
STRUCTURE FUNCTION ANALYSIS OF THE DEUBIQUITYLATING ENZYME FAM

Poon-Yu Khut, B.Sc

**School of Molecular & Biomedical Science (Biochemistry)
University of Adelaide
Adelaide, South Australia 5005**

October 2006

DECLARATION OF ORIGINALITY

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Poon-Yu Khut 04/10/06

ACKNOWLEDGMENTS

This thesis has no doubt been a painful experience for all concerned. To those who cry inside when they think of it, this thesis is dedicated to you.

First and foremost, many thanks are given to my endearing supervisor Stephen Wood for his humour, patience, insight and humility. It has been an unusual journey that was always met with your understanding. I am grateful that I was allowed the freedom to learn and explore both science and life (deep). To the members of the Wood lab that have assisted me over the years. The effervescent Michaela and omnipotent Susan have been steadfast in guiding me with hearts of gold. Susan incidentally performed the work to generate the Cys mutant that I later used to great effect. Mark's scepticism towards my scientific methods was probably founded in the end however Lachlan's disposition nurtured them. I know they are and will continue to flourish into incredibly astute scientists.

I am indebted to the Lardelli lab for their help and support with my work. It is a lab of remarkable camaraderie that was generously extended to me. Michael's expertise in analysing *usp9*'s expression pattern and frank guidance was invaluable, particularly in the writing of my first paper (to be submitted at the same time as this thesis). Svanhild has been motherly and taught me embryo manipulation and injection skills. Christine has always been happy to assist me and lend me her reagents, blindly randomised my injection samples, and embedded and sectioned embryos for me. Ben Tucker's speed and expertise in phylogenetic analysis and all things too genetic was greatly appreciated.

Maria Makrides from CHRI's Nutrition group guided me in the preparation of the phenotype data from the dominant-negative domain screen and performed the statistical analysis. Most importantly she told me what the numbers meant. Sassan Asgari is a true Iranian gentleman who is now a lecturer in the University of Brisbane, and saved me a hell of a lot of time and money in performing the baculoviral infections of the insect cells, subsequent preparation of viral stocks and harvesting of cell pellets. The boys at the IMVS protein lab Chris Bagley and Ian Milne were helpful with my ignorant questions regarding mass spectrometry and returned me a whole heap of numbers. Anne Chapman-Smith went

above the call of duty in the helping me to clean up an extremely dirty gel filtration column and the ill-fated filtration of FAM protein.

I must also thank my friends and family who both doubted me and supported me over these many years. My parents will be more relieved than I, thankyou for looking after me particularly when I faced the darkness of injury. Thankyou mum for running around to organise the final stages of this thesis while I have been in Japan. My siblings and my sister Chiew-Yee (Qiu Yi) in particular have been incredibly sympathetic and supportive. Of this I am most grateful. Chiew-Yee leaves Adelaide today for aid work in Vietnam after returning to us from aid work in East-Timor. Her generous nature is reflected in these exploits and is what was extended to me. To my Kung-fu group who have been a source of inspiration, I thank you for providing my mind a procrastination tool. I attribute my increased internal skills to the boredom my thesis provided. Special thanks to Derek, Zac, Jon, Benny, Justin (Kung-fu), Ming, Eric, Lap, Kev (Races), Tara, Lyndon, Gareth, Flem, Zkye, Nick, Jacob, Doolz, Eddy, Simon J, Toby, Ben, Dieter, Dion (High priests and Musicians) John Mansfield and Lucie (Eternal friends).

Overwhelming gratitude is extended to the people that aided my physical and spiritual rehabilitation in particular Rob Human, Garry Evans, Kingsley Wood and Nada. Lastly I thank the spirits and universe for guiding me through these difficult times and pushing me to do the soul searching and growing I needed. I would not have guessed that in injury and immobility would I gain true freedom. Thankyou for looking out for me then, now and in the future.

ABSTRACT

The ubiquitin pathway is a highly conserved post-translational modification system best characterised for its roles in protein degradation and intracellular trafficking and is involved in a diverse spectrum of cellular processes. Ubiquitylation is opposed by deubiquitylating enzymes (Dubs), and the ubiquitin specific peptidase (USP) class of Dubs remove ubiquitin from specific substrates, thereby affecting protein fate. USPs exhibit broad sequence diversity except over their catalytic cores and it has been suggested that this sequence variation constitutes their individual substrate-specific binding sites.

Fat Facets in Mouse (Fam) is a developmentally regulated USP whose function is crucial for mouse pre-implantation development. *Fam* is expressed in a complex fashion throughout development in a number of diverse tissue types and time points, well beyond its critical role in the early embryo. *Fam*'s orthologue in fly, *Fat Facets (faf)* is also developmentally regulated and is required for both drosophila eye and syncytial stage development.

Given the strengths of the zebrafish system as a developmental tool, the zebrafish orthologue of *Fam*, *usp9* was identified and found to be highly conserved. Analysis of its expression pattern found considerable overlap with the published mouse patterns. Given the similarities between the mouse and zebrafish systems, a series of cross-species experiments were conducted to determine whether exogenous expression of highly conserved regions of FAM, could cause dominant-negative phenotypes in developing zebrafish embryos.

Outside of the catalytic core, FAM's large N and C-terminal extensions consist of novel sequence bearing no similarity to any known domain. To delineate FAM domains, full-length FAM was expressed in insect cells and subjected to partial proteolysis. Combining this data with recent structural predictions and computer analyses of the FAM sequence, four FAM domains were characterised with the first domain containing three possible subdomains. The predominant helical nature of the N and C-terminal extensions of FAM were predicted to form scaffolding structures, well suited to protein binding.

TABLE OF CONTENTS

TITLE	I
DECLARATION OF ORIGINALITY	II
ACKNOWLEDGMENTS	III
ABSTRACT	V
TABLE OF CONTENTS	VI
FIGURES INDEX	IX
TABLES INDEX	XI
CHAPTER 1 - INTRODUCTION	1
UBIQUITIN: A MULTIPURPOSE TAG	1
<i>UBIQUITIN AND PROTEIN DEGRADATION</i>	<i>1</i>
<i>UBIQUITIN AND TRAFFICKING</i>	<i>2</i>
UBIQUITYLATION	4
<i>REGULATION AND SPECIFICITY OF UBIQUITYLATION</i>	<i>4</i>
DEUBIQUITYLATION	7
<i>CATALYTIC CORE STRUCTURE OF DUBS</i>	<i>9</i>
DROSOPHILA FAT FACETS (FAF), A DEVELOPMENTALLY REGULATED USP	12
<i>FAF AND COMPOUND-EYE DEVELOPMENT</i>	<i>12</i>
<i>FAF AND OOCYTE DEVELOPMENT</i>	<i>14</i>
FAT FACETS IN MOUSE (FAM)	15
<i>TEMPORAL AND SPATIAL REGULATION OF FAM DURING DEVELOPMENT</i>	<i>16</i>
<i>FAM SUBSTRATES AND BINDING PROTEINS</i>	<i>18</i>
<i>SUBCELLULAR LOCALISATION OF FAM</i>	<i>20</i>
STRUCTURE AND FUNCTION OF USPS	21
AIMS	23
CHAPTER 2 - MATERIALS AND METHODS	25
ABBREVIATIONS	25
MATERIALS	28
<i>ANTIBODIES</i>	<i>28</i>
<i>BACTERIAL STRAINS</i>	<i>29</i>
PLASMIDS	29
<i>Usp9 SPECIFIC REGION FOR IN SITU HYBRIDISATION</i>	<i>30</i>
ENTRY CLONES	31
<i>HIGHLY CONSERVED REGIONS (HCRS)</i>	<i>31</i>
<i>EGFP</i>	<i>31</i>
<i>USP CATALYTIC CORES</i>	<i>31</i>
<i>FAM N AND C-TERMINAL EXTENSIONS</i>	<i>32</i>
<i>FULL-LENGTH FAM (SSSRF±)</i>	<i>32</i>
<i>FAM-CYS MUTANT (C1566S)</i>	<i>32</i>
DESTINATION VECTORS	33
<i>GWPCS2⁺ MT</i>	<i>33</i>
<i>COMMERCIALY AVAILABLE DESTINATION VECTORS</i>	<i>33</i>
EXPRESSION CLONES	34
PLASMID PREPARATION	35
<i>RECOMBINANT BACMID PREPARATION</i>	<i>35</i>
CELL CULTURE AND 293T TRANSFECTIONS	36
WHOLE-MOUNT IN SITU HYBRIDISATION	37

MRNA INJECTIONS INTO ZEBRAFISH EMBRYOS	37
PROTEIN PROCESSING	38
<i>BACTERIAL INDUCTION</i>	38
<i>PREPARATION OF MAMMALIAN CELL LYSATES</i>	39
<i>PREPARATION OF INSECT CELL LYSATES</i>	39
<i>PROTEIN PURIFICATION</i>	39
<i>PARTIAL PROTEOLYSIS</i>	40
<i>PROTEIN PRECIPITATION</i>	40
<i>PROTEIN ELECTROPHORESIS</i>	40
PROTEIN DETECTION	41
<i>WESTERN BLOTTING</i>	41
<i>SILVER STAIN</i>	41
<i>COLLOIDAL COOMASSIE STAINING</i>	42
MASS SPECTROMETRY	42
COMPUTER ANALYSES.....	43
<i>ASSEMBLY OF USP9 SEQUENCE</i>	43
<i>PHYLOGENETIC ANALYSIS</i>	43
<i>STATISTICAL ANALYSIS OF INJECTION PHENOTYPES</i>	44
<i>FAM PEPTIDE MASS ANALYSIS</i>	44
<i>SECONDARY STRUCTURE PREDICTION</i>	45
<i>INTRINSIC PROTEIN DISORDER, DOMAIN & GLOBULARITY PREDICTION</i>	45
<i>STATISTICAL ANALYSIS OF PROTEIN SEQUENCES</i>	45
<i>FOLD PREDICTION</i>	45
CHAPTER 3 - CHARACTERISATION OF USP9 AND ESTABLISHMENT OF A DOMINANT-NEGATIVE DOMAIN SCREEN FOR FAM-BINDING PROTEINS IN ZEBRAFISH	47
INTRODUCTION.....	47
<i>IDENTIFICATION OF FAM-BINDING PROTEINS IN ZEBRAFISH</i>	47
<i>DOMINANT-NEGATIVE DOMAIN SCREEN FOR FAM-BINDING PROTEINS</i>	48
<i>CHAPTER DESCRIPTION</i>	51
ASSEMBLY OF THE ZEBRAFISH USP9 SEQUENCE	52
PHYLOGENETIC ANALYSIS OF FAM ORTHOLOGUES.....	54
IN SITU ANALYSIS OF USP9 IN EARLY ZEBRAFISH DEVELOPMENT	54
DOMINANT-NEGATIVE DOMAIN SCREEN	56
<i>CLONING OF FAM'S HIGHLY CONSERVED REGIONS (HCRs)</i>	56
<i>VALIDATION OF HCR PROTEIN PRODUCTION</i>	57
<i>DOMINANT-NEGATIVE DOMAIN SCREEN PILOT STUDY</i>	57
<i>EXPERIMENTAL DESIGN OF THE DOMINANT-NEGATIVE DOMAIN SCREEN</i>	59
<i>DATA COLLECTION</i>	61
<i>STATISTICAL ANALYSIS</i>	62
EXPRESSION OF FAM HCR PROTEIN	64
<i>GST FUSIONS</i>	65
<i>MAMMALIAN EXPRESSION</i>	67
SUMMARY	69
CHAPTER 4 - STRUCTURAL ANALYSIS OF FAM	72
INTRODUCTION.....	72
<i>ANALYSIS OF PROTEIN DOMAIN STRUCTURE BY PARTIAL PROTEOLYSIS</i>	72
<i>FAM EXPRESSION IN THE BACULOVIRUS EXPRESSION VECTOR SYSTEM</i>	73
<i>CHAPTER DESCRIPTION</i>	74
EXPRESSION OF FAM CATALYTIC CORE.....	75
<i>IDENTIFICATION OF FAM-DUB</i>	75
<i>CLONING AND EXPRESSION OF FAM-DUB</i>	76
EXPRESSION OF FAM'S N AND C-TERMINAL EXTENSIONS.....	76
<i>CLONING AND PREPARATION OF RECOMBINANT BACMID DNA</i>	77
<i>TRANSFECTION OF FAM-N RECOMBINANT BACMID DNA</i>	78
<i>EXPRESSION OF FAM-C GST PROTEIN</i>	78

CLONING OF FULL-LENGTH FAM.....	79
OPTIMISATION OF GST-FAM PROTEIN EXPRESSION.....	80
<i>TRANSFECTION TIME</i>	80
<i>MONOLAYER VERSES SUSPENSION CULTURES</i>	81
<i>FAM-CYS MUTANT EXPRESSION</i>	81
PARTIAL PROTEOLYSIS.....	82
<i>TRYPSIN DIGESTS</i>	83
<i>V8 PROTEASE DIGESTS</i>	84
<i>CHYMOTRYPSIN DIGESTS</i>	84
<i>LARGE-SCALE DIGESTS</i>	85
MASS SPECTROMETRY.....	86
FAM SEQUENCE COMPUTER ANALYSIS.....	87
FAM DOMAIN PREDICTIONS.....	88
<i>FAM[†]: 60-187 AA (GST-FAM: 296-423 AA) SUBDOMAIN 1.1</i>	89
<i>FAM[†]: 219-867 AA (GST-FAM: 455-1103 AA) SUBDOMAIN 1.2</i>	90
<i>FAM[†]: 888-975 AA (GST-FAM: 1124-1211 AA) UBL SUBDOMAIN 1.3</i>	90
<i>FAM[†]: 60-975 AA (GST-FAM: 296-1211 AA) DOMAIN 1</i>	91
<i>FAM[†]: 1009-1549 AA (GST-FAM: 1245-1785 AA) DOMAIN 2</i>	91
<i>FAM[†]: 1556-2025 AA (GST-FAM: 1792-2261 AA) DOMAIN 3: CATALYTIC CORE</i>	91
<i>FAM[†]: 2039-2480 AA (GST-FAM: 2275-2716 AA) DOMAIN 4: C-TERMINAL EXTENSION</i>	92
FAM DOMAIN FOLD PREDICTIONS.....	93
SUMMARY.....	94
CHAPTER 5 - FINAL DISCUSSION.....	96
INTRODUCTION.....	96
CONSERVATION OF FAM AND ITS ORTHOLOGUES.....	97
<i>CONSERVATION OF FAM SEQUENCE</i>	97
<i>CONSERVATION OF FAM EXPRESSION</i>	98
FAM DOMAIN STRUCTURE.....	102
<i>FAM'S PUTATIVE SUBDOMAINS</i>	102
<i>POSSIBLE MODES OF FAM FUNCTION</i>	103
ANALYSIS OF PREVIOUS FAM FRAGMENTS.....	104
<i>FAM'S HIGHLY CONSERVED REGIONS (HCRs)</i>	104
<i>FAM'S N AND C-TERMINAL EXTENSIONS</i>	104
FUTURE WORK.....	106
APPENDIX A.....	108
APPENDIX B.....	121
REFERENCES.....	126

FIGURES INDEX

CHAPTER 1

FIGURE 1.1: THE UBIQUITIN PROTEASOME PATHWAY	1a
FIGURE 1.2: UBIQUITIN AND PROTEIN TRANSPORT REGULATION.....	3a
FIGURE 1.3: UBIQUITIN CONJUGATION.....	4a
FIGURE 1.4: ROLES OF DEUBIQUITYLATING ENZYMES.....	7a
FIGURE 1.5: CONSERVATION OF DEUBIQUITYLATING ENZYME CATALYTIC CORES.....	7b
FIGURE 1.6: THE UCH CATALYTIC CORE	9a
FIGURE 1.7: THE USP CATALYTIC CORE	10a
FIGURE 1.8: RESCUE OF THE <i>FAF</i> MUTANT EYE PHENOTYPE BY <i>FAM</i> TRANSGENES.....	12a
FIGURE 1.9: <i>FAM</i> IS ESSENTIAL FOR PRE-IMPLANTATION MOUSE DEVELOPMENT.....	15a
FIGURE 1.10: <i>FAM</i> EXPRESSION IN WHOLEMOUNT MOUSE EMBRYOS.....	16a
FIGURE 1.11: <i>FAM</i> EXPRESSION IN THE DEVELOPING MOUSE EYE	17a
FIGURE 1.12: SUBCELLULAR LOCALISATION OF <i>FAM</i> DURING MOUSE PRE-IMPLANTATION DEVELOPMENT.....	20a
FIGURE 1.13: PREDICTED DOMAIN STRUCTURE OF <i>FAM</i>	23a

CHAPTER 2

FIGURE 2.1: ENTRY VECTOR AND THE BP REACTION	29a
FIGURE 2.2: CONSTRUCTED DESTINATION VECTOR GWpCS2+MT.....	33a
FIGURE 2.3: COMMERCIALY AVAILABLE DESTINATION VECTORS	33b

CHAPTER 3

FIGURE 3.1: SCHEMATIC OF THE ZEBRAFISH DOMINANT-NEGATIVE DOMAIN SEARCH.....	49a
FIGURE 3.2: CORRECTIONS TO THE PREDICTED <i>Usp9</i> SEQUENCE	53a
FIGURE 3.3: ALIGNMENT OF <i>FAM</i> AND <i>Usp9</i>	54a
FIGURE 3.4: PHYLOGENETIC ANALYSIS OF <i>FAM</i> ORTHOLOGUES.....	54b
FIGURE 3.5: EXPRESSION OF <i>USP9</i> IN THE EARLY STAGES OF ZEBRAFISH DEVELOPMENT	54c
FIGURE 3.6: EXPRESSION OF <i>USP9</i> IN THE LATER STAGES OF ZEBRAFISH DEVELOPMENT	55a
FIGURE 3.7: DESIGN OF FOUR <i>FAM</i> HIGHLY CONSERVED REGIONS	56a
FIGURE 3.8: WESTERN ANALYSIS OF HCR TRANSIENT TRANSFECTIONS INTO 293T CELLS	57a
FIGURE 3.9: PILOT INJECTIONS WITH TDEE mRNA.....	58a
FIGURE 3.10: PILOT INJECTIONS WITH ERL mRNA.....	58b
FIGURE 3.11: PILOT INJECTIONS WITH RKE mRNA	58c
FIGURE 3.12: PILOT INJECTIONS WITH NPF mRNA.....	58d
FIGURE 3.13: HCR mRNA INJECTION PHENOTYPE CATEGORIES.....	61a
FIGURE 3.14: EXPRESSION OF HCR-GST FUSIONS IN BACTERIA	66a

CHAPTER 4

FIGURE 4.1: BAC-TO-BAC BACULOVIRUS EXPRESSION SYSTEM	74a
FIGURE 4.2: SEQUENCE ALIGNMENT OF HAUSP WITH OTHER USP FAMILY MEMBERS.....	75a
FIGURE 4.3: COMPARISON OF USP CATALYTIC CORE POSITION	75b
FIGURE 4.4: WESTERN ANALYSIS OF CATALYTIC CORE EXPRESSION IN BACTERIA.....	76a
FIGURE 4.5: BACULOVIRUS EXPRESSION OF <i>FAM-N</i> IN Sf9 INSECT CELLS	78a
FIGURE 4.6: EXPRESSION AND PURIFICATION OF GST- <i>FAM-C</i> FROM Sf9 INSECT CELLS.....	78b
FIGURE 4.7: CONSERVED MUTATIONS IDENTIFIED IN GATEWAY CLONED <i>FAM</i> SSSRF±.....	79a
FIGURE 4.8: EXPRESSION AND PURIFICATION OF FULL-LENGTH GST- <i>FAM</i> PROTEIN.....	80a
FIGURE 4.9: FULL-LENGTH <i>FAM</i> EXPRESSION OPTIMISATION.....	81a
FIGURE 4.10: PARTIAL PROTEOLYSIS OF GST- <i>FAM-CYS</i> WITH TRYPSIN	83a
FIGURE 4.11: PARTIAL PROTEOLYSIS OF GST- <i>FAM-CYS</i> WITH V8 PROTEASE.....	84a
FIGURE 4.12: PARTIAL PROTEOLYSIS OF GST- <i>FAM-CYS</i> WITH CHYMOTRYPSIN.....	84b
FIGURE 4.13: SCALED-UP PARTIAL PROTEOLYSIS DIGESTS.....	85a
FIGURE 4.14: SEQUENCE ANALYSIS OF GST- <i>FAM</i>	87a
FIGURE 4.15: <i>FAM</i> DOMAIN PREDICTIONS	88a
FIGURE 4.16: PREDICTED <i>FAM</i> DOMAIN FOLDS.....	93c

CHAPTER 5

FIGURE 5.1: COMPARISON OF HCRs WITH PREDICTED <i>FAM</i> DOMAINS.....	104a
FIGURE 5.2: COMPARISON OF <i>FAM</i> FRAGMENTS WITH DOMAIN MAP	104b

TABLES INDEX

CHAPTER 2

TABLE 2.1: PRIMERS FOR SEQUENCING AND TRANSCRIPTION	29b
TABLE 2.2: PRIMERS FOR PCR	30a
TABLE 2.3: TABLE OF EXPRESSION CLONES	34
TABLE 2.4: BACMID SEQUENCING PRIMERS.....	36

CHAPTER 3

TABLE 3.1: TABLE OF <i>FAM</i> ORTHOLOGUE SEQUENCES.....	53
TABLE 3.2: STATISTICAL ANALYSIS OF HCR mRNA INJECTIONS.....	62a
TABLE 3.3: ASSESSMENT OF PIGMENTATION PHENOTYPE	63a

CHAPTER 4

TABLE 4.1: TABLE OF PREDICTED FAM DOMAINS.....	89
TABLE 4.2: TABLE OF PREDICTED FAM SUBDOMAIN FOLDS.....	93a
TABLE 4.3: TABLE OF PREDICTED FAM DOMAIN FOLDS.....	93b