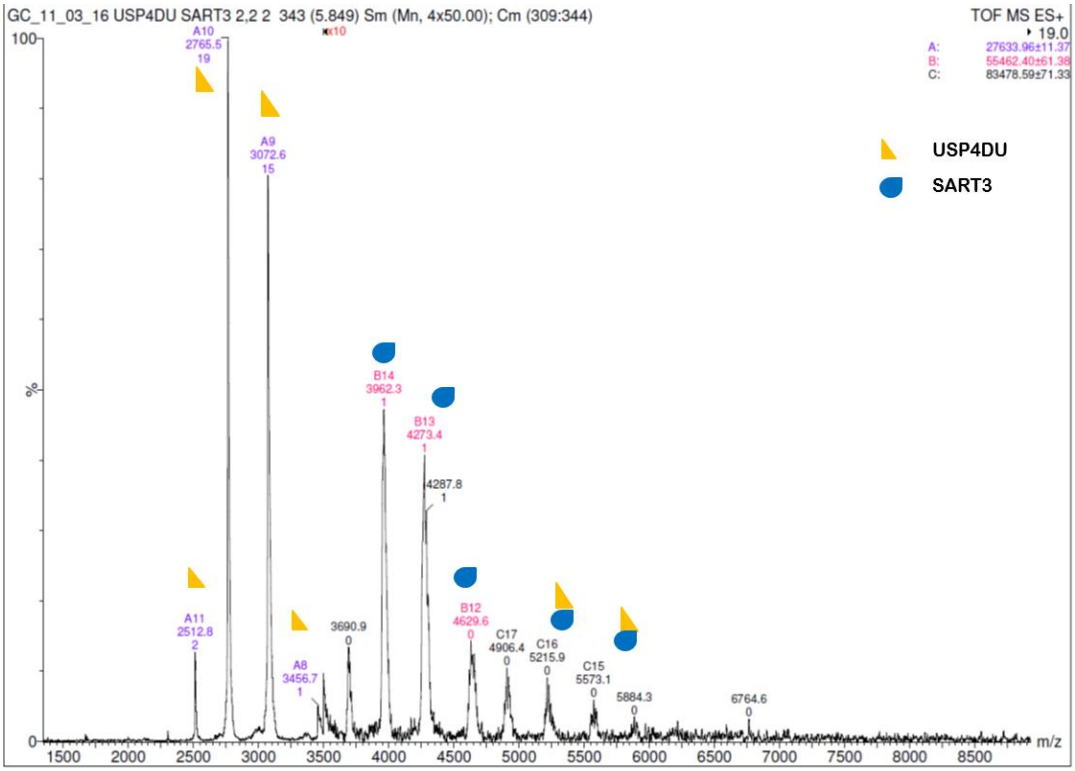


Figure 7.16. Mass spectrometry result of protein active site USP4DU-SART3 (3rd sample-4) The spectra showed the presence of monomer SART, monomer USP4DU, and complex monomer SART3-USP4DU

Most of the complex showed the presence of dimer SART3, both in unbound formation or bound with monomer USP4DU and dimer USP4DU. Park et al., (2016) showed that USP4DU and SART3 bind in homodimer formation.

Table 7.7. The summary of mass spectrometry analysis of complex USP4DU (1-244) –SART3 (96-574)

Protein or Complex Protein	MW based on Protparam (Da)	MW based on MS result (Da)	Prediction mass	Prediction of protein and/or interaction	Purification / Sample preparation
USP4DU-SART3	29222.8 (USP4DU) 59953.0 (SART3)	27641.39 ± 3.75	28141.6	USP4DU without His Tag	GF: 300 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		55299.43 ± 3.19	56482.2	SART3 without His Tag	
		59904.89 ±7.06	59953.0	SART3	
		87583.22 ± 2.04	89175.8	USP4DU-SART3 (1:1)	
		147596.95 ± 28.79	149128.8	SART3- SART3-USP4DU (2 : 1)	
		175390.06 ± 57.26	178351.6	USP4DU-USP4DU-SART3-SART3 (2 : 2)	
USP4DU-SART3	29222.8 (USP4DU) 59953.0 (SART3)	27637.66 ± 5.18	28141.6	USP4DU without His Tag	GF: 100 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		59905.98 ± 7.61	59953.0	SART3	
		120105.94 ± 18.93	119906	SART3-SART3 (dimer)	
		147968.98 ± 66.40	149128.8	SART3-SART3 – USP4DU (2:1)	
		169692.70 ± 31.06	171410	SART3-SART3 – USP4DU-USP4DU (2:2)	
		175974.00 ± 40.33	178351.6	SART3-SART3- USP4DU-USP4DU (2:2)	
USP4DU-SART3	29222.8 (USP4DU) 59953.0 (SART3)	27680.14 ± 8.28	28141.6	USP4DU without His Tag (28141.6)	GF: 100 mM NaCl, 1 % Glycerol, 50 mM Tris pH 7.5 Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		599926.92 ± 17.93	59953.0	SART3	
		120398.49 ± 68.53	119906	SART3-SART3 (dimer)	
		147960.81 ± 41.40	149128.8	SART3- SART3 – USP4DU (2:1)	
		176283.33 ± 24.36	178351.6	SART3- SART3 - USP4DU –USP4DU (2:2)	
		27619.96 ± 0.79	28141.6	USP4DU without His Tag	Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		55243.39 ± 3.53	56482.2	SART3 without His Tag	
		55266.62 ± 25.27	56482.2	SART3 without His Tag	
		27622.17 ± 8.46	28141.6	USP4DU without His Tag	Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		59895.62 ± 28.07	59953.0	SART3	
		87667.31 ± 35.51	89175.8	USP4DU - SART3 (1:1)	
		119967.63 ± 23.64	119906	SART3- SART3 (dimer)	
		147822.97 ± 19.71	149128.8	SART3-SART3 - USP4DU (2:1)	
		175913.53 ± 19.04	178351.6	SART3- SART3 - USP4DU-USP4DU (2:2)	
		27633.96 ±11.37	28141.6	USP4DU without His Tag	Desalting by dialysis with 150 mM Ammonium acetate pH 6.8
		55462.40 ± 61.38	56482.2	SART3 without His Tag	
		83478.59 ± 71.33	84623.8	USP4DU - SART3 (1:1) without His Tag	

7.6. Discussion

The binding interaction between various active site mutant USP4 and Ubiquitin variant using mass spectrometry and ITC analysis confirmed how the molecules bind each other. The USP4FL, in active or mutant construct, was a large protein molecule which needs optimisation in the mass spectrometry analysis to prevent the degradation. ITC analysis showed that USP4FL-C311S has tightest binding interaction with diubiquitin.

In the binding activity of USP4C1C2-C311S with Ubiquitin variant, this construct has slightly increased the affinity with Ub-GGG rather than Ubiquitin. On the other hand, in the diubiquitin construct resulted in a weaker interaction. These data supported by mass spectrometry analysis that the complex of USP4C1C2-H881N with Ub-GGG has more probability, about 70%, compared to the complex with diubiquitin-L73X which only 51%.

Moreover, the other construct showed that the USP4ΔDU-C311S tends to bind with diubiquitin rather than ubiquitin (refer to K_D value). The ESI-mass spectrometry data only showed the interaction of complex USP4ΔDU-H881N with ubiquitin. In the USP4ΔDU construct, which has insert region, mass spectrometry data showed there were no unbound USP4ΔDU-H881N. All USP4ΔDU-H881N bind to Ubiquitin. This mass spectrometry data supports the role of the insert region on increasing Ubiquitin-binding as Clerici et al., (2014)

was revealed it. In addition, for comprehensive result, binding analysis with diubiquitin using ESI-MS was needed to confirm the interaction of USP4ΔDU.

Interaction between USP4DU and SART3, the data showed that the complex could be consists of monomer USP4DU, monomer SART3, dimer SART3, complex monomer USP4DU with monomer SART3, complex dimer SART3 with monomer USP4DU and dimer SART3 and dimer USP4DU.

The investigation of interaction was also continued by crystallisation trial. Unfortunately, the complex USP4DU and SART3 on various screen, which incubated at 10°C, did not show the crystal formation. The heterogeneity of the samples could have been an issue. Subsequently a crystal structure of SART3 in complex with USP15DU and USP4 was published in 2016 (Zhang, et al., 2016; Park, et al., 2016).

Chapter 8

Discussion and Future Outlook

8.1. Expression and Purification of USP4

The various constructs of USP4, namely USP4FL, USP4htt, USP4httΔUbl2, USP4C1C2, USP4ΔDU and USP4DU were expressed and purified well. Only for USP4htt and USP4ΔDU a lower yield was obtained. The yield of USP4htt and USP4ΔDU was 14-fold and almost 8-fold lower respectively than USP4FL. Interestingly, USP4ΔDU showed a dimer in solution when gel filtration purification was performed. Moreover, the specific enzymatic activity of various active USP4 constructs showed that the constructs USP4htt had the lowest activity. USP4htt specific activity has less 125 times than USP4C1C2 and less 25 times than USP4FL, which suggest they are probably less stable overall.

These data indicate the importance of the DUSP-Ubl domains for protein expression and enzymatic activity of USP4. Interestingly, the USP4httΔUbl2, which also lacks the DUSP-Ubl domains, has a high yield was obtained.

8.2. Coexpression of Active Site Mutants of USP4 with Ubiquitin Variants and Crystallisation trials

Based on the coexpression experiments between active site mutants of various USP4 constructs with ubiquitin variants, the best result was obtained with the protein complex of active site mutant USP4FL-H881N with diubiquitin. The crystal had grown after around three months and diffracted to 3 Å, although the interpretation of the dataset was difficult and molecular replacement has

not yet provided a solution to solve the structure of USP4. Optimisation of the crystallisation conditions unfortunately did not reproduce the crystals.

The ubiquitin Ub-GGG variant in complex with USP4C1C2-H881N and USP4htt-H106N indicate preliminary crystal growth but no diffraction quality crystals were obtained. A longer experiment and optimisation will be needed to form crystals.

8.3. Characterisation Binding Interactions of USP4 with Ubiquitin variants

The binding interaction between active site mutant USP4FL-C311S, USP4C1C2-C311S and USP4ΔDU-C311S with Ubiquitin, Ub-GGG and diubiquitin showed interesting pattern. USP4FL-C311S has the tightest affinity with diubiquitin then Ub-GGG and the lowest was Ubiquitin. In USP4FL construct, additional ubiquitin (to be linear diubiquitin) increases the affinity 40-folded higher, whereas additional Glycine residue, in the Ub-GGG, only increase 4 times compare to the Ubiquitin affinity itself.

The construct of USP4ΔDU, has the same type of Ubiquitin variant affinity with USP4FL. Affinity of USP4ΔDU with diubiquitin was the tightest and the lowest was with Ubiquitin. Truncation of DUSP-Ubl domain in the construct USP4ΔDU has not affected the interaction binding with Ubiquitin variant. Diubiquitin and Ub-GGG affinity has 15 and 2.5 times higher respectively compare to Ubiquitin.

The catalytic core domain USP4C1C2-C311S showed the different pattern. The affinity with diubiquitin was the lowest. The value was 42-fold and 65-fold lower than Ubiquitin and Ub-GGG respectively. In the construct without insert and DUSP-Ubl domain, additional Glycine residue (in Ub-GGG) possibly made fit in the binding site USP4C1C2-C311S so can increase the affinity interaction, however the additional Ubiquitin construct (in diubiquitin) made the flexible domain in the binding site then decrease the binding affinity.

Another comparison showed Ubiquitin and Ub-GGG has the tightest affinity to USP4C1C2-C311S construct rather than USP4FL-C311S and USP4ΔDU-C311S. The Ubiquitin binding affinity was 6.5-fold and 14-fold lower, more over the Ub-GGG affinity was 2.5-fold and 9-fold lower respectively. This interaction showed that without DUSP-Ubl and insert domain, Ubiquitin and Ub-GGG bind properly in the binding site of USP4C1C2-C311S. Additional insert and DUSP-Ubl possibly change the conformation of binding site then inhibit the interaction and decrease the affinity (Figure 8.1).

In the diubiquitin interaction, USP4FL-C311S has the tightest binding followed by USP4ΔDU-C311S and USP4C1C2-C311S. The affinity of diubiquitin to USP4ΔDU-C311S was 6 times lower and significantly decrease more than 250 times to USP4C1C2. These data indicated that without DUSP-Ubl domain, the diubiquitin still able to bind in the binding site of USP4ΔDU-C311S, the insert domain might be having a role to stabilise the interaction. But without DUSP-

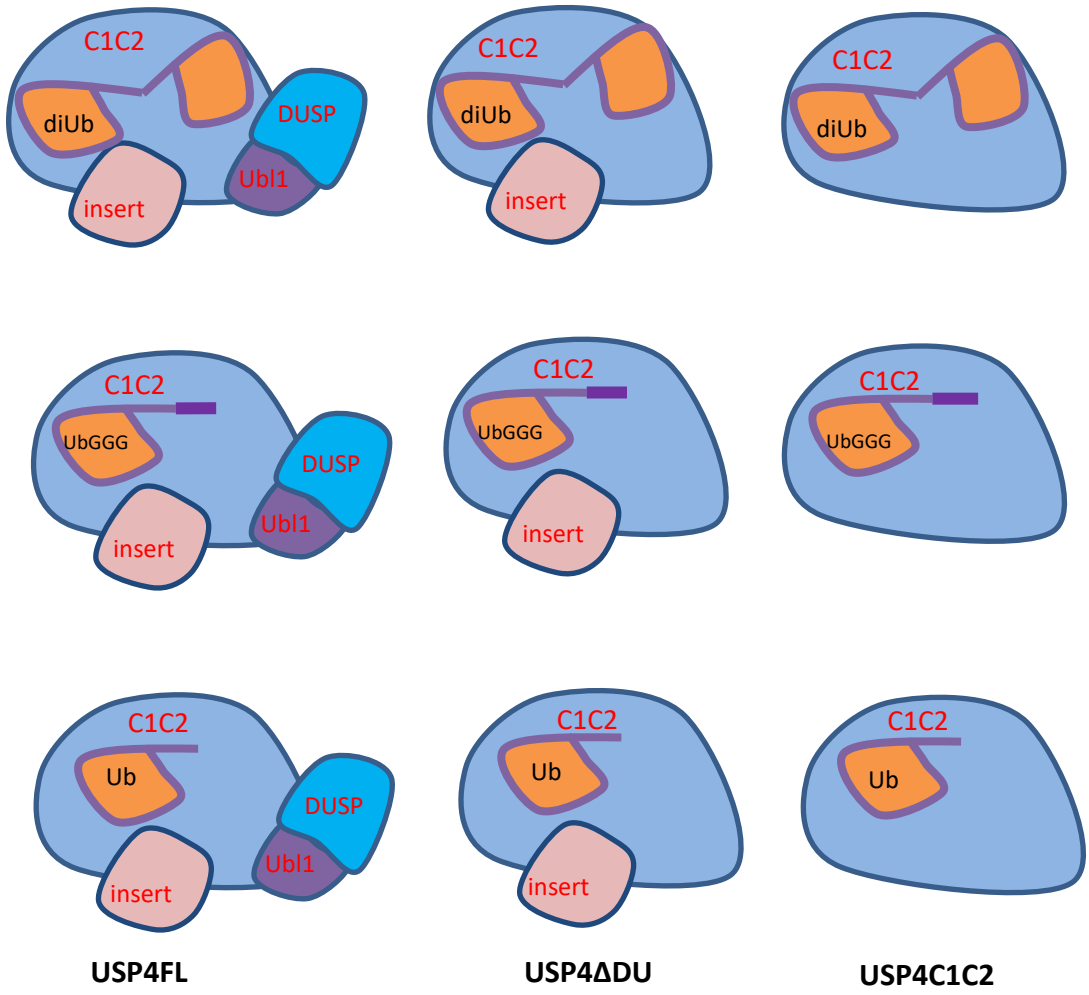


Figure 8.1. Schematic diagram showing the interaction between various USP4 with Ubiquitin variant

Ubl and insert region, the conformational interaction of diubiquitin in the binding site of USP4C1C2-C311S possibly more difficult so decrease the affinity sharply.

8.4. Future Outlook

Further experiments are required to reproduce the crystal of USP4FL-H881N in complex with diubiquitin and USP4FL-C311S with diubiquitin. Molecular replacement analysis still needed to be done to solve the structure and gain insights into the interaction with high affinity. ITC data showed that active site mutant USP4FL-C311S binds to diubiquitin.

Further ESI-mass spectrometry analysis for various USP4-C311S constructs with Ubiquitin variants (Ubiquitin, diubiquitin and Ub-GGG) should be performed to characterise the complexes and relate the result to the ITC data.

Inhibition activity assays and crystallisation trials using inhibitor agents, such as mitoxantrone should be continued to unravel the USP4 inhibitor interactions. The result can be compared to the paralogues USP11 and USP15.

References

ABDUL-REHMAN, S.A. KRISTARIYANTO, Y.A. CHOI, S.Y. NKOSI, P.J. WEIDLICH, S. LABIB, K. HOFMANN, K. & KULATHU, Y. 2016. MINDY-1 is a member of an evolutionarily conserved and structurally distinct new family of deubiquitinating enzymes. *Mol Cell*, 63(1): 146-55

AKASHI, S. 2006. Investigation of molecular interaction within biological macromolecular complexes by mass spectrometry. *Med Res Rev,* 26**,** 339-68.

ANDERSEN, K.R. LEKSA, N.C. & SCHWARTZ, T.U. 2013. Optimized E.coli expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins*, 81 (11): 1857-61

BALDWIN, M. A. 2004. Protein identification by mass spectrometry: issues to be considered. *Mol Cell Proteomics,* 3**,** 1-9.

BARUCH, A., JEFFERY, D. A. & BOGYO, M. 2004. Enzyme activity – it's all about image. *Trends in Cell Biology,* 14**,** 29-35.

BEN-NISAN, G. & SHARON, 2018. The application of ion-mobility mass spectrometry for structure/function investigation of protein complexes. *Curr Opin Chem Biol,* 42: 25-33.

BERTHOUZE, M., VENKATARAMANAN, V., LI, Y. & SHENOY, S. K. 2009. The deubiquitinases USP33 and USP20 coordinate beta2 adrenergic receptor recycling and resensitization. *EMBO J,* 28**,** 1684-96.

BLOW, D. 2002. *Outline of Crystallography for Biologists*. OUP Oxford, 146-149

BRADER, M. L., BAKER, E. N., DUNN, M. F., LAUE, T. M. & CARPENTER, J. F. 2017. Using X-ray Crystallography to Simplify and Accelerate Biologics Drug Development. *J Pharm Sci,* 106**,** 477-494.

CATANZARITI, A.M. SOBOLEVA, T.A. JANS, D.A. BOARD, P.G. & BAKER, R.T. 2004. An efficient syatem for high-level expression and easy purification of authentic recombinant proteins. *Protein Sci,* 13 (5): 1331-9

CHAYEN, N. E. 2005. Methods for separating nucleation and growth in protein crystallisation. *Prog Biophys Mol Biol,* 88**,** 329-37.

CHEUNG, R.C. WONG, J.H. & NG, T.B. 2012. Immobilized metal ion affinity chromatography : a review on its applications. *Appl Microbiol Biotechnol,* 96 (6): 1411-20

CIECHANOVER, A., 2005. Proteolysis: from the lysosome to ubiquitin and proteasome. *Mol Cell Bio*, 6, 79-86

CLAGUE, M. J., BARSUKOV, I., COULSON, J. M., LIU, H., RIGDEN, D. J. & URBE, S. 2013. Deubiquitylases from genes to organism. *Physiol Rev,* 93**,** 1289-315.

CLAGUE, M. J., COULSON, J. M. & URBE, S. 2012. Cellular functions of the DUBs. *J Cell Sci,* 125**,** 277-86.

CLERICI, M., LUNA-VARGAS, M. P. A., FAESEN, A. C. & SIXMA, T. K. 2014. The DUSP–Ubl domain of USP4 enhances its catalytic efficiency by promoting ubiquitin exchange. *Nat Commun,* 5. 5399

CLEYNEN, I., JUNI, P., BEKKERING, G. E., NUESCH, E., MENDES, C. T., SCHMIED, S., WYDER, S., KELLEN, E., VILLIGER, P. M., RUTGEERTS, P., VERMEIRE, S. & LOTTAZ, D. 2011. Genetic evidence supporting the association of protease and protease inhibitor genes with inflammatory bowel disease: a systematic review. *PLoS One,* 6**,** e24106.

CUBRILOVIC, D. 2014. Native mass spectrometry: A powerful tool to study protein-ligand complexes in drug discovery. *Thesis,* 29

DANG, L.C. MELANDRI, F.D. & STEIN, R.L. 1998. Kinetic and mechanistic studies on the hydrolysis of ubiquitin C-terminal 7-amido-4-methylcoumarin by deubiquitinating enzymes. *Biochemistry,* 37 (7): 1868-79

DESSAU, M. A. & MODIS, Y. 2011. Protein crystallization for X-ray crystallography. *J Vis Exp*, 16; (47). pii: 2285

DIKIC, I., WAKATSUKI, S. & WALTERS, K. J. 2009. Ubiquitin-binding domains - from structures to functions. *Nat Rev Mol Cell Biol,* 10**,** 659-71.

DOMON, B. & AEBERSOLD, R. 2006. Mass Spectrometry and Protein Analysis: Review. *Science*, 312 (5771). 212-217.

DU, X., LI, Y., XIA, Y. L., AI, S. M., LIANG, J., SANG, P., JI, X. L. & LIU, S. Q. 2016. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci,* 17.

ELLIOTT, P. R., LIU, H., PASTOK, M. W., GROSSMANN, G. J., RIGDEN, D. J., CLAGUE, M. J., URBE, S. & BARSUKOV, I. L. 2011. Structural variability of the ubiquitin specific protease DUSP-UBL double domains. *FEBS Lett,* 585**,** 3385-90.

FREYER, M.W. & LEWIS, E.A. 2008. Isothermal Titration Calorymetri : Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. *Methods in Cell Biol*, 84: 79-113

GARMAN, E. F. & OWEN, R. L. 2006. Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr,* 62**,** 32-47.

GAVIRA, J. A. 2016. Current trends in protein crystallization. *Arch Biochem Biophys,* 602**,** 3-11.

GILCHRIST, C. & BAKER, R. 2000. Characterization of the ubiquitin-specific protease activity of the mouse/human Unp/Unph oncoprotein. *Biochimica et Biophysica Acta,* 1481**,** 297 - 309.

GOLDBERG, A. L. 2005. Nobel committee tags ubiquitin for distinction. *Neuron,* 45**,** 339-44.

GRAY, D. A., INAZAWA, J., GUPTA, K., WONG, A., UEDA, R. & TAKAHASHI, T. 1995. Elevated expression of Unph, a proto-oncogene at 3p21.3, in human lung tumors. *Oncogene,* 10**,** 2179-83.

GRAZETTE, A. 2015. Molecular characterisation of squamous cell carcinoma antigen recognised by T-Cells 3, An Adaptor Protein of Ubiquitin Specific Protease 15. *Thesis,*

GU, J. SHIMBA, S. NOMURA, N. & REDDY, R. 1998. Isolation and characterisation of a new 110 kDa human nuclear RNA binding protein 9p110nrb). *Biochim Biophys Acta,* 1399 (1):1-9

GUPTA, K. COPELAND, NG. GILBERT, DJ. JENKINS, NA. & GRAY, DA. 1993. Unp, a mouse gene related to the tre oncogene. *Oncogene,* 8(8): 2307-10

HARADA, K. YAMADA, A. YANG, D. ITOH, K. & SHICHIJO, S. 2001. Binding of a SART3 tumor-rejection antigen to a pre-mRNA splicing factor RNPS1: a possible regulation of splicing by a complex formation. *Int J Cancer,* 93(5):623-8

HARPER, S., BESONG, T. M., EMSLEY, J., SCOTT, D. J. & DREVENY, I. 2011. Structure of the USP15 N-terminal domains: a beta-hairpin mediates close association between the DUSP and UBL domains. *Biochemistry,* 50**,** 7995-8004.

HARPER, S., GRATTON, H. E., CORNACIU, I., OBERER, M., SCOTT, D. J., EMSLEY, J. & DREVENY, I. 2014. Structure and catalytic regulatory function of ubiquitin specific protease 11 N-terminal and ubiquitin-like domains. *Biochemistry,* 53**,** 2966-78.

HECK, A. J. 2008. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat Methods,* 5**,** 927-33.

HEDLUND, H. 2006. Preparative size-exclusion chromatography. *CSH Protoc,* 2006 (1)

HENDRICKSON, W.A., 2014, Anomalous diffraction in crystallographic phase evaluation. *Q Rev Biophys,* 47 (1): 49-93

HERSHKO, A. & CIECHANOVER, A. 1998. The Ubiquitin System. *Annual Review of Biochemistry,* 67**,** 425-479.

HICKE, L., SCHUBERT, H. L. & HILL, C. P. 2005. Ubiquitin-binding domains. *Nat Rev Mol Cell Biol,* 6**,** 610-21.

HONG, P. KOZA, S. & BOUVIER, E.S. 2012. Size Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *J Liq Chromatgr Relat Technol,* 35 (20) 2923-2950

HOSPENTHAL, M. K., FREUND, S. M. V. & KOMANDER, D. 2013. Assembly, analysis and architecture of atypical ubiquitin chains. *Nat Struct Mol Biol,* 20**,** 555-565.

HOU, X., WANG, L., ZHANG, L., PAN, X. & ZHAO, W. 2013. Ubiquitin-specific protease 4 promotes TNF-alpha-induced apoptosis by deubiquitination of RIP1 in head and neck squamous cell carcinoma. *FEBS Lett,* 587**,** 311-6.

HUANG, O. W. & COCHRAN, A. G. 2013. Regulation of deubiquitinase proteolytic activity. *Curr Opin Struct Biol,* 23**,** 806-11.

HUSNJAK, K & DIKIC,L. 2012. Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu Rev Biochem,* 81:291-322

IRVINE, G. B. 1997. Molecular weight estimation for neuropeptides using size exclusion high performance liquid chromatography. *Methods Mol Biol,* 73: 89-95

ILARI, A. & SAVINO, C. 2017. A Practical Approach to Protein Crystallography. *Methods Mol Biol,* 1525**,** 47-78.

JIANG, W. & ROBINSON, R.A.S. 2013. Ion Mobility-Mass Spectrometry. *Encyclopedia of Analytical Chemistry*. John Wiley & Sons. 1-21.

KANTARDJIEFF, K.A. & RUPP, B. 2003. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci,* 12 (9): 1865-71

KOMANDER, D. 2009. The emerging complexity of protein ubiquitination. *Biochem Soc Trans,* 37**,** 937-53.

KOMANDER, D., CLAGUE, M. J. & URBE, S. 2009. Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol,* 10**,** 550-63.

KOMANDER, D. & RAPE, M. 2012. The ubiquitin code. *Annu Rev Biochem,* 81**,** 203-29.

KONIJNENBERG, A. BUTTERER, A. SOBBOT, F. 2013. Review: Native ion mobility-mass spectrometry and related methods in structural biology. *Biochim Biophys Acta*, 1834, 1293-1256.

LANUCARA, F. HOLMAN, S.W. GRAY, C.J. & EYERS, C.E. 2014. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics, Nat Chem, 6, 281-294

LENEY, A. C. & HECK, A. J. 2017. Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectrom,* 28**,** 5-13.

LI,W. & YE,Y. 2008. Polyubiquitin chains: functions, structures, and mechanisms. *Cell Moll Life Sci*, 65 (15): 2397-406

LI, H. AI, L. FAN, S. WANG, Y. & SUN, D. 2017. Rapid determination of 18 glucocorticoids in serum using reusable on-line SPE polymeric monolithic column coupled with LC quadrupole/orbitrap high resolution mass spectrometer. *J Chromator B Analyt Technol Biomed Life Sci*, 1065-1066: 79-86

LIN, R. NIE, J. REN, J. LIANG, R. LI, D. WANG, P. GAO, C. ZHUO, C. YANG, C. & LI, B. 2017. USP4 interacts and possitively regulates IRF8 function via K48-linked deubiquitination in regulatory T cells. *FEBS Lett,* 591(12): 1677-1688

LONG, L. THELEN, J.P. FURGASON, M. HAJ-YAHYA, M. BRIK, A. CHENG, D. PENG, J. & YAO, T. 2014. The U4/U6 recycling factor SART3 has histone chaperone activity and associates with USP15 to regulate H2B deubiquitination. *J Biol Chem,* 289(13): 8916-30

LOSSL, P. van de WATERBEEMD, M. & HECK, A.J.R. 2016. The diverse and expanding role of mass spectrometry in structural and molecular biology. *EMBO*, 35: 2634-2657

McLAFFERTY, F.W. 2008. Mass spectrometry across the sciences. *PNAS*, 105, 18088 - 18089.

McPHERSON, A. 2004. Introduction to protein crystallization. *Methods,* 34**,** 254-65.

McPHERSON, A. & GAVIRA, J. A. 2014. Introduction to protein crystallization. *Acta Crystallogr F Struct Biol Commun,* 70**,** 2-20.

MILOJEVIC, T., REITERER, V., STEFAN, E., KORKHOV, V. M., DOROSTKAR, M. M., DUCZA, E., OGRIS, E., BOEHM, S., FREISSMUTH, M. & NANOFF, C. 2006. The ubiquitin-specific protease Usp4 regulates the cell surface level of the A2A receptor. *Mol Pharmacol,* 69**,** 1083-94.

MIRZA, S. P. & OLIVIER, M. 2008. Methods and approaches for the comprehensive characterization and quantification of cellular proteomes using mass spectrometry. *Physiol Genomics,* 33**,** 3-11.

NATH, D. & SHADAN, S. 2009. The ubiquitin system, *Nature*, 458 (7327): 421

NIEDZIALKOWSKA, E. GASIOROWSKA, O. HANDLING, K.B. MAJOREK, K.A. POREBSKI, P.J. SHABALIN, I.G. ZASADZINSKA, E. CYMBOROWSKI, M. & MINOR, W. 2016. Protein purification and crystallization artifacts: The tale usually not told. *Protein Sci*, 25 (3): 720-33

NIJMAN, S. M. LUNA-VARGAS, M. P. VELDS, A., BRUMMELKAMP, T. R. DIRAC, A. M. SIXMA, T. K. & BERNARDS, R. 2005. A genomic and functional inventory of deubiquitinating enzymes. *Cell,* 123**,** 773-86.

PAL, A., YOUNG, M.A., & DONATO, N.J. 2014. Emerging potential of Therapeutic Targeting of Ubiquitin-Specific Proteases in the Treatment of Cancer. *Cancer Res*, 74 (18), 4955-4966

PARK, J.K. DAS, T. SONG, E.J. & KIM, E.E. 2016. Structural basis for recruiting and shuttling of spliceosomal deubiquitinase USP4 by SART3. *Nucleic Acids Res*, 44(11): 5424-37.

PERRETT, C.A. LIN, D.Y. & ZHOU, D. 2011. Interactions of bacterial proteins with host eukaryotic ubiquitin pathways. *Front Microbiol*, 2 (143): 1-18.

PFAUNMILLER, EL. PAULEMOND, M.L. DUPPER, CM. HAGE, DS. 2013. Affinity monolith chromatography : a review of principles and recent analytical applications. *Anal Bioanal Chem*, 405 (7) 2133-45.

PICKART, C. M. & EDDINS, M. J. 2004. Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta,* 1695**,** 55-72.

PICKERING, A.M. DAVIES, K.J. 2012. A simple fluorescence labeling method for studies of protein oxydation, protein modification, and proteolysis. *Free Radic Biol Med*, 52(2): 239-46

PIERCE, M.M. RAMAN, C.S. & NALL, B.T. 1999. Isothermal Titration Calorimetry of Protein-Protein Interactions. *Methods*, 19: 213-221.

PIKE, A.C.W. GARMAN, E.F. KROJER, T. von DELFT, F. & CARPENTER, E.P., 2016, An overview of heavy-atom derivatization of protein crystals. *Acta Cryst,* D72, 303-318.

PORATH, J. CARLSSON, J. OLSSON, I. & BELFRAGE, G. 1975. Metal chelate affinity chromatography, a new approach to protein fractination. *Nature*, 258 (5536): 598-9

PyMOL, The PyMOL Molecular Graphis System, version 1.3 Schrodinger, LLC

RAHIGHI, S. & DIKIC,I. 2012. Selectivity of the ubiquitin-binding modules. *FEBS Lett*, 586 (17): 2705-10

REYES-TURCU, F. E., VENTII, K. H. & WILKINSON, K. D. 2009. Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu Rev Biochem,* 78**,** 363-97.

REYMOND, J. L., FLUXA, V. S. & MAILLARD, N. 2009. Enzyme assays. *Chem Commun (Camb)***,** 34-46.

RHODES, G. 2006. *Crystallography made crystal clear*. 3rd Ed. Academia Press. London. 37 - 46

RITTINGER, K. & IKEDA, F. 2017. Linear ubiquitin chains: enzymes, mechanisms and biology. *Open Biol,* 7.

ROWLETT, R.S. 2005. *Protein X-ray Crystallography Methods*. 2nd Ed. Department of Chemistry, Colgate University. 3-6

ROSSI, A. M. & TAYLOR, C. W. 2011. Analysis of protein-ligand interactions by fluorescence polarization. *Nat Protoc,* 6**,** 365-87.

RUSSEL, N.S. & WILKINSON, K.D. 2005. Deubiquitinating enzyme purification, assay inhibitors, and characterization. *Methods Mol Biol,* 301: 207-19

SIPPL, W., COLLURA, V., & COLLAND, F., 2011. Ubiquitin-specific proteases as cancer drug targets. *Future Oncol*, 7(5), 619-632.

SMYTH, M.S. & MARTIN, J.H.J. 2000. x Ray Crystallography. *J Clin Pathol: Mol Pathol,* 53: 8–14

SONG, E. J., WERNER, S. L., NEUBAUER, J., STEGMEIER, F., ASPDEN, J., RIO, D., HARPER, J. W., ELLEDGE, S. J., KIRSCHNER, M. W. & RAPE, M. 2010. The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev,* 24**,** 1434-47.

SUN, L., CHEN, Z.J. 2004. The novel function of ubiquitination in signalling. *Curr Opin Cell Biol*, 16: 119 - 126

TAYYAB, S. QAMAR, S. & ISLAM, M. 1991. Size exclusion chromatography and Size Exclusion HPLC of Proteins. *Biochem Educ*, 19 (3)

TOMLINSON, E., PALANIYAPPAN, N., TOOTH, D. & LAYFIELD, R. 2007. Methods for the purification of ubiquitinated Proteins. *Proteomics,* 7**,** 1016-1022.

URAS, I. Z., LIST, T. & NIJMAN, S. M. 2012. Ubiquitin-specific protease 4 inhibits mono-ubiquitination of the master growth factor signaling kinase PDK1. *PLoS One,* 7**,** e31003.

URH, M., SIMPSON, D. & ZHAO, K. 2009. Chapter 26: *Affinity Chromatography*, 463**,** 417-438.

VARSHAVSKY, A. 2012. The Ubiquitin System, an Immense Realm. *Annu Rev Biochem,* 81**,** 167-176.

VENTII, K. H. & WILKINSON, K. D. 2008. Protein partners of deubiquitinating enzymes. *Biochem J,* 414**,** 161-75.

VLASSCHAERT, C. XIA ,X. COULOMBE, J. GRAY,DA. 2015. Evolution of the highly networked deubiquitinating enzymes USP4, USP15, and USP11. *BMC Evol Biol*, 15:230

WADA, K. & KAMITANI, T. 2006. UnpEL/Usp4 is ubiquitinated by Ro52 and deubiquitinated by itself. *Biochem Biophys Res Commun,* 342**,** 253-8.

WADA, K., TANJI, K. & KAMITANI, T. 2006. Oncogenic protein UnpEL/Usp4 deubiquitinates Ro52 by its isopeptidase activity. *Biochem Biophys Res Commun,* 339**,** 731-6.

WANG, G., GAO, Y., LI, L., JIN, G., CAI, Z., CHAO, J. I. & LIN, H. K. 2012. K63-linked ubiquitination in kinase activation and cancer. *Front Oncol,* 2**,** 5.

WANG, L., ZHAO, W., ZHANG, M., WANG, P., ZHAO, K., ZHAO, X., YANG, S. & GAO, C. 2013. USP4 positively regulates RIG-I-mediated antiviral response through deubiquitination and stabilization of RIG-I. *J Virol,* 87**,** 4507-15.

WARRENS, A.N. JONES, M.D. & LECHLER, R.I. 1997. Splicing by overlap extension by PCR using asymmetric amplification: an improved technique for the generation of hybrid proteins of immunological interest. *Gene*, 186 (1): 29-35

WELCHMAN, R. L., GORDON, C. & MAYER, R. J. 2005. Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat Rev Mol Cell Biol,* 6**,** 599-609.

WIJNHOVEN, P. KONIETZNY, R. BLACKFORD, A.N. TRAVERS, B.J. KESSLER, B.M. NISHI, R. & JACKSON, S.P. USP4 Auto-Deubiquitylation Promotes Homologous Recombination. *Mol Cell*, 60, 362-373

WLODAWER, A. LI, M. & DAUTER, Z. 2017. High-Resolution Cryo-EM Maps and Models: A Crystallographer's Perspective. *Structure*, 25, 1589–1597

WOOD, D. W. 2014. New trends and affinity tag designs for recombinant protein purification. *Curr Opin Struct Biol,* 26**,** 54-61.

XIAO, N., LI, H., LUO, J., WANG, R., CHEN, H., CHEN, J. & WANG, P. 2012. Ubiquitin-specific protease 4 (USP4) targets TRAF2 and TRAF6 for deubiquitination and inhibits TNFalpha-induced cancer cell migration. *Biochem J,* 441**,** 979-86.

YAN, E. K., ZHANG, C. Y., HE, J. & YIN, D. C. 2016. An Overview of Hardware for Protein Crystallization in a Magnetic Field. *Int J Mol Sci,* 17.

YIN, S. T., HUANG, H., ZHANG, Y. H., ZHOU, Z. R., SONG, A. X., HONG, F. S. & HU, H. Y. 2011. A fluorescence assay for elucidating the substrate specificities of deubiquitinating enzymes. *Biochem Biophys Res Commun,* 416**,** 76-9.

ZHANG, L., ZHOU, F., DRABSCH, Y., GAO, R., SNAAR-JAGALSKA, B. E., MICKANIN, C., HUANG, H., SHEPPARD, K. A., PORTER, J. A., LU, C. X. & TEN DIJKE, P. 2012. USP4 is regulated by AKT phosphorylation and directly deubiquitylates TGF-beta type I receptor. *Nat Cell Biol,* 14**,** 717-26.

ZHANG, W., SULEA, T., TAO, L., CUI, Q., PURISIMA, E., VONGSAMPHANH, R., LACHANCE, P., LYTVYN, V., QI, H., LI, Y. & MENARD, R. 2011a. Contribution of Active site residues to substrate hydrolysis by USP2: Insight into Catalysis by Ubiquitin Specific Protease. *Biochemistry,* 50**,** 4775 - 4785.

ZHANG, X., BERGER, F. G., YANG, J. & LU, X. 2011b. USP4 inhibits p53 through deubiquitinating and stabilizing ARF-BP1. *EMBO J,* 30**,** 2177-89.

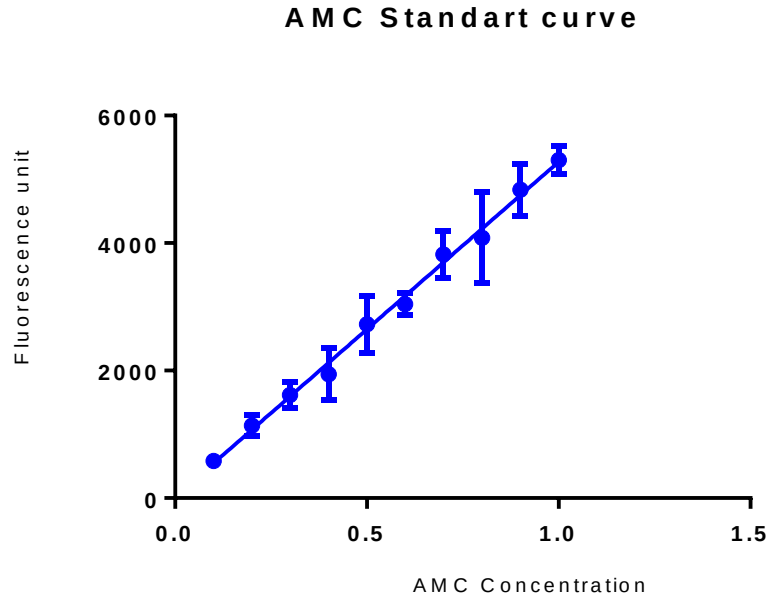
ZHAO, B., SCHLESIGER, C., MASUCCI, M. G. & LINDSTEN, K. 2009. The ubiquitin specific protease 4 (USP4) is a new player in the Wnt signalling pathway. *J. Cell. Mol. Med.,* 13**,** 1886 - 1895.

ZHAO, B., VELASCO, K., SOMPALLAE, R., PFIRRMANN, T., MASUCCI, M. G. & LINDSTEN, K. 2012. The ubiquitin specific protease-4 (USP4) interacts with the S9/Rpn6 subunit of the proteasome. *Biochem Biophys Res Commun,* 427**,** 490-6.

ZHOU, F. XIE, F. JIN, K. ZHANG, Z. CLERICI, M. GAO, R. VAN DINTHER, M. SIXMA, TK. HUANG, H. ZHANG, L. TEN DIJKE, P. 2017. USP4 inhibits SMAD4 monoubiquitination and promotes activin and BMP signaling. *EMBO J.*, 36(11): 1623-1639

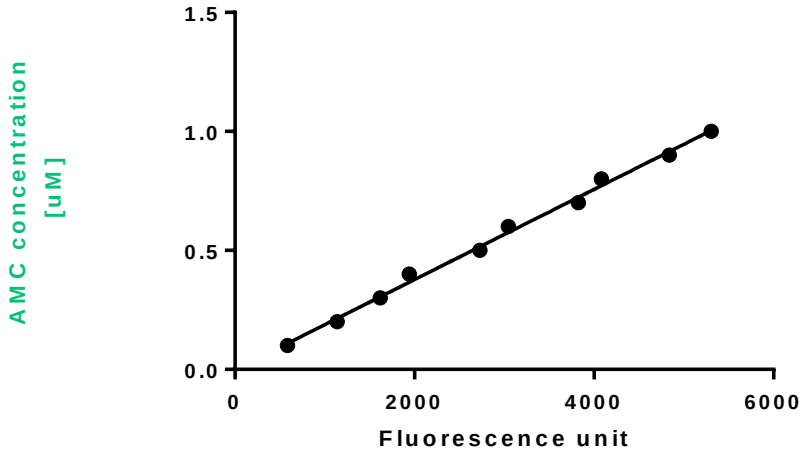
ZHU, X., MENARD, R. & SULEA, T. 2007. High incidence of ubiquitin-like domains in human ubiquitin-specific proteases. *Proteins,* 69**,** 1-7.

APPENDIX



$y=5253x+20.87$ ($R^2=0.96$)

Standard curve for enzymatic activity assay



$y=0.00019x - 0.0014$ ($R^2=0.99$)

Standard curve for kinetic parameter of USP4FL