Analysis of Binary Interactions between OTUB1 and E2 Ubiquitin-Conjugating Enzymes

UNIVERSITY OF
LIVERPOOL

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

Nurulisa Zulkifle

October 2012

Untuk Ezra yang disayangi selama-lamanya

Anak unta anak kuda,
Kejar mengejar keliling batu,
Rindu cinta sentiasa ada,
Segar menjalar seiring waktu.

Analysis Of Binary Interactions Between OTUB1 And E2 Ubiquitin-Conjugating Enzymes

Nurulisa Zulkifle

Abstract

Post-translational modification of proteins via ubiquitination is mediated by three enzyme families; E1 activating enzymes, E2 conjugating enzymes and E3 ligases, all of which work in a hierarchical manner to facilitate different forms of protein ubiquitin ranging from mono-ubiquitination to the formation of different forms of ubiquitin chains (Ciechanover *et al.*, 2000). Deubiquitinating enzymes (DUBs) act to remove ubiquitin from modified substrates. Apart from the classic interactions within the E1-E2-E3 enzymatic cascade, an unusual non-hierarchical interaction has been observed between some E2 enzymes and a DUB called OTUB1 (Markson *et al.*, 2009). This observation raises interesting questions concerning the molecular mechanisms and specificity of this unusual E2:DUB partnership. In this study, systematic yeast two-hybrid (Y2H) screens were performed between all human E2 and DUB proteins to analyse the extent of E2:DUB interactions. Putative partnerships between OTUB1 and UBE2D1, UBE2D2, UBE2D3, UBE2D4, UBE2E1, UBE2E2, UBE2E3 and UBE2N were identified. These data correlate well with data from other independent studies, including HTP Y2H screens (Markson *et al.*, 2009) and mass spectrometry (Sowa *et al.*, 2009). An N-terminal truncated form of OTUB1 ($^{\Delta N}$OTUB1) was generated by removing a predicted 39aa N-terminal disordered region (Edelmann *et al.*, 2009). Using this construct in combination with wild type (WT) OTUB1, complementary biophysical studies were performed to investigate the formation of complexes with UBE2D2 and UBE2E1 as these represented the strongest interactions detected in preliminary Y2H studies. Gel filtration chromatography showed convincing complex formation for both $^{\Delta N}$OTUB1:UBE2D2 and $^{\Delta N}$OTUB1:UBE2E1 in 1:1 stoichiometry. The thermodynamic profile of each complex was measured by ITC suggested a stronger affinity between $^{\Delta N}$OTUB1:UBE2D2 (K_d 3.89 μM) than observed for the $^{\Delta N}$OTUB1:UBE2E1 complex (K_d 16.55 μM). The *n* values for both complexes are 1.16±0.06 sites and 0.92±0.03 sites respectively, confirming that both complexes adopt a 1:1 stoichiometry. Observing the UBE2D2 ($^{1}H^{15}N$)-HSQC NMR spectral changes that occurred upon addition of unlabelled $^{\Delta N}$OTUB1 allowed the identification of potential residues of contact between the two proteins. From this study, we were able to predict that the 1st α-helix, the L1 loop of the 3rd and 4th β-sheet, the L2 loop connecting the 4th β-strand and the H2 α-helix within UBE2D2 were likely to be the binding surfaces for

OTUB1. Point mutants corresponding to predicted contact residues in UBE2D2 were generated and tested in Y2H studies to determine their role in facilitating the formation of both E2:OTUB1 and E2:E3-RING complexes. This data suggests that in some, but not all cases, OTUB1 and E3-RINGs bind competitively to the same interface on E2 proteins. Preliminary immunofluorescence studies show that partner proteins predominantly co-localise in the cytoplasm, except UBE2E1 which is predominantly nuclear. Data from this study allowed us to propose a model of how OTUB1:UBE2D2 complex may forms and functions. Significantly, many of these predictions have now been verified by independent structural studies and subsequent live cell microscopy studies in our lab.

<div align="right">(482 words)</div>

Table of Contents

Chapter Two: Materials and Methods

Chapter Six: Discussion and Conclusion

References

List of Figures

List of Tables

List of Abbreviations

1D	One-dimensional
2D	Two-dimensional
3AT	3-amino-1,2,4-triazole
3D	Three-dimensional
AD	GAL4 activating domain
Ade	Adenine
ATP	Adenosine triphosphate
APS	Ammonium persulphate
BC-PCR	Bacterial colony-PCR
BD	GAL4 DNA binding domain
BioGRID	Biological General Repository for Interaction Datasets
BMRB	Biological Magnetic Resonance Bank
bp	base pair
C-terminal	Carboxy-terminal / COOH-terminal
$CaCl_2$	Calcium chloride
cDNA	Complementary DNA
CO_2	Carbon dioxide
D_2O	Deuterium oxide
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
dH_2O	Distilled water
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DO	Dropout
DTT	Dithiothreitol
DUB	Deubiquitinating enzyme
EMBL	The European Molecular Biology Laboratory
FCCS	Fluorescence cross correlation spectroscopy
$FeCl_2$	Iron (II) chloride / ferrous chloride
FL	Full-length
GF	Gel filtration
GFP	Green fluorescent protein
GW	Gateway®

HECT	Homologous to E6-AP carboxy terminus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HSQC	Heteronuclear Single Quantum Coherence
HPRD	Human Protein Reference Database
IEC	Ion exchange chromatography
IPTG	Isopropyl-β-D-thio-galactopyranoside
ITC	Isothermal titration calorimetry
K_a	Binding affinity
K_d	Dissociation constant
kcal	Kilocalories
KH_2PO_4	Monopotassium phosphate
*Lac*Z	Gene encoding β-galactosidase
Lys	Lysine
$MgCl_2$	Magnesium chloride
MHz	Megahertz
MINT	Molecular Interactions Database
$MnSO_4$	Manganese (II) sulphate
MWCO	Molecular weight cut-off
N-terminal	Amino terminal / NH_2-terminal
Na_2HPO_4	Sodium phosphate dibasic
NaN_3	Sodium azide
NaOH	Sodium hydroxide
NEAA	Non-essential amino acids
NH_4Cl	Ammonium chloride
Ni^{2+}	Nickel
NMR	Nuclear magnetic resonance
NCBI	National Centre for Biotechnological information
OD	Optical density
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
pI	Isoelectric point
PMSF	Phenylmethanesulfonylfluoride
PPI	Protein-protein interaction

ppm	parts per million
rf	radiofrequency
RING	Really Interesting New Gene
RNA	Ribonucleic acid
RONN	Regional Order Neural Network
rpm	revolutions per minute
RT	Room temperature
SD	Synthetic defined
SDS	Sodium dodecyl sulphate
SOB	Super optimal broth
SOC	SOB with catabolite repression (added glucose)
TCEP	Tris(2-carboxyethyl)phosphine
TEMED	N,N,N',N'-tetramethylethylenediamine
TEV	Tobacco etch virus
T_m	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
UAS	Upstream activating sequence
Ub / Ubq	Ubiquitin
UBL	Ubiquitin-like
UPS	Ubiquitin-Proteasome System
UV	Ultraviolet
WT	Wild type
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Y2H	Yeast two-hybrid
YC-PCR	Yeast colony-PCR
ΔG	Change in Gibbs energy
ΔH	Change in enthalpy
ΔS	Change in entropy

Acknowledgements

First and foremost, I would like to thank the Ministry of Higher Education Malaysia (MOHE) and Universiti Sains Malaysia (USM) for providing the scholarship and study leave. Without their support, this PhD would not have been possible.

I owe my deepest gratitude to my supervisor Professor Chris M. Sanderson for the continuous encouragement, excellence supervision and extraordinary patience throughout the completion of this work. I am also hugely indebted to Professor Lu-Yun Lian for the access to the equipments and facilities in School of Biosciences and NMR centre. To Dr. Igor Barsukov, Dr. Paul Elliott and Dr. Martyna Pastok, a sincere thanks for the advice and technical assistance that they provided.

To current and former members of the Sanderson lab: Julie, Jonathan, Russell, Helen, Rob, Kelly, Amy, Emily and Dave, all have contributed to a happy tune in my daily lab life, therefore a very huge thanks to all of you. To Yvonne, Jia Lih and the rest of the 5th floor Nuffield Wing occupants, thank you for all of your help along the way. Further thanks go to the rest of the staff and fellow students in the department, always friendly and helpful since day one.

I would like to acknowledge everyone outside of the lab especially my ex-housemates in Liverpool, always get me smile and laugh along this tearful journey. Izyan, Shafiqah, Hada, Ad, Eeka, Maryam, Najwa and Intan; these girls are indeed true friends and the most irreplaceable pieces of my life. Also to my loyal old friends in Malaysia thank you for always being there. Especially to Linda, you are the only one that never leave.

To my other half Yazdi, thank you for the encouragement, support, patience and belief in me. I have put my life and future in good hands and only God knows how glad I am. And lastly to my beloved family: Mak, Ayah, Asma, Ezakry, Ain, Amir, Nani, Haziq and Ammar, my final but thoroughly heartfelt thank you is for all of you. Especially to my parents, you are the reason I thrive to be better. Your pride for me is my main goal in life.

Chapter One:
INTRODUCTION

1.1   Overview of the human ubiquitin system

Many vital body functions such as temperature regulation and maintenance of blood pressure serve to maintain the state of homeostasis in our body. The same principle applies at the level of a single cell where control of protein homeostasis is essential for regulating cellular physiology or responding to adverse signals. One of the most important modes of post-translational regulation is protein ubiquitination, a highly dynamic process that governs nearly every function in human cells (Hershko and Ciechanover, 1998). Over the past two decades, ubiquitination has been increasingly acknowledged as more than simply a signal for protein degradation, but also a main regulator of a diverse array of different biological processes (Weissman, 2001). It is a reversible post-translational modification of protein, almost as common as phosphorylation, which involves the addition of the ubiquitous small protein aptly named ubiquitin to the lysine residue of a target substrate (Hershko and Ciechanover, 1998).

In general, ubiquitin attachment occurs through a sequential enzymatic cascade involving an E1 ubiquitin-activating enzyme, an E2 ubiquitin-conjugating enzyme and an E3 ubiquitin-protein ligase, resulting in the modification of a specific protein substrate (Pickart, 2001). In the classical view of ubiquitination, a polyubiquitin chain is synthesised by serial addition of ubiquitin moieties to the Lys48 residue of the previously conjugated ubiquitin (Hershko and Ciechanover, 1998). This particular polyubiquitinated substrate is recognised as a target for degradation by a multi-subunit ATP-dependent protease complex known as the 26S proteasome (Tashiro *et al.*, 1997; Pickart, 2001). The reversibility of ubiquitination is provided by deubiquitinating enzymes (DUBs) which either prevent ubiquitin attachment to

the target substrate or remove ubiquitin from specific target proteins (Komander *et al.*, 2009; Reyes-Turcu *et al.*, 2009).

1.2 **Ubiquitin and ubiquitin like molecules**

Ubiquitin is a 76 amino acid globular protein with a molecular mass around 8.5 kDa and is highly conserved throughout eukaryotes (Weissman, 2001). Since its discovery over three decades ago, it has become abundantly clear that the ubiquitin system is an essential feature of all aspects of eukaryotic biology. Due to its pervasive action, ubiquitin does not seem to be produced in excess but the free pool of ubiquitin monomer is maintained at an adequate level to ubiquitinate the large number of its potential substrates in human cells (Kimura and Tanaka, 2010). In yeast as well as in higher eukaryotes, ubiquitin is initially expressed in the form of a precursor either as polyubiquitin, a linear fusion protein consisting of four or more ubiquitin copies in a head-to-tail configuration, or as fusion proteins between ubiquitin and large and small essential ribosomal polypeptides, L40 and S27, respectively (Finley *et al.*, 1989; Redman and Rechsteiner, 1989). These ubiquitin precursors are cleaved by DUBs to release identical functional monomeric ubiquitin units.

In the ubiquitination process, ubiquitin covalently attaches to the lysine residues of target proteins via its carboxy-terminal glycine residue to form an isopeptide linkage in an ATP-dependent fashion. Multiple ubiquitin can be covalently added to a substrate successively by E1, E2 and E3 enzymes, producing a substrate conjugated with polyubiquitin (Pickart, 2001; Dye and Schulman, 2007). The key feature of ubiquitin is the fact that it contains seven lysines, each of which can potentially mediate attachment to other ubiquitin molecules, allowing the formation of a range of structurally distinct polyubiquitin chains (Figure 1.1) (Komander, 2009).

N- MQIFV**K**TLTG**K**TITLEVEPSDTIENV**K**A**K**IQD**K**EGIPPDQQRLIFAG**K**QLEDG

RTLSDYNIQ**K**ESTLHLVLRLRGG **-C**
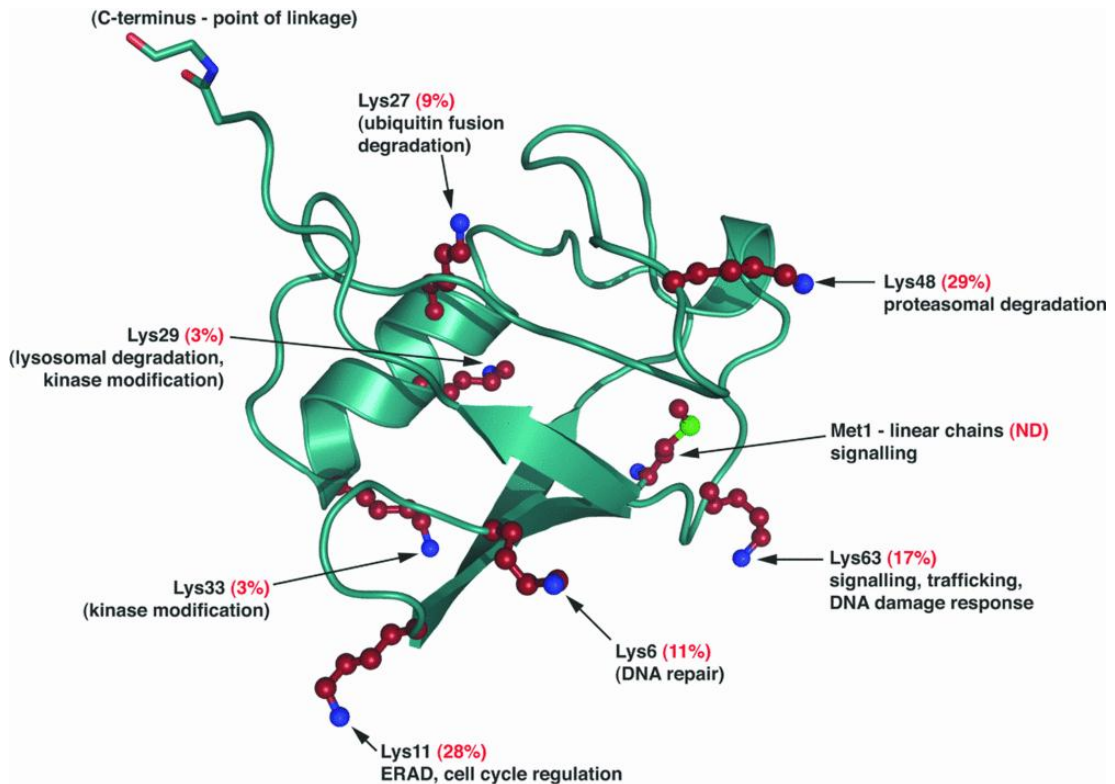
Figure 1.1   Ribbon diagram representing the tertiary structure of ubiquitin and the functional implications of specific linkages: The structure of ubiquitin reveals that all seven lysine residues (red, with blue nitrogen atoms) and methionine (with a green sulphur atom) reside on different surfaces of the molecule. The percentage numbers refer to the relative abundance of the particular linkage in *S. cerevisiae* (Komander, 2009). Above is the amino acids sequence of ubiquitin with the seven lysine residues highlighted in red.

In addition to ubiquitin itself, multiple polypeptides that are distinct from but related to ubiquitin are also enzymatically conjugated to target substrates via processes very similar to ubiquitination (Gill, 2004). These ubiquitin-like proteins (UBLs) participate in many biological processes including gene transcription, signal transduction, autophagy and cell-cycle control (Welchman *et al.*, 2005). So far, SUMO2/3 and NEDD8 are the only UBLs

known to participate in chain formation (Tatham *et al.*, 2001; Xirodimas *et al.*, 2008; Matic *et al.*, 2008). Members of the UBL family are listed in the table below:

Table 1.1   Known and putative UBLs in *Saccharomyces cerevisiae*: Reviewed in Hochstrasser (2009). ND = Not detectable by standard BLAST search.
* = For each two ubiquitin-related domains.

UBL	Identity with ubiquitin (%)
Known UBLs	
Rub1 (NEDD8)	55
FUBI (also known as MNSF-β or FAU)	38
FAT10	32 and 40*
ISG15	32 and 37*
Smt3 (SUMO1, SUMO2, SUMO3)	18
Atg8	ND
Atg12	ND
Urm1	ND
UFM1	ND
Putative UBLs	
BUBL1, BUBL2	Variable (up to 80%)
UBL-1	40
SF3A120	30
Oligoadenylate synthetase	30 and 42*

1.3   Roles of ubiquitination

1.3.1   Intracellular protein degradation by ubiquitin-proteasome pathway: The classical ubiquitin role

Protein degradation in eukaryotic cells is carried out either by lysosomal or proteasomal degradation (Cooper, 2000). Lysosomes are membrane-enclosed organelles that contain a large variety of hydrolytic enzymes which digest extracellular proteins taken up by endocytosis. Most of the proteolysis of cytosolic proteins that occurs in lysosomes is non-specific (Knopp *et al.*, 1993). In contrast, the Ubiquitin-Proteasome System (UPS) uses ubiquitin as a marker which can target cytosolic and nuclear proteins for selective destruction by proteasomes (Pickart, 2001). It is the responsibility of the UPS to identify and exterminate damaged and faulty proteins or those simply surplus for requirement to maintain the right amount of proteins within cells. The UPS can malfunction in two ways: it can either become overzealous and destroy useful protein inappropriately, or it can be restrained in some way resulting in the build up of potentially harmful proteins, which can then reach toxic levels. An imbalance in the UPS is thought to occur in common diseases especially cancer (Scheffner *et al.*, 1990; Loda *et al.*, 1997; Joazeiro *et al.*, 1999; Maxwell *et al.*, 1999; Waterman *et al.*, 1999; Bignell *et al.*, 2000). The UPS marks a target protein for destruction by the addition of a specific form of ubiquitin modification, in essence, proteins tagged with polyubiquitin chains of more than four ubiquitin residues (Thrower *et al.*, 2000). There is a lot of ubiquitin present in cells but it cannot attach itself to protein at random due to its highly regulated and controlled process which has evolved ways of avoiding any unwanted protein degradation.

The multi-stage ubiquitination process occurs by the sequential action of three different enzymes (Pickart, 2001; Weissman, 2001; Markson *et al.*, 2009). As can be seen in Figure 1.2, ubiquitination is initiated by the formation of high energy thioester intermediates, E1-Cys~ubiquitin generated upon binding between the cysteine active site of E1 ubiquitin-activating enzyme with the C-terminal glycine (G76) of ubiquitin in an ATP-dependent

process (Groettrup *et al.*, 2008). The activated ubiquitin moiety is then transferred from the E1 to the E2 ubiquitin-conjugation enzyme by transthiolation, again involving the carboxyl terminus of ubiquitin to generate an E2-Cys~ubiquitin intermediate (Michelle *et al.*, 2009). The E2 protein acts as an escort for ubiquitin to its next destination, which is either an E3 HECT (homologous to E6-AP terminus) ligase or directly to a specific substrate protein via E3 RING (Really Interesting New Gene). Unlike the situation with HECT type which really involves in accepting and delivering the ubiquitin molecule to the substrate, E3-RINGs primarily act as a platform on which the active E2-ubiquitin complex and target protein substrate can meet and interact. The E3 proteins represent a pivotal part of ubiquitin cascade processes as they define both substrate specificity and the recruitment of selective E2 enzymes. In each case, an E2 enzyme loaded with activated ubiquitin interacts with one of a specific subset of E3 proteins in order to transfer ubiquitin to the target protein by the formation of an isopeptide bond between ubiquitin's C-terminal glycine and the ε-amino group of a lysine residue within the protein substrate or a selective lysine residue in previously added ubiquitin, thereby forming polymeric chains of selective structure and function (Deshaies and Joazeiro, 2009). The Lys48 polymeric chains target proteins to the cell's waste disposal unit, the proteasome (Hersko and Ciechanover, 1998). The proteasome binds and removes the polyubiquitin chain and unfolds the protein. The protein is then threaded through the proteasome chamber before being chopped up into component building blocks, which are reused for the synthesis of new proteins while the ubiquitin is recycled (reviewed in Hochstrasser (1996)).
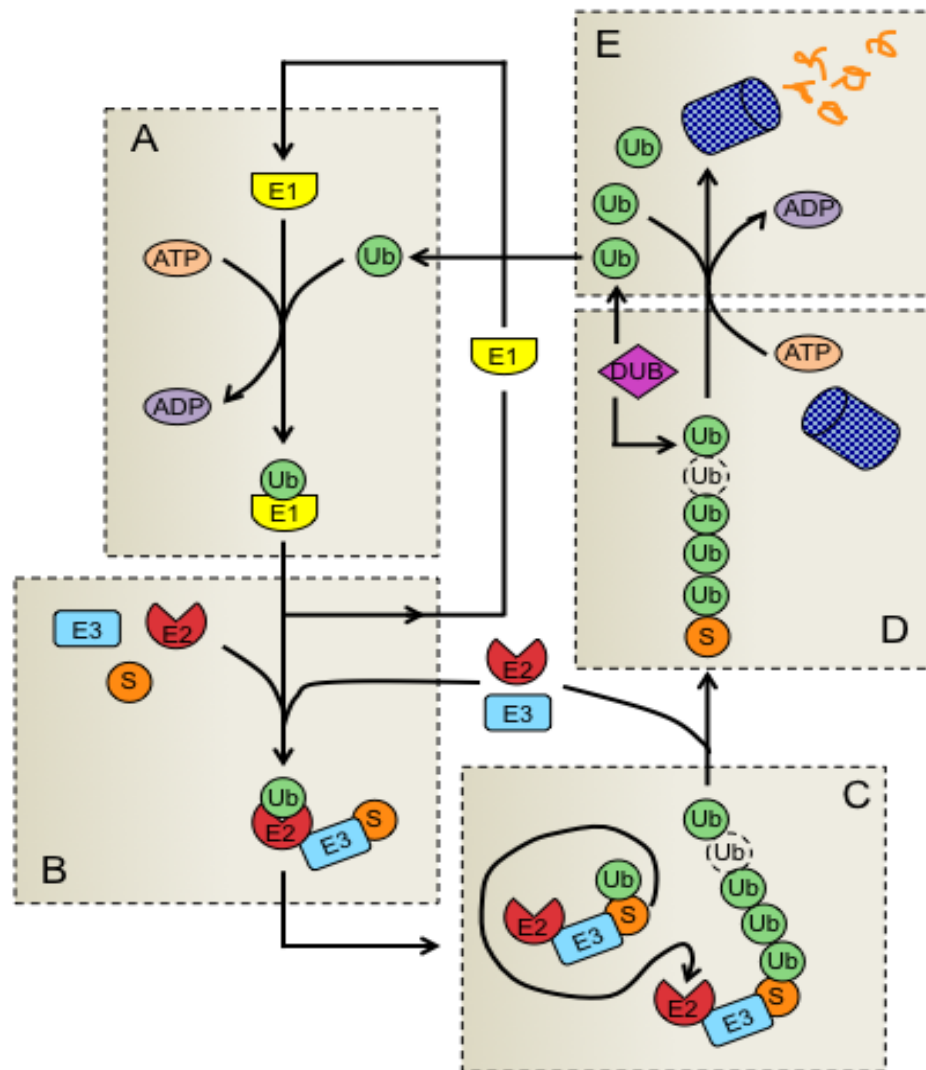
Figure 1.2 **Protein degradation through the UPS: (A)** ATP-dependent ubiquitin activation by E1 followed by ubiquitin delivery to E2. **(B)** Complex formation by E2-Cys~Ub, E3 and the substrate, involving formation of thioesters between the active-site cysteines of E1 and E2 enzymes and the carboxy-terminal carboxylate of ubiquitin. **(C)** Transfer of ubiquitin(s) to the substrate lysine(s) to earmark the substrate with a polyubiquitin chain. **(D)** Polyubiquitinated substrates are released from the E3. Proteasome recognises the polyubiquitin chain as a signal for deubiquitination and subsequent degradation. **(E)** The proteasome unfolds the substrate in an ATP-dependent manner, removing the ubiquitin chain through a proteasome-associated ubiquitin hydrolase activity before threading the unfolded protein into the proteasome. Here, the protein is degraded and released as peptides together with the ubiquitin molecules.
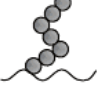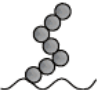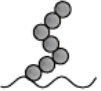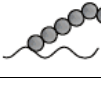
1.3.2   The diversity of ubiquitin chain configurations

Protein ubiquitination is essential to the process of protein homeostasis. However, different forms of protein ubiquitination confer different affects other than proteasome mediated degradation (Weissman, 2001). Proteins can be modified through the conjugation of monoubiquitin or polyubiquitin chains of variable length on any of the seven ubiquitin Lys residues (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 or Lys63) or the amino-terminal Met (Met1) of the ubiquitin monomer. Ubiquitin chains can thus be connected by at least eight different homotypic linkages, as well as by a range of atypical chains such as heterologous, forked or mixed chains (Ikeda and Dikic, 2008; Iwai and Tokunaga, 2009; Ye and Rape, 2009). The various conformations of ubiquitin chain create a range of molecular signals in cells (Table 1.2).

Among these examples are non-canonical polyubiquitination through the Lys63 residues, which are involved in DNA damage response (Nakada *et al.*, 2010), stress response (Arnason and Ellison, 1994), mitochondrial DNA inheritance (Fisk and Yaffe, 1999) and ribosomal function (Spence *et al.*, 2000). Another mode of conjugation is exemplified by linking ubiquitin molecules via Lys29 of ubiquitin, which may also act as a signal for degradation. However, Lys29-ubiquitinated Deltex, a regulator of Notch signalling was found to be degraded by lysosomal rather than proteasomal degradation pathways (Chastagner *et al.*, 2006). Meanwhile, Lys11-linked chains have been implicated in ERAD (endoplasmic-reticulum-associated degradation) where Lys11-linked ubiquitin chains were co-purified with UBA/UBX family proteins (Alexandru *et al.*, 2008) in which, the UBX domains interact with the AAA (ATPase associated with various cellular activities) protein Cdc48/p97, an important regulator of ERAD (Ye *et al.*, 2001; Schuberth and Buchberger, 2008). On the other hand, monoubiquitination (or multi-monoubiquitination) of surface receptors acts as a signal for internalisation mediated by ESCRT (endosomal sorting complex required for transport) (Haglund *et al.*, 2003; Raiborg and Stenmark, 2009). This could also be observed in yeast in which a monoubiquitin conjugate attached to membrane

proteins or associated transport modifiers serves as a signal for internalisation into the endocytotic pathway, ultimately resulting in lysosomal proteolysis (Hicke and Dunn, 2003).

Table 1.2 **Functions of various ubiquitin chains:** The variety of biological functions of ubiquitination inside the human cells resulting from different types of ubiquitin chain formation.

Type of ubiquitin chains		Biological function	References
	Monoubiquitination	Trafficking, endocytosis, virus budding, nuclear export, DNA repair, gene expression	(Haglund *et al.*, 2003)
	Linear ubiquitin	Signalling	(Rahighi *et al.*, 2009)
	Lys6 chain	Proteasomal degradation	(Baboshina and Haas, 1996; Xu *et al.*, 2009)
	Lys11 chain	Proteasomal degradation, protein stability	(Baboshina and Haas, 1996; Nishikawa *et al.*, 2004; Xu *et al.*, 2009)
	Lys27 chain	Proteasomal degradation	(Xu *et al.*, 2009)
	Lys29 chain	Lysosomal degradation, proteasomal degradation, regulation of kinase activity, protein stability	(Nishikawa *et al.*, 2004; Chastagner *et al.*,2006; Al-Hakim *et al.*, 2008; Xu *et al.*, 2009)
	Lys33 chain	Proteasomal degradation, regulation of kinase activity	(Al-Hakim *et al.*, 2008; Xu *et al.*, 2009)
	Lys48 chain	Proteasomal degradation	(Chau *et al.,* 1989)
	Lys63 chain	Trafficking, kinase activation, proteolytic degradation of misfolded proteins	(Olzmann *et al.*, 2007; Tan *et al.*, 2008; Wooten *et al.*, 2008)
	Forked chain	Negative effect on proteasomal degradation	(Kim *et al.*, 2007; Kim *et al.*, 2009)

1.4   Ubiquitin conjugation components

1.4.1   Human E1 ubiquitin-activating enzyme

The ubiquitin activating enzyme (UBE1) stands on the top of the ubiquitin hierarchy, being responsible for activating ubiquitin and preparing it for one of a large number of distinct E2 ubiquitin-conjugating enzymes. Before the formation of the thioester bond between the C-terminus of ubiquitin and E1, the ubiquitin was activated by adenylation via binding with MgATP, leading to the formation of a ubiquitin adenylate intermediate that serves as the donor of ubiquitin to a cysteine in the E1 active site (Haas *et al.*, 1982; Groettrup *et al.*, 2008). Each charged E1 molecule carries two molecules of activated ubiquitin: one as a thioester, and one as an adenylate. Only the thiol-linked ubiquitin is transferred to the E2 by transthiolation. E1 is an efficient enzyme, justified by its maximum turnover number of ATP-AMP exchange (1-2 s^{-1}) involving all steps from ATP binding through thioester formation (Haas *et al.*, 1982) when compared to the catalytic rate (k_{cat}) of substrate ubiquitination which is 10- to 100-fold slower (Mastrandrea *et al.*, 1999). This allows the production of sufficient activated ubiquitin for all cellular ubiquitination reactions, even though the concentration of the E1 protein is thought to be lower than the total concentration of E2 proteins (Pickart, 2001).

The human genome encodes only two E1 genes: A1S9T (Zacksenhaus and Sheinin, 1990; Zacksenhaus *et al.*, 1990), which appeared to be an ubiquitin-activating enzyme and designated later as UBE1, and UBA6 which was recently identified by homology searches using one of the two ThiF-homology motifs that compose the adenylation domain (Pickett, 2007).

1.4.2   Human E2 ubiquitin-conjugating enzyme

The second component in the ubiquitin cascade are the E2 enzymes, which are the key enzymes in a ubiquitin pathway. In contrast with E1 enzymes, E2 proteins will not forge thioester linkages with ubiquitin in the absence of E1. Instead, the charged E1 will bind to

the E2 conjugating enzyme through its ubiquitin fold domain to bring the catalytic cysteine on the E2 into close enough proximity to the E1 active site, thereby allowing transfer of the thioester bound from ubiquitin from the E1 to the E2 (Pickart, 2001). Once loaded with ubiquitin, E2 proteins can donate the thioester linked ubiquitin to the active cysteine of a HECT domain E3 ubiquitin ligase (E3) which then transfers ubiquitin to a protein substrate. Alternatively, E2 proteins can transfer ubiquitin directly onto a lysine residue of a target substrate protein. In this case, an E3 RING protein may assist the E2 protein in recognising the appropriate protein target (Ptak *et al.*, 2001).

The importance of E2 proteins lies in their role of determining the lysine preferences in ubiquitination, hence dictates the topology of polyubiquitination (David *et al.*, 2010). This can be seen in BRCA/BARD, an E3 ligase which acts with UBE2K to generate Lys48-conjugated chains, whereas with the UBE2N-UBE2V1 combination of E2 proteins, these catalyse the formation of polyubiquitin chains conjugated through Lys63 (Christensen *et al.*, 2007). Also, the Pellino1 E3 protein acts with UBE2N-UBE2V1 to generate lysine Lys63 chains, but when acting with UBE2R1, it actually catalyses the formation of Lys48 chains. However, when this E3 functions with members of the UBE2D family (Ubc4/5), it promotes the formation of Lys11 and Lys48 chains (Ordureau *et al.*, 2008).

1.4.2.1 E2 structure

E2 proteins are distinguished by the presence of a UBC (ubiquitin conjugating) domain that is approximately 35% conserved throughout eukaryotes. This domain consists of about 150 amino acids (Pickart, 2001), which is organised into four standard alpha helices (α1-4), a short 3_{10} helix, and four-stranded (S1-4) antiparallel β-sheets as demonstrated in Figure 1.3 (Pickart, 2001; Özkan *et al.*, 2005). The β-sheets and α2 form a central region that is bordered by the α1 region at one end and α3/α4 regions at the other. The cysteine residue of the active site lies in a long loop that connects S4 to α2, and sits in a shallow groove formed by upstream residues of the same loop on one side, and residues of the α2-α3 on the other

(Pickart, 2001). A large proportion of the most highly conserved E2 residues are found around the active site. Some of these residues interact with ubiquitin, and others are presumed to interact with the E1 protein (Pickart, 2001). In contrast, the region of the E2 protein that faces the active site contains some of the most poorly conserved residues found within the protein, perhaps to make the structural region easier to define. It is suggested that a proportion of these residues may have diverged under low selective pressure due to a lack of essential function. However, it is also possible that these differences may actually facilitate specific partner interactions (Pickart, 2001). Support for the theory that the region opposite the active site may be involved in specific E2 interactions is provided by the analysis of the UbcH10 (UBE2C) protein, which shows that many residues in this region are conserved among orthologues of UBE2C. However, these residues are not conserved among different E2 enzymes within the same species, thus implying that these residues are specific to the function of UBE2C (Pickart, 2001).
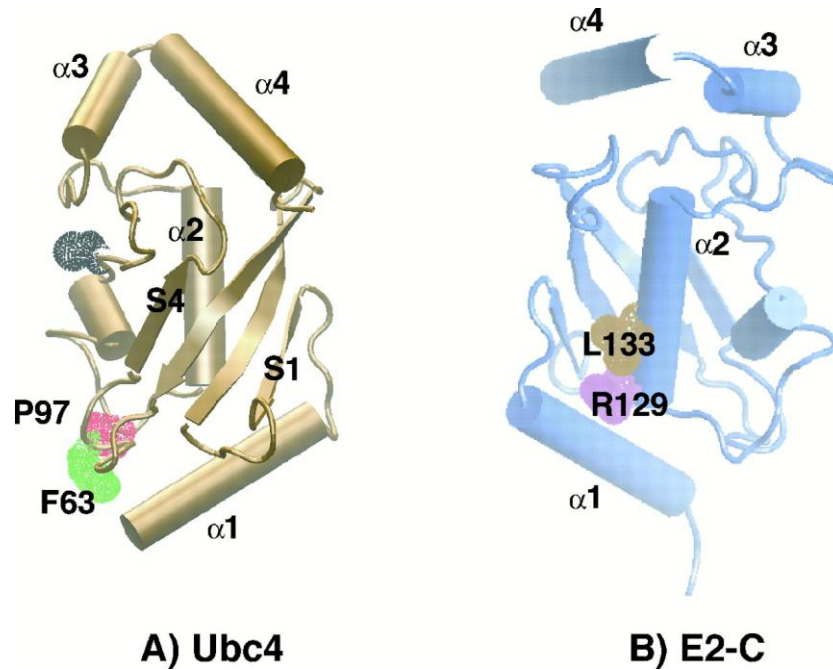
Figure 1.3 **Conservation of E2 structures between species**: **(A)** Yeast Ubc4 (UBE2D1 homologue). The active site cysteine (C86) is highlighted black, along with side chains corresponding to F63 (green) and P97 (red) of UBE2C. **(B)** Clam E2-C (UBE2C homologue). Relative to (A), the molecule was rotated ~180° along the *y* axis. The side chains of R129 (purple) and L133 (brown), corresponding to the first and fourth residues of a destruction box located in α2, are highlighted. Picture from Pickart (2001).

The structures of UbcH7 (UBE2L3) bound to RING E3 c-Cbl and of UbcH5B (UBE2D2) bound to CNOT E3 RING have been solved (Zheng *et al.*, 2000; Dominguez *et al.*, 2004) and would therefore provide a general explanation of E2 specificity. The CNOT4 RING finger interacts specifically with UBE2D2 and not with UBE2L3 despite the fact that in both complexes the same regions of the E2s are involved (helix α1, loops L1 and L2) in the interaction with the RING (Winkler *et al.*, 2004). This suggests that although the three regions of UBE2D2 and UBE2L3 involved in the binding are similar, the binding properties must be different. For example, several contacts between the H1 helix of UBE2D2:CNOT4 are not present in the UBE2L3:c-Cbl complex (Figure 1.4). Furthermore, UBE2L3:c-Cbl involved mainly hydrophobic or uncharged residues while UBE2D2:CNOT4 governs the binding by charged residues.

UBE2D2:CNOT4   UBE2L3:c-Cbl
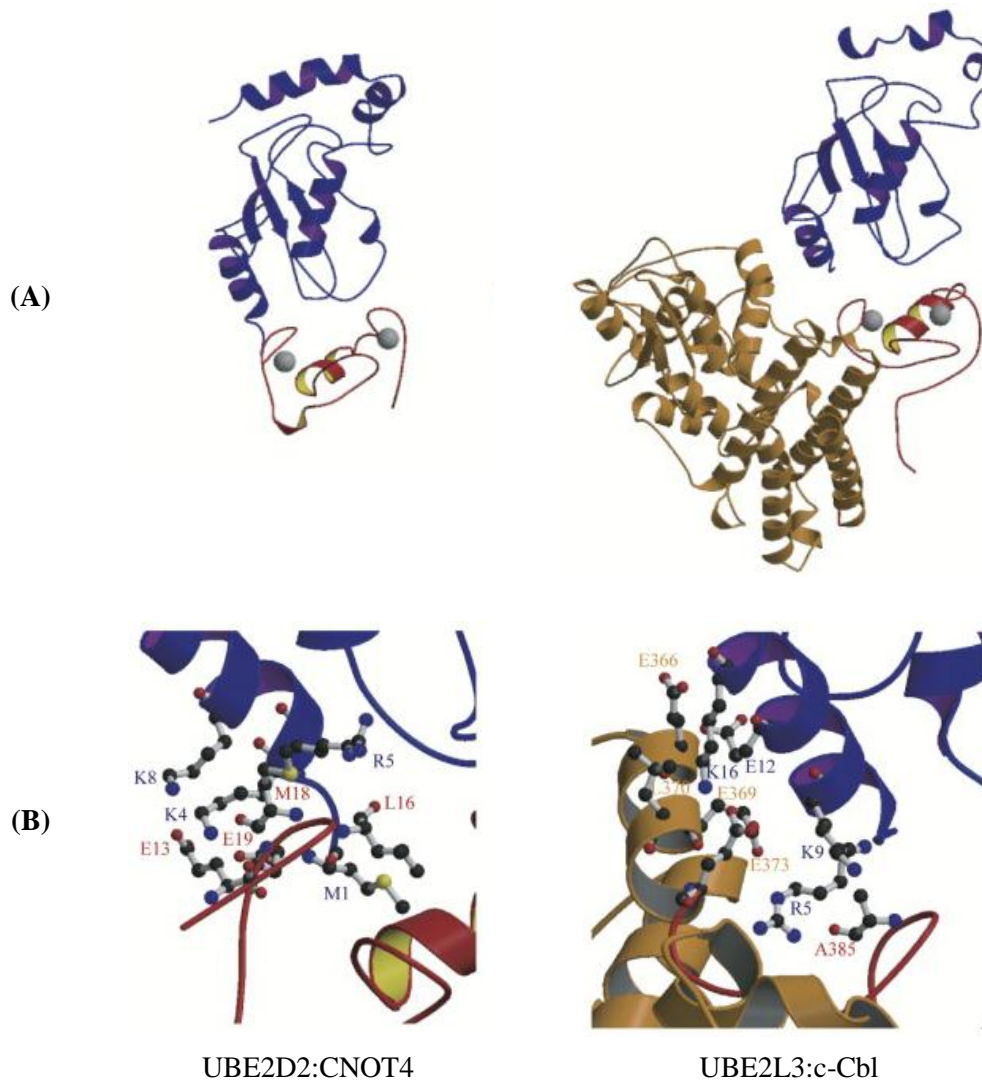
Figure 1.4   Comparison between the UBE2D2:CNOT4 docking model and the UBE2L3:c-Cbl crystal structure: (A) The orientation of the RING domain compared to the E2 enzyme is similar in both complexes. UBE2D2 and UBE2L3 are coloured blue, CNOT4 and c-Cbl RING domains are coloured red, and the other domains of c-Cbl are coloured brown. **(B)** The helix α1 of UBE2D2 makes many contacts with the CNOT4 RING domain, whereas the helix α1 of UBE2L3 interact mainly with the linker region of c-Cbl. Residues 1, 4, 5, and 8 of UBE2D2 and 5, 9, 12, and 16 of UBE2L3 are labelled in blue. Residues 13, 16, 18, and 19 of CNOT4 and 385 of c-Cbl RING are labelled in red. Residues 366, 369, 370, and 373 corresponding to the linker region of c-Cbl are labelled in brown. Picture from Dominguez *et al.* (2004).

1.4.2.2   E2 family diversity

The existence of at least 37 E2 enzymes suggest some degree of specificity in ubiquitination events with regards to partners, substrates, and functions. Generally, the E2 family are classified based on the existence of additional extensions to the UBC catalytic core (Figure 1.5). E2 class I consist of the catalytic domain only, while others with additional N- or C-terminal extensions are classified as Class II and Class III, respectively. E2s that have extension in both termini are categorised Class IV. These extensions are involved in functional differences between E2s, which involve differences in subcellular localisation, stabilisation of the interaction with E1 enzymes, or modulation of the activity of the interacting E3s (van Wijk and Timmers, 2010).

While the majority of E2 enzymes contain predominantly the UBC domain alone along with N- or C-terminal flanking regions, there are also a few examples of domains outside of the UBC region, which are required for functions, both within and external to the ubiquitin cascade. For example, it has been hypothesised that the ubiquitin binding UBA domain C-terminal of the UBC domain in UBE2K may be involved in ubiquitin chain formation based on a similar domain in the yeast Ubc1 protein (Merkley and Shaw, 2004). The massive atypical E2 BIRC6 is annotated for an anti-apoptotic BIR domain distal to the C-terminal UBC domain. At almost 5 000 amino acids in length it is likely however that BIRC6 contains many ordered domains. It is thought to be the sole chimeric E2:E3 ligase that utilises the BIR domain to bind and mediate ubiquitination of SMAC and caspase-9 (Bartke *et al.*, 2004).
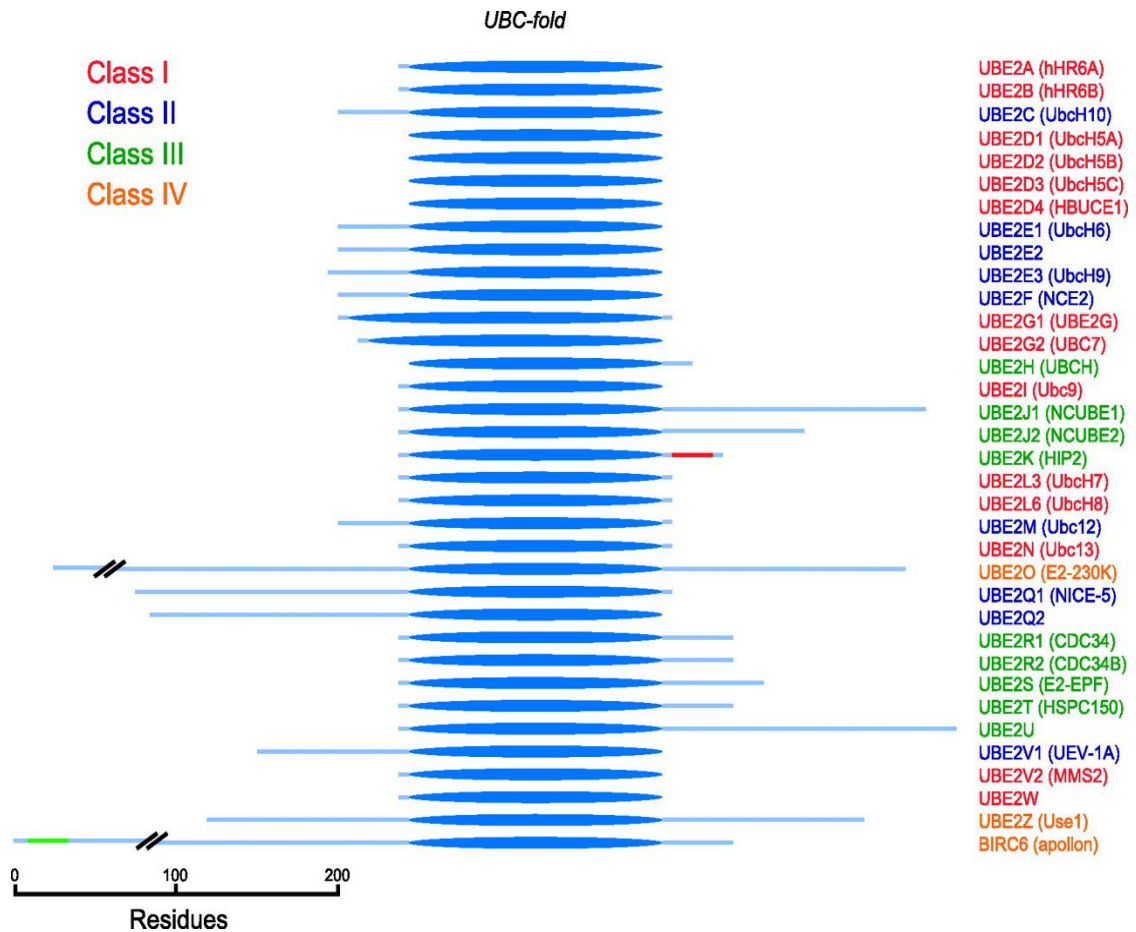
Figure 1.5 **The family of human E2 ubiquitin-conjugating enzymes:** Schematic overview of the superfamily of active, human E2 enzymes. The UBC fold is indicated as a dark-blue thick line, the N- and/or C- termini are represented by the light blue line, UBA domain of UBE2K is indicated as a red line and the BIR domain of BIRC6 as a green line. Different classes of E2 enzymes are indicated in different colours. Scale bar indicates protein lengths (aa). Picture from van Wijk and Timmers (2010).

1.4.3   Human E3 ubiquitin-protein ligases

Human E3 ligases perform the final step in the ubiquitin pathway, having a vital role in the recognition of specific protein substrates and mediating ubiquitin transfer from E2 enzymes to specific protein substrates. Among all of the components of the ubiquitin cascade, they are the most numerous and display the greatest diversity. Even though E3s are heterogeneous, they can nevertheless be classified into two primary classes according to domain homology and mechanism of action: HECT domain and RING finger-containing E3s (Glickman and Ciechanover, 2002).

1.4.3.1 HECT domain E3s

HECT domain proteins harbour a 350aa sequence homologous to the E6-AP carboxyl-terminal domain. This domain contains a conserved catalytic Cys residue that transfers the activated ubiquitin from an E2 to an internal Cys residue within the E3 before conjugation of ubiquitin to an NH_2 group in the substrate (Glickman and Ciechanover, 2002). Indeed the HECT domain has at least four biochemical activities: (1) it binds specific E2s; (2) it accepts ubiquitin from the E2, forming a ubiquitin-thioester intermediate with its active-site cysteine; (3) it transfers ubiquitin to the ε-amino groups of lysine side chains on the substrate by catalysing the formation of an isopeptide bond; and (4) it transfers additional ubiquitin molecules to the growing end of the multi-ubiquitin chain (Wang *et al.*, 1999).

1.4.3.2 RING domain E3s

RING finger proteins were first thought to play a role in the dimerisation of proteins. It was only in the late nineties that RING finger domains were identified as ubiquitin ligases (Lorick *et al.*, 1999). The RING finger proteins have been defined by a pattern of conserved Cys and His residues that form a cross-brace structure allowing the binding of Zinc cations. The conserved RING finger consensus is CX2CX(9-39)CX(1-3)HX(2-3)C/HX2CX(4-48)CX2C (Borden and Freemont, 1996; Joazeiro and Weissman, 2000). RING finger

domains fall into three categories: RING-HC, RING-H2 and RING-CH depending on whether a Cys or His occupies the fifth coordination site respectively. While RING fingers are structurally diverse, all contain two interleaved Zinc-binding sites (Glickman and Ciechanover, 2002). Unlike HECT E3s, RING E3s do not have recognisable catalytic active sites that define classical enzymes. Instead, these E3s have large binding interfaces and act as scaffold proteins that bring together the participant E2 and substrate proteins.

1.5 Hierarchical structure and specificity of the ubiquitin system

The structure of the ubiquitin system appears to be hierarchical. In human cells, only two main E1 proteins carry out the activation of ubiquitin required for all modifications. These enzymes are responsible for transferring ubiquitin to different E2 enzymes. Subsequently, each E2 acts in conjunction with either one or more different E3 enzymes. Beside the two main E1 activating enzymes and approximately 40 types of E2 conjugating enzymes, there are about 600-800 E3 ligases in human cells, in other word there are nearly 50 000 different potential combinations might operate in ubiquitin cascades, making the ubiquitin pathway both versatile and highly complex (illustrated in Figure 1.6). The diversity of ubiquitin, besides being contributed by more than 600 E3 ligases, is further expanded by the generation of different forms of ubiquitin chains. Even though the type of ubiquitin chain formation is depend on the E2 enzyme, it is still possible that the E3 ligase determines which substrate to bind ubiquitin to (Nakada *et al.*, 2010).
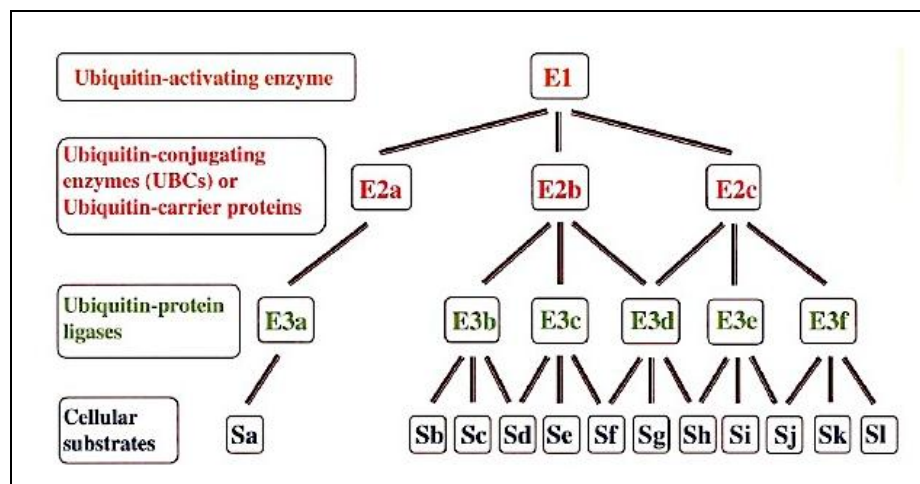
Figure 1.6 **Illustration of the hierarchical structure and specificity within the ubiquitin system:** One E1 (coloured orange) interacts with several different E2s (labelled E2a, E2b, E2c and coloured red), in which each of them may interact with multiple different E3s (labelled E3a to E3f and coloured green). Each E3 may target several different substrates (labelled Sa to Sl, coloured black).

1.6   The deubiquitinating enzyme

1.6.1   Classification of deubiquitinating enzyme

Deubiquitinating enzymes (DUBs) belong to the superfamily of proteases and they function to remove covalently attached ubiquitin from proteins, either from the polyubiquitin chain or from the substrate (Wilkinson, 1997). Among all ubiquitin machineries, the functions, targets and regulation of DUBs are the most poorly understood due to their non-uniform structure and function, though it is becoming increasingly apparent that DUBs regulate various cellular processes (Ventii and Wilkinson, 2008). The human genome encodes approximately 95 DUBs, which fall into five major classes (Nijman *et al.*, 2005) namely ubiquitin specific proteases (USPs), ubiquitin C-terminal hydrolases (UCHs), ovarian tumour proteases (OTUs), Josephins and the Jab1/MPN/MOV34 metalloenzymes (JAMMs, also known as MPN$^+$). The USP, UCH, OTU and Josephin families are cysteine proteases, whereas the JAMM/MPN$^+$ family members are zinc metalloproteases. Figure 1.7 show the structures of the catalytic domains of all five subclasses of the DUBs. For the cysteine proteases, the enzymatic activity relies on the thiol group of a cysteine in the active site. The adjacent histidine, which is polarised by an aspartate residue helps in cysteine deprononation hence these three make up the catalytic triad. During catalysis, the cysteine performs a nucleophilic attack at the peptide bond between the target and the ubiquitin resulting in the release of the target protein and formation of a covalent intermediate with the ubiquitin moiety. Upon reaction of this intermediate with a water molecule, free enzyme and ubiquitin are released (Nijman *et al.*, 2005). On the other hand, metalloproteases generally use a Zn^{2+} bound polarised water molecule to generate a non-covalent intermediate with the substrate. The metal atom is stabilised by an aspartate and two histidine residues (Ambroggio *et al.*, 2003). The intermediate is further broken down by proton transfer from a water molecule causing the release of the DUB (Nijman *et al.*, 2005).

In spite of what appears to be a hierarchical system, an interaction between a DUB called OTUB1 and E2 enzymes was observed (Markson *et al.*, 2009). This is really against the

conventional hierarchy and therefore much attention should be paid to elucidate the purposes of such interaction.

1.6.1.1 USP domain

The largest and most diverse cysteine protease, the domain contains well-conserved Cys and His boxes, which include all the catalytic triad residues as well as other residues in the active site pocket (Amerik and Hochstrasser, 2004).

1.6.1.2 UCH domain

Generally small proteins and were originally identified by their ability to hydrolyse small amides and esters at the C-terminus of ubiquitin (Amerik and Hochstrasser, 2004).

1.6.1.3 OTU domain

A novel family of cysteine proteases which display structural similarity in a presumed catalytic core domain containing conserved Cys, His and Asp residues thought to comprise the proteolytic/catalytic triad. OTU has been proven to have DUB activity by the ability of its prominent members, Otubain 1 and 2, to cleave ubiquitin from either a ubiquitin–GFP fusion protein or a tetraubiquitin fusion (Balakirev *et al.*, 2003).

1.6.1.4 Josephin domain

Its representing member, Ataxin-3 has the typical properties of DUBs: the enzyme disassembles ubiquitin–lysozyme conjugates, cleaves ubiquitin-7-amido-4-methylcoumarin (ubiquitin-AMC), and binds to the DUB inhibitor ubiquitin aldehyde (Ubal) (Burnett *et al.*, 2003). The Josephin domain, which is found in over 30 predicted proteins, most of unknown function, includes segments that show weak similarity to the His and Cys boxes of UBPs and UCHs, suggesting that this region of Ataxin-3 and its relatives will also assume the papain-

like protease fold that characterises the other cysteine proteases (Amerik and Hochstrasser, 2004).

1.6.1.5 JAMM domain

The MPN$^+$/JAMM motif DUB called AMSH (associated molecule with the SH3 domain of STAM) was found to have deubiquitinating activity as well (McCullough *et al.*, 2004). This metalloprotease motif includes two absolutely conserved His residues and an Asp residue that together coordinate a zinc ion important for proteolytic activity.
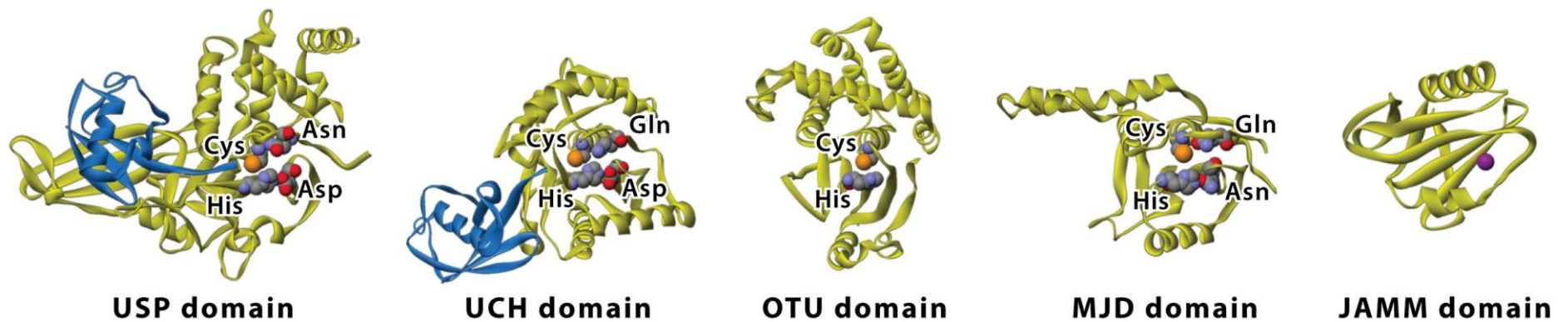
Figure 1.7   Structures of the catalytic domains of the five subclasses of DUBs (yellow) with ubiquitin (blue): Structures show the remarkable variability in secondary structure between the DUB classes. Catalytic centers are shown as Van der Waals spheres (carbon, grey; nitrogen, blue; oxygen, red; sulphur, orange; zinc, purple) and have been aligned for easy comparison. The OTU domain of OTUB2 lacks the conserved Asp in the catalytic centre and the Asn/Glu/Gln that is normally used to stabilise the oxyanion hole in these proteases. Picture from Nijman *et al.* (2005).

1.6.2   General function of DUB

Generally, DUBs activity can be divided into three major functional categories. Firstly, the DUBs are responsible for processing linear ubiquitin precursor proteins into single ubiquitin molecules. As explained earlier, ubiquitin is encoded as a polyubiquitin gene or ribosomal fusion gene. Upon expression of ubiquitin protein, DUBs play a vital role in separating the polyubiquitin and cleave the ribosomal protein to generate a free ubiquitin (Komander *et al.*, 2009a). Mutations in several DUBs have been shown to cause ubiquitin reduction and therefore resulting in various serious defect in living cells. In yeast, deletion of DUB encoding genes, including DOA4 and UBP6 were shown to reduce the amount of monomeric ubiquitin (Swaminathan *et al.*, 1999).

DUBs also functioned to remove ubiquitin chains from post-translationally modified proteins, leading to the reversal of the ubiquitin signal hence rescuing the protein from either proteasomal or lysosomal degradation. However, if a commitment to these degradative machines has been made, DUBs act to recycle the ubiquitin released from the protesome (or lysosome), thereby maintaining the free ubiquitin pool. Any ubiquitin that is released as an oligomer could also be disassembled by DUBs (Komander *et al.*, 2009a). Thirdly, DUBs can also be used to edit the form of ubiquitin modification by trimming ubiquitin chains from the distal end of the chain (Komander *et al.*, 2009a).

The knowledge about cellular functions of DUB has increased significantly in recent years. Due to the diversity in DUBs structure and functions, the prediction of DUBs main role in cellular regulation has proved to be quite tricky. Nevertheless, essential progress that largely focuses in elucidating the role of DUBs in the area of membrane trafficking, cell signalling and regulation of nuclear events has been observed in recent years (Clague *et al.*, 2012).

1.7 Protein-protein interactions

Upon completion of Human Genome Project in 2003, approximately 23 000 genes of the human genome were identified and mapped. However, this knowledge about the entire human genome could not provide an understanding even of the basic principles in human cellular systems. Entering the post-genomic era, the importance of protein-protein interactions is becoming even more apparent in order to explain how the genetic information in the form of DNA manage to generate functions. In order to tackle this question, we need to understand how the gene products, particularly proteins, interact with each other to perform many biological functions that build a living organism. Protein-protein interactions (PPI) are fundamental to all biological processes. Determination of the PPI that take place within an organism provides a framework for understanding the links between molecular and cellular biology. Alteration of PPI are thought to be involved in the development in many diseases for example neurodegenerative disorders, cancers and infectious diseases. Hence, examination of when and how they are controlled is essential for understanding diverse biological processes and elucidating the molecular basis of disease as well as identifying potential targets for therapeutic interventions.

Various methodologies can be used to detect PPI. Each has its own strengths and weaknesses with regard to the sensitivity and specificity of the method. Generally methods for identifying interacting proteins can be divided into two types: (1) biology/biochemical methods for example co-immunoprecipation and genetic manipulation yeast two-hybrid methods; (2) biophysical methods such as NMR, ITC and crystallography.

1.7.1   Genetic *in vivo* methods

1.7.1.1 Yeast two-hybrid screen

Since its description in 1989 (Fields and Song, 1989), yeast two-hybrid (Y2H) analysis has become widely used to detect binary protein-protein interaction. In this system, a protein of interest (the bait) is fused to the DNA binding domain (BD) of a transcription factor (such as GAL4) as can be seen in Figure 1.8. The bait's potential interacting partner (the prey protein) is fused to the transcription factor's activation domain (AD). These fusions are carried out by DNA cloning methods, allowing expression of the subsequent bait and prey fusion proteins in the nucleus of the yeast host. The yeast strain used in this system carries a set of reporter constructs which are under the control of an upstream sequence containing the binding sites for the BD. If the bait-BD and prey-AD fusions interact, then a functional transcription factor is reconstituted and expression of the reporter gene is activated (Fields and Song, 1989). The reporter genes produce a scorable phenotype such as growth on selective media or colour change. For example, activation of the reporter genes *ADE2* and *HIS3*, enables growth on media lacking adenine and histidine respectively, and the activation of the *lacZ* reporter gene produces a blue colour in an X-gal assay.

The first genome-wide Y2H interaction map was generated for bacteriophage T7, and large-scale yeast two-hybrid screens have been conducted for several viruses, *H. pylori*, *S. cerevisae*, *P. falciparum*, *C. elegans*, *Drosophila* and human (Flajolet *et al.*, 2000; McCraith *et al.*, 2000; Uetz *et al.*, 2000; Guo *et al.*, 2001; Ito *et al.*, 2001; Giot *et al.*, 2003; Li *et al.*, 2004; Stanyon *et al.*, 2004; Formstecher *et al.*, 2005; LaCount *et al.*, 2005; Rual *et al.*, 2005; Stelzl *et al.*, 2005; Uetz *et al.*, 2006). The data generated from these screens will be useful for individual studies and for system wide studies (Parrish *et al.*, 2006). For example the human pathogen interaction maps have generated lead proteins that may interact during pathogenesis and identified potential drug targets (Parrish *et al.*, 2006). The potential for proteome-wide interactome mapping in humans has been demonstrated by two major studies. Rual *et al.* (2005) screened approximately 7 200 human full-length ORFs, randomly

chosen on the basis of clone availability, and identified 2 754 protein interactions. Stelzl *et al.* (2005) screened two-hybrid arrays generated from a human foetal brain cDNA library and a collection of full-length ORFs, identifying 3 156 interactions. Combined, these two datasets identified over 5 900 protein interactions, a large proportion of which are novel.

Our lab has used the Y2H system to perform both library screen and targeted matrix experiments. A library screen allows the identification of novel interactive partners, whereas a matrix approach enables the detection of an interaction between two specific proteins. A library screen is applied as a tool for the discovery of binding partners for a specific bait, by screening it against an AD-cDNA library (which contains potential cDNA prey clones fused to the GAL4 AD). Direct yeast colony PCR of the diploid yeast colonies using prey specific primers permits identification of the preys via sequencing. Meanwhile, the matrix method involves pairs of proteins which are systematically tested for a binary two-hybrid interaction. Here, arrays of yeast of one mating type containing bait clones are mated with arrays of yeast of the opposing mating type containing individual prey clones. The advantage of the matrix method is that there is no need to identify the interacting protein, as the clone is already defined by its position in the array.
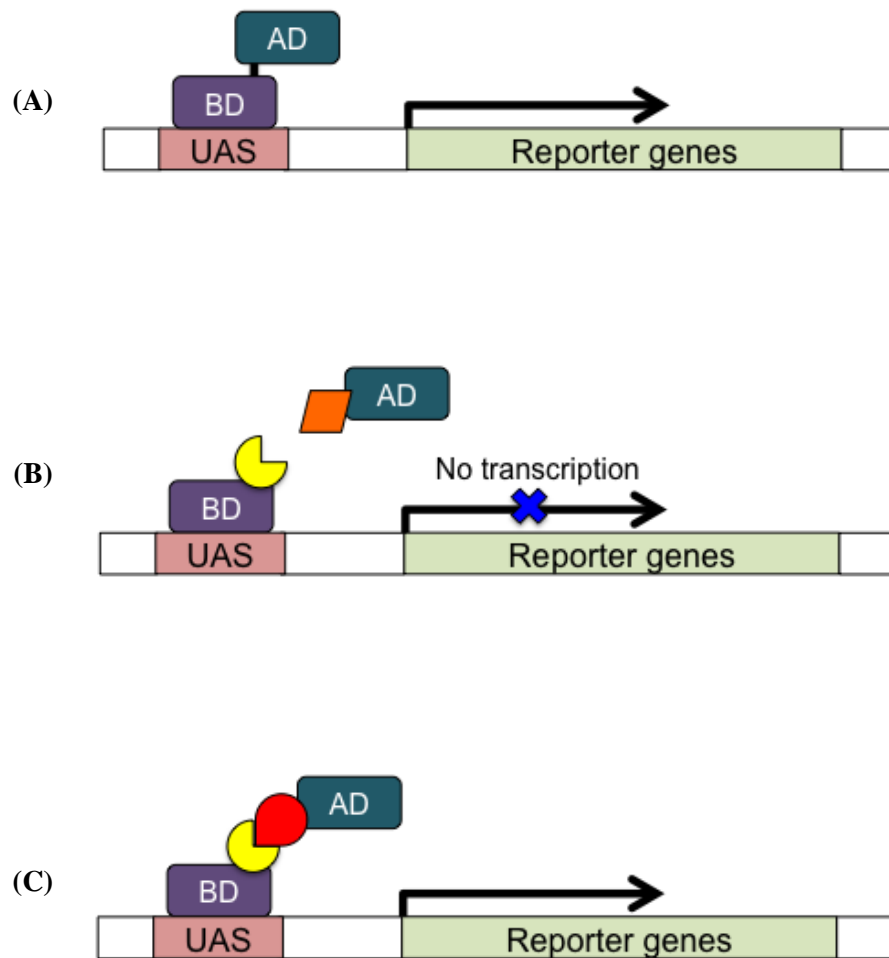
Figure 1.8   **The Y2H system:** The yeast GAL4 transcription factor can be divided into two domains, the DNA binding domain (in purple, labelled BD), and the transcription activating domain (in green, labelled AD). **(A)** The unmodified GAL4 transcription factor can activate the expression of a reporter gene by binding a GAL4-specific promoter (in pink) and confer a survival phenotype to yeast growing on selective media. **(B)** Expression of either the DNA binding domain or the transcription activating domains in isolation cannot activate reporter gene expression and therefore yeast cells growing on selective media cannot survive. **(C)** To assess the interactions between a protein X and a protein Y, fusion constructs can be co-expressed in yeast cells to produce BD-X and AD-Y chimeras. If protein X and protein Y interact the survival phenotype is restored.

1.7.2   Biophysical and theoretical methods

1.7.2.1 Isothermal titration calorimetry (ITC)

ITC has been extensively used to study protein-ligand binding and its application in protein-protein binding are growing rapidly (Leavitt and Freire, 2001). It is the only technique that determines directly the thermodynamic parameters of a given reaction which are the ΔG, ΔH and ΔS. Theoretically, the binding of two proteins cause a change in the thermodynamic potentials (ΔG, ΔH and ΔS) and these are measured directly by highly sensitive calorimetry (Velazquez-Campoy *et al.*, 2004). ITC uses stepwise injections of one protein into a calorimetric cell containing the second protein to measure the heat of the reaction for both exothermic and endothermic processes (Figure 1.9). The underlying principles of ITC lies on the chemical scheme based on a reversible association equilibrium:
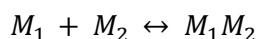
$$M_1 + M_2 \leftrightarrow M_1 M_2$$

where M_1 and M_2 are the interacting macromolecules. The strength of the interaction is described by the association constant K_a or the dissociation constant K_d :

$$K_a = [M_1 M_2]/[M_1][M_2] \text{ and } K_a = 1/K_d$$

where $[M_1]$ and $[M_2]$ are the concentrations of the free reactants and $[M_1 M_2]$ is the concentration of the complex. These constants are related to the Gibbs energy of association ΔG_a and dissociation ΔG_d and can be expressed in terms of the enthalpy, ΔH and entropy, ΔS, change in the process:

$$\Delta G_a = -RT \ln K_a = \Delta H_a - T\Delta S_a$$

$$\Delta G_d = -RT \ln K_d = \Delta H_d - T\Delta S_d$$

where R is the gas constant (1.9872 cal/Kmol) and T is the absolute temperature (kelvins).

Among the information available by doing ITC are simultaneous determination of the association constant, the enthalpy of binding and the stoichiometry.

An ITC composed of two identical cells made of a highly efficient thermal conducting and chemically inert material, surrounded by an adiabatic jacket (O'Brien *et al.*, 2000). Sensitive thermocouple circuits are used to detect temperature differences between the reference cell filled with buffer and the sample cell filled with the first protein. Before the addition of second protein, a constant power is applied to the reference cell. This directs a feedback circuit, activating a heater located on the sample cell (MicroCal ITC manual). During the experiment, ligand is titrated into the sample cell in known amount, causing heat to be either taken up or evolved, depending on the nature of the reaction.
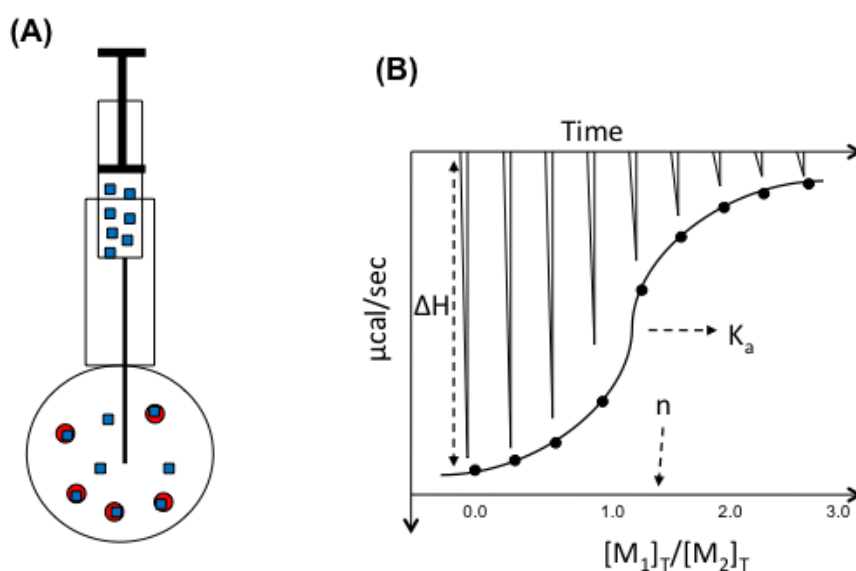
Figure 1.9   **Illustration of the configuration of an ITC reaction cell: (A)** The cell is filled with the protein 1 (red) while the injection syringe is filled with protein 2 (blue). At specified time intervals (see B) a small volume of protein 2 is injected into the cells triggering the binding reaction and producing a characteristic peak sequence in the recorded signal. **(B)** After integration of the area under each peak, the individual heats are plotted against the molar ratio from which, through non-linear regression, is possible to estimate the thermodynamic parameters such as the stoichiometry of the interaction (n), K_a and ΔH.

1.7.2.2 Nuclear magnetic resonance (NMR)

The fundamental questions in protein-protein interaction are the location of binding interface. A lot of biochemical or biophysical techniques are currently available to address this question but among the most widely used is NMR chemical shift perturbation due to its sensitivity to subtle changes in the chemical environment of proteins. NMR can provides both specific and qualitative information for examples: (1) give specific information on physical properties of individual functional groups such as ionisation states, pKa, and hydrogen bonds; (2) provide site-specific information on backbone and side-chain dynamic motions in solution; (3) be used to identify contacts between individual atoms of a protein and its binding partners, as well as to study the kinetics of ligand binding (Gao *et al.*, 2004). Table 1.3 below shows the comparison between ITC and NMR.

Table 1.3 **The comparison between ITC and NMR:** Pros and cons of ITC and NMR. NMR, however, has an extra advantage since it can provides binding site information.

	ITC	**NMR**
Measure	ΔH	$\Delta \delta$
Concentration P1	>1 μM	>10 μM
Concentration P2	>10*[P]	Any (>10*[P] opt)
Volume	0.3 ml	0.5 ml
Titration time	~30 min	~1 day
Binding parameter	ΔH, K_d, $T\Delta S$	K_d

Chemical shift arises from anisotropies in the electronic environment surrounding magnetic nuclei. Some nuclei like ^{1}H, ^{13}C, ^{19}F, ^{31}P have nuclear spin that can create a magnetic moment. By placing these nuclei in a magnetic field, they absorb electromagnetic radiation at particular frequencies governed by their chemical environment and give an electronic signal displayed as a peak (Figure 1.10). Their chemical environment is influenced by various factors such as chemical bonds, molecular conformations and dynamic processes. So

even the nuclei of the same element give rise to distinct spectral lines. The relative positions of these spectral lines are therefore called chemical shifts. In other words, upon addition of ligand or protein partner, shifting peaks, line width, and/or intensity changes indicate which residues experience a change in their environment, and this in turn allow the identification of the binding interface. In practice, this can be achieved by simply titrating the unlabelled protein into an isotopically enriched (^{15}N or ^{13}C) protein sample and monitoring peak perturbations in spectra of the labelled protein (Gao *et al.*, 2004).
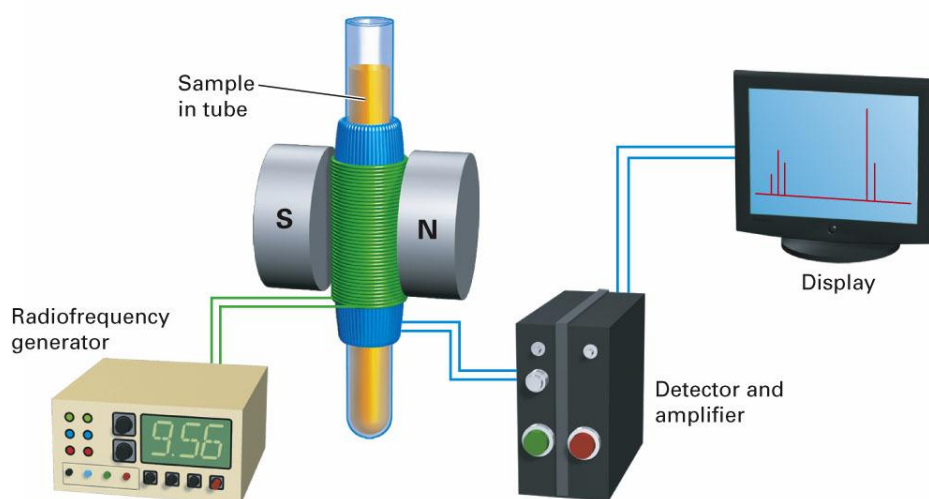
Figure 1.10 **Schematic operation of basic NMR spectrometer:** Data is usually collected from 300-600 µl protein in concentration range of 0.1-0.3 mM. The sample is placed in a thin glass tube between the poles of a strong magnet and irradiated with rf energy. A sensitive detector monitors the absorption of rf energy, and the electronic signal is then amplified and displayed as a peak.

1.7.2.3 X-ray crystallography

X-ray crystallography is the oldest technique in determining protein structure with a vast number of Nobel prize medals and contribution to majority of PDB depositions. It is a method of determining the atom arrangement within a crystal, in which a beam of X-rays strikes a crystal and causes the beam of light to scatter into many specific directions. A 3D picture of the electron density within the crystal can be determined from the angles and intensities of these diffracted beams. From this electron density, the mean positions of the atoms in the crystal can be determined, as well as their chemical bonds, their disorder and various other information. In the past decade, biological crystallography has experienced an improvement in data collection and structure and has deposited more than 58 000 crystal structures into the PDB (Ennifar, 2012).

Proteins can be prompted to form crystals when placed in the appropriate conditions. In order to crystallise a protein, the purified protein undergoes slow precipitation from an aqueous solution. As a result, individual protein molecules align themselves in a repeating series of unit cells by adopting a consistent orientation. The crystal that forms is held together by non-covalent interactions (Rhodes, 1993). The complete step-by-step method of protein crystallography is illustrated in Figure 1.11.

Protein in solution needs to be incubated with precipitants in order to form crystal. The general function of precipitants in protein crystallisation experiments is to decrease the solubility of the protein. In theory, the precipitants compete with the protein solutes for water, thus leading to supersaturation of the proteins (McPherson, 2001). Salts, polymers, and organic solvents are the most popular precipitants.
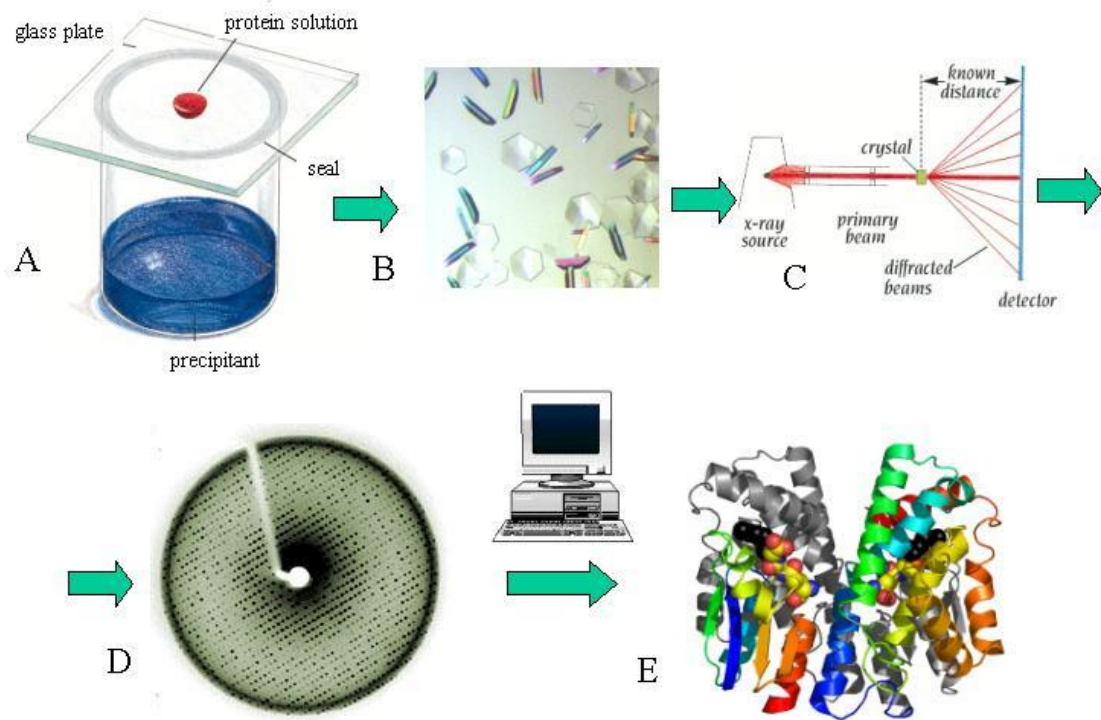
Figure 1.11   **Steps of protein X-ray crystallography: (A)** and **(B)** Proteins in solution are set up with precipitant(s) to obtain a good-diffraction crystal. **(C)** The electrons surrounding the molecule diffract as the X-rays hit them and **(D)** forms a pattern called the 'X-ray diffraction pattern' as the electrons leave. **(E)** The diffraction data is then processed by Fourier transform and the structure of the atom or molecule is deduced and visualised.

1.8   Thesis overview and aims

Research advancement in the ubiquitin field has developed rapidly since its inception. Having its major role protein degradation extensively explored, the current interest is now focused on elucidating the non-degradative functions covering every cellular process governed by ubiquitination. The expansion of ubiquitin research has also revealed the non-classical partnership among the ubiquitin enzymes for example the interaction between E2 and DUB. In this thesis, I aim to investigate and discuss the behaviour of interactions between the E2, particularly UBE2D2 and the DUB represent by OTUB1 in the human ubiquitination system. In the first results chapter, I will concentrate on a Y2H assay between E2 and DUB clones that has been established in our lab. The next chapter will focus on validation of the Y2H data by means of biophysical analysis and the discussion of my unfortunate attempt to have the structure of OTUB1:E2 complex solved. Following this, the third results chapter will discuss on Y2H targeted analysis on selective binding residue mutants UBE2D2 and co-localisation experiment in order to get a better understanding about the characteristics of the OTUB1:E2 partnership. Finally, in the last chapter, I will put emphasis on the discussion of OTUB1:UBE2D2 complex structure which has been solved in recent literature and how our Y2H targeted analysis matched this finding.

Chapter Two:
MATERIALS AND METHODS

2.1   Introduction

This project employed a range of experimental methods including: molecular biology, yeast two-hybrid, protein biochemistry, biophysical analysis of purified proteins and protein complexes, cell biology and bioinformatic analysis of protein interaction networks and pathway prediction methods. For ease of presentation, these methods have been divided into six general areas: Y2H methods, molecular biology, protein expression methods, biophysical procedures, cell biology methods and bioinformatic methods. Experimental procedures are depicted in detail in this chapter while explanation about the principle or limitations of particular methods or reagents are addressed more comprehensively in succeeding chapters.

2.2   Preparation of clones for Y2H screening

2.2.1   Reagents

Human KOD Hot Start DNA Polymerase was obtained from Novagen (distributed by Merck Chemicals Ltd., Nottingham, UK). DNA HyperLadder, agarose and BIOTaq DNA polymerase were obtained from Bioline (London, UK). Sets of dNTPs (100 mM solutions) were obtained from GE Healthcare (Buckhinghamshire, UK). SYBR® Safe DNA gel stain was from Invitrogen (Paisley, UK). TBE buffer was obtained from VWR International Ltd (Lutterworth, UK). Gel extraction purification kit was from Qiagen (Crawley, UK). Peptone, yeast nitrogen base w/o amino acids and glucose were from Formedium (Norfolk, UK). Yeast extract and BioAgar were from BioGene (Cambridge, UK). Salmon testes DNA and primers were from Invitrogen (Paisley, UK). Polyethylene glycol (PEG), all amino acids and other chemicals were from Sigma-Aldrich (Poole, UK) unless otherwise stated. The yeast two-hybrid host strain used for bait clones in conventional yeast two-hybrid experiments was

PJ69-4A (*MATa trp1-901*, *leu2-3*, *112 ura3-52*, *his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*), while prey clones were constructed using the complementary PJ69-4α strain. Both strains were kindly provided by Phil James, (University of Wisconsin, USA).

2.2.2 Proofreading PCR-amplification of protein coding inserts from pDONR223 entry clones

A selected of pDONR223 entry clones containing specific DUB expression clones were PCR-amplified using primers flanking with Gateway®- and yeast- sequence to allow *in vivo* gap repair cloning into *Bam*HI linearised pACTBE-B or pGBAE-B vectors as described by Markson *et al.* (2009). Primers were typically designed to be 18-22 bp long with melting temperatures (T_m) around 60°C. PCR mastermix was prepared on ice by mixing the PCR reaction in sterile 1.5 ml Eppendorf tubes. The mastermix was vortexed at full speed for 2 second before aliquoting into thin-walled PCR plates or PCR tubes. pDONR223 constructs containing ORFs of interest were pipetted separately into each aliquot and PCR cycle was performed on thermocycler (Table 2.1). The PCR conditions were occasionally optimised to increase product yield by increasing cycle number or adjusting the annealing temperature when using different primer combinations. Typically the annealing temperature is about 3-5°C below the T_m of the primers used. All PCR products were evaluated by agarose gel electrophoresis to ascertain correct band size.

Table 2.1   Reaction mixture and cycling parameters used for typical KOD Hot Start PCR reactions: Components for a single PCR reaction are shown in (A). For multiple samples, mastermix was prepared by multiplying the volume (excluding DNA template) with total number of reactions. Typical PCR reaction parameters are shown in (B).

(A)

	Ingredient	Volume	Final conc.
1	DNA template	1.0 µl	1-5 ng/µl
2	Forward primer (10 mM)	0.5 µl	0.5 mM
3	Reverse primer (10 mM)	0.5 µl	0.5 mM
4	dNTPs (32 mM total)	2.5 µl	0.96 mM
5	KOD Hot Start buffer	2.5 µl	1x
6	$MgSO_4$ (50 mM)	1.5 µl	2.5 mM
7	KOD Hot Start DNA pol.	0.5 µl	-
8	dH_2O	15.0 µl	-
	Total	24.0 µl	-

(B)

Segment	Cycles	Temperatures	Duration
1	1	94°C	2 min
2	30	94°C	30 s
		60°C	30 s
		70°C	1 min
3	1	70°C	5 min

2.2.3   Agarose gel electrophoresis

0.8% to 1.2% (w/v) agarose gels were prepared by adding appropriate amount of agarose into 0.5x TBE buffer before heating in a microwave until completely dissolved. Gel mixtures were allowed to cool before adding 5 µl of SYBR® Safe DNA gel stain to per 100 ml agar. Gels were then left to set at RT before being submerged in 0.5x TBE buffer in an electrophoresis tank (BioLine, UK). 5 µl of PCR product were mixed with 2 µl Orange G (5% (w/v) sucrose, 0.05% (w/v) orange G) sample buffer before loading into each well. 5 µl of HyperLadder marker was loaded into the first lane on each gel and electrophoresis was

performed at 100 V for 30 minutes to 1 hour to resolve PCR products before visualising DNA bands on an ultraviolet light (UV) source.

2.2.4 Purification of PCR products by gel extraction

DNA fragments were excised from agarose gels using a clean scalpel and the gel slices were trimmed to minimise gel volume before being transferred into 1.5 ml Eppendorf tubes. After weighing, three volume of QG neutralisation buffer (5.5 M guanidine thiocyanate, 20 mM Tris-HCl pH 6.6) was added to one volume of gel (100 mg≈100 μl) and gel slices were dissolved by heating at 50°C for 10 minutes. Tubes were vortexed every 2-3 minutes during the incubation period in order to speed up gel solubilisation. The resulting sample was then loaded onto a spin column with a 2 ml collection tube before being centrifuged for 1 minute to bind the DNA and remove excess buffer. The column was washed twice with 750 μl of PE wash buffer (20 mM NaCl, 2 mM Tris-HCl pH 7.5, 80% (v/v) ethanol) to remove excess salt. 50 μl of dH_2O was added to elute bound DNA by spinning for 1 minute at 13 500 rpm in a bench top centrifuge. Purified PCR products were either used directly for yeast transformation or stored at -20°C for future use.

2.2.5 Yeast media preparation

Media required for yeast transformation and Y2H screening protocols were prepared in-house from individual components. The list below provides a brief description of the constituents of all types of media used in these studies:

Table 2.2   List of Y2H medium and their uses.

	Name	**Description**	**Used for:**
1	YPAD	Nutrient rich **Y**east Extract-**P**eptone-**A**denine-**D**extrose	Routine growth of yeast strains and mating of haploid bait and prey clones
2	SD-W	**S**ynthetic **D**efined medium, deficient in tryptophan	Bait (pGBAE-B) selection and growth
3	SD-U	Deficient in uracil	Bait (pGBDU-GW) selection and growth

4	SD-L	Deficient in leucine	Prey pACTBE-B selection and growth
5	SD-W, low adenine	Deficient in tryptophan, insufficient amount of adenine hemisulphate	Selection and growth of Mat-a (bait) clones
6	SD-L, low adenine	Deficient in leucine, insufficient amount of adenine hemisulphate	Selection and growth of Mat-α (prey) clones
7	SD-WH (3-AT)	Deficient in tryptophan and histidine, supplemented with 3-aminotriazole (3-AT)	Testing for *HIS3* autoactivation when using pGBAE-B (bait) constructs
8	SD-LH (3-AT)	Deficient in leucine and histidine, supplemented with 3-AT	Testing for *HIS3* autoactivation when using pACTBE-B (prey) constructs
9	SD-WA	Deficient in tryptophan and adenine	Testing for *ADE2* autoactivation in bait pGBAE-B construct
10	SD-LA	Deficient in leucine and adenine	Testing for *ADE2* autoactivation when using pACTBE-B prey constructs
11	SD-WL	Deficient in tryptophan and leucine	Select of diploid clones containing pGBAE-B (bait) and pACTBE-B (prey) constructs
12	SD-UL	Deficient in uracil and leucine	Select of diploid clones containing pGBDU-GW (bait) and pACTBE-B (prey) constructs
13	SD-WLH (3-AT)	Deficient in tryptophan, leucine and histidine, supplemented with 3-AT	Less stringent triple selection media, used to select positive interactions between pGBAE-B (bait) and pACTBE-B (prey) clones
14	SD-WLA	Deficient in tryptophan, leucine and adenine	More stringent triple selection media, used to select positive interaction between pGBAE-B (bait) and pACTBE-B (prey) clones
15	SD-ULH (3-AT)	Deficient in uracil, leucine and histidine, supplemented with 3-AT	Less stringent triple selection, used to select positive interaction between pGBDU-GW (bait) and pACTBE-B (prey) clones
16	SD-ULA	Deficient in uracil, leucine and adenine	More stringent triple selection, used to select positive interaction between pGBDU-GW (bait) and pACTBE-B (prey) clones

Table 2.3   **Recipes for basic yeast media: (A)** YPAD and **(B)** SD-X, all media were made up with 500 ml dH$_2$O and autoclaved. Media were allowed to cool before pouring into 50 or 90 mm petri dishes which were left at room temperature to set.

(A) YPAD	(B) SD-X
10 g D-glucose	10 g D-glucose
10 g Peptone	3.35 g Yeast Nitrogen Base
5 g Yeast extract	Appropriate amount of amino acid mix (Table 2.4A and 2.4B)
0.1 g Adenine hemisulphate	
10 g BioAgar (for solid medium)	10 g BioAgar (for solid medium)

Table 2.4   **SD-X amino acids supplement mixes: (A)** Media for yeast transformation and autoactivation test. **(B)** Media for double and triple selection on Y2H screening. A/H/L/W/U DO is listed in Table 2.5.

(A)	SD-X mix (Grams/10 L)								
	-W	-W*	-U	-L	-L*	-WA	-WH	-LA	-LH
Adenine	0.6	0.06	0.6	0.6	0.06	-	0.6	-	0.6
Histidine	0.2	0.2	0.2	0.2	0.2	0.2	-	0.2	-
Leucine	1.0	1.0	1.0	-	-	1.0	1.0	-	-
Uracil	0.2	0.2	-	0.2	0.2	0.2	0.2	0.2	0.2
Tryptophan	-	-	**	**	**	-	-	**	**
A/H/L/W/U DO	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3
Total g/L	**0.73**	**0.73**	**0.71**	**0.63**	**0.63**	**0.67**	**0.71**	**0.57**	**0.61**

*Low adenine media for selecting positive transformants.

**20 mg/l filter-sterilised Trp was added after autoclaving the media to avoid Trp degradation by high heat.

(B)	SD-X mix (Grams/10 L)					
	-WL	-UL	-WLA	-WLH	-ULA	-ULH
Adenine	0.6	0.6	-	0.6	-	0.6
Histidine	0.2	0.2	0.2	-	0.2	-
Leucine	-	-	-	-	-	-
Uracil	0.2	-	0.2	0.2	-	-
Tryptophan	-	**	-	-	**	**
A/H/L/W/U DO	5.3	5.3	5.3	5.3	5.3	5.3
Total g/L	**0.63**	**0.61**	**0.57**	**0.61**	**0.55**	**0.59**

Table 2.5   **Dropout (DO) amino acid mixture:** DO mix was added to all SD-X media with the appropriate amino acids as stated above (Table 2.4A and 2.4B).

A/H/L/W/U DO	
Amino acid	**Grams/100 L**
Arginine	2 g
Isoleucine	3 g
Lysine	3 g
Methionine	2 g
Phenylalanine	5 g
Threonine	20 g
Tyrosine	3 g
Valine	15 g
Total (g/L)	**0.53**

2.2.6   Yeast transformation / gap repair

2 ml of YPAD broth was inoculated with fresh Mat-a or Mat-α yeast and grown overnight in a shaking incubator at 30°C, 220 rpm. An additional 8 ml of YPAD was added next morning and the cells were incubated for a further 5 hours at 30°C, 220 rpm. Cells were then harvested by centrifugation at 2 300 rpm for 5 minutes in a bench top centrifuge and the supernatant was discarded. Cell pellets were resuspended in 5 ml 100 mM LiOAc and 1.5 ml of each was transferred to a 1.5 ml Eppendorf tube. Cells were again harvested by centrifugation at 2 300 rpm for 5 minutes before being resuspended in 320 μl transformation mixture containing 230 μl 50% (w/v) PEG 3350, 35 μl 1 M LiOAc, 45 μl dH_2O, 9 μl of 10.5 mg/ml denatured salmon testes DNA and 20 ng/μl *Bam*HI-linearised yeast vector. The mixture was vortexed and transferred into 10 individual PCR tubes (32 μl per tube) and 4 μl of purified target DNA was added to each tube before mixing. Samples were then incubated in a thermal cycler at 30ºC for 30 minutes, 42ºC for 25 minutes and 30ºC for 1 minute. Transformation reactions were then plated onto appropriate SD-X selection medium (see

Table 2.2) containing low adenine before being incubated for three to five days at 30°C to allow the growth of red and white colonies (Semple *et al.*, 2005).

2.2.7 Diagnostic yeast colony PCR (YC-PCR)

Red colonies indicating positive gap repaired reactions were picked for analysis by YC-PCR in order to confirm correct insert size. For each reaction, 30 μl dH$_2$O and 3 μl 0.02 M NaOH were prepared in separate PCR tubes/wells. PCR mastermix was set up on ice and mixed by vortexing at full speed for 2 seconds. Red/pink yeast colonies were then picked with sterile toothpicks, dipped three times into 30 μl dH$_2$O before being resuspended in 3 μl 0.02 M NaOH at RT in order to the lyse yeast cells. After 10-15 minutes, 12 μl of PCR mastermix was aliquoted into each tube containing lysed cells, giving a final reaction volume of 15 μl. YC-PCR reactions were then cycled in thermocycler as described below in the table 2.6B:

Table 2.6   **Reaction mixture and cycling parameters for a typical YC-PCR: (A)** Ingredients for YC-PCR reaction mixture for one sample. **(B)** The typical cycling parameters conditions were occasionally altered to suit with the primers T_m.

(A)		Ingredient	Volume	Final conc.
	1	Forward primer (10 mM)	0.75 μl	0.5 mM
	2	Reverse primer (10 mM)	0.75 μl	0.5 mM
	3	dNTPs (32 mM total)	0.45 μl	0.96 mM
	4	Taq pol. buffer	1.5 μl	1x
	5	$MgCl_2$ (50 mM)	0.75 μl	2.5 mM
	6	Taq pol.	0.15 μl	-
	7	DMSO	0.3 μl	-
	8	dH_2O	7.35 μl	-
	9	NaOH yeast suspension	3.0 μl	-
		Total	15.0 μl	-

(B)	Segment	Cycles	Temperatures	Duration
	1	1	95°C	5 min
	2	35	95°C	1 min
			68°C	1 min
			72°C	3 min 30 s
	3	1	72°C	5 min

2.2.8   Autoactivation assay

Following verification by YC-PCR, 5 μl of the corresponding inoculated dH_2O sample (see section 2.2.6) were spotted onto appropriate SD-X plates: single dropout (e.g. SD-L) for colony growth and double dropout plates lacking interaction reporter (e.g. SD-LA, SD-LH(3-AT)) for autoactivation testing. Plates were then incubated at 30°C for up to 14 days. Any growth of haploid yeast on double dropout plates is indicative of autoactivation, hence these clones were discarded from further Y2H experiments. Non-autoactivating clones were maintained on appropriate media and archived as glycerol stock for future use.

2.2.9 Glycerol stocks

Fresh colonies were picked and inoculated into 200 µl SD-X broth medium in a 1.5 ml Eppendorf tube. They were grown with shaking at 30°C, 220 rpm for 24 hours before adding 80 µl of 80% (v/v) autoclaved glycerol to each culture, which was then thoroughly mixed before storing at -80°C.

2.2.10 Y2H matrix mating

Yeast clones from glycerol stocks were spotted onto appropriate SD-X agar and allowed to grow for three days at 30°C. Alternatively yeast colonies were picked from fresh single dropout autoactivation plates. After three days (or directly after autoactivation test), healthy yeast colonies were picked using sterile toothpicks and resuspended in 200 µl dH$_2$O (until dH$_2$O turned cloudy, similar opacity in every experiment assessed by eyes) in 1.5 ml Eppendorf tube. 2 µl was then spotted onto a YPAD plate in a 8x12 grid format (96 x 2 µl = 192 µl, hence 200 µl suspension needed for each prey clones). The step is repeated for every Mat-α (prey) or Mat-a (bait) clone using a different YPAD plate. In the meantime, each transfected Mat-a bait yeast colonies were resuspended in X µl of dH$_2$O in 96-well Y2H mating plate to similar opacity so each well representing different bait clones. 2 µl from each 96 well were transferred on YPAD on top of the prey spots plate using multichannel pipette. (X = *2 µl x total amount of YPAD plates with prey spots x 2*; for example if there were 10 YPAD plates, X wil be 2 µl x 10 x 2 = 40 µl). All the YPAD plates were allowed to dry at RT before being incubated at 30°C for 24 hours. On the following day, yeast grown on the YPAD plates were replica-plated onto double selection SD-X plates using a sterilised velvet cloth in order to select for diploid cells. Double selection plates were incubated at 30°C for another two days and then transferred by velvet replication onto triple selection SD-X plates to select growth reporter activators. Yeast colonies were allowed to grow on triple selection SD-X plates at 30°C for 7-10 days before growth was scored as will be discussed comprehensively in Chapter 3.

2.3   Molecular biology

2.3.1   Reagents

α-select silver efficiency cells (F⁻ *deo*R *end*A1 *rec*A1 *rel*A1 *gyr*A96 *hsd*R17(r$_k^-$, m$_k^+$) *sup*E44 *thi*-1 *pho*A Δ(*lac*ZYA-*arg*F)U169 Φ80*lacZ*ΔM15 λ⁻) were from BioLine (London, UK). Gateway® BP and LR Reaction Kits, Champion™ pET151 Directional TOPO® Expression Kit and DH5α subcloning efficiency (F⁻ *end*A1 *rec*A1 *rel*A1 *gyr*A96 *hsd*R17(r$_k^-$,m$_k^+$) *sup*E44 *thi*-1 *pho*A Δ(*lac*ZYA-*arg*F)U169 Φ80*lacZ*ΔM15 λ⁻) were from Invitrogen (Paisley, UK). Antarctic phosphatase, all restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs (Hertfordshire, UK). QuikChange® II Site-Directed Mutagenesis Kits were from Agilent Technologies (Berkshire, UK). Miniprep kits were from Promega (Southampton, UK). Midi and maxiprep purification kits were from Qiagen (Crawley, UK). Tryptone was from Fisher (Loughborough, UK). Primers were purchased from either Invitrogen (Paisley, UK) or Sigma Aldrich (Poole, UK) and all chemicals were obtained from Sigma-Aldrich (Poole, UK) unless otherwise stated.

2.3.2   Cloning

2.3.2.1 Gateway® LR reaction to generate expression vector

LR reactions were performed to transfer E2 and OTUB1 ORFs and mutant constructs from pDONR207 and pDONR223 entry vectors into required destination vectors including pc-Myc, pDEST27, pEGFP-N2 and pG-cherry-mGR. For each reaction, approximately 200 ng of entry clone and destination vector were combined with 2 μl of LR clonase, 2 μl of clonase buffer and dH$_2$O to bring the total volume to 10 μl. The LR reaction mixture was then briefly centrifuged and incubated at 25°C for 16 hours. On the following day, 0.5 μl of proteinase K was added to the reaction mixture before incubating at 37°C for 10 minutes. 2 μl of the reaction mix was then used transform chemically competent α-select cells.

2.3.2.2 Conventional cloning into the pETM-11vector

2.3.2.2.1Target DNA amplification and restriction digest

ORFs of interests were cloned into the pETM-11 expression vector in order to generate His-tag proteins in bacterial cells. The selected E2s and OTUB1 ORFs were amplified by KOD PCR from sequence-verified pDONR207 or pDONR223 entry clones using gene-specific primers containing *Nco*I/*Hin*dIII restriction sites. 5 µl of PCR products were analysed by agarose gel electrophoresis and the DNA bands with the correct size were purified by PCR purification kit. Meanwhile, the initial aliquot of His-tag vector pETM-11 (kindly provided by Paul Elliott, University of Liverpool) was transformed and amplified in α-select chemically competent cells and purified by maxiprep. Purified PCR products ('inserts') and the pETM-11 vector were both double digested with 2 µl *Nco*I and 2 µl *Hin*dIII in 1x NEBuffer 2 with 1x BSA at 37°C for 10 hours. The linearised vector were then treated with 2 µl antarctic phosphatase and 1x antarctic phosphatase reaction buffer for 45 minutes at 37°C to remove phosphate groups from DNA overhangs and prevent vector reannealing without an insert. Finally the reactions were heated at 65°C for 20 minutes to inactivate the restriction enzymes.

2.3.2.2.2Ligations of double digested DNA inserts and pETM-11

Double digested inserts and vectors were mixed in 1.5 ml Eppendorf tubes with a molar ratio of insert to vector of 0.5:1 in 1x ligase buffer containing 0.5 µl T4 DNA ligase. They were then incubated in a thermocycler for 16 hours at 25°C followed by heat inactivation at 65°C for 10 minutes. After incubation, approximately 1-5 µl ligation product were transformed into 25 µl of α-select chemically competent cells. The insert:vector ratio (and to some extent the DNA ligase concentration) were sometimes adjusted to 2:1 to increase the number of positive colonies as the ligation reaction for some ORFs proved to be problematic.

2.3.2.3 Directional TOPO® cloning

TOPO® cloning was performed as an alternative method to conventional pETM-11 cloning. Targeted ORFs were amplified by proofreading KOD PCR from pDONR entry vectors with forward primers flanked with an over-hanging CACC sequence at the 5' end. PCR products were resolved by agarose gel electrophoresis and correct sized DNA products were purified by gel extraction method. Approximately 5 ng of purified PCR products were then incubated with 1 μl of salt (1.2 M NaCl, 0.06 M MgCl$_2$), 1 μl 15 ng/μl pET151/D TOPO® vector (both provided in the TOPO® kit) and dH$_2$O was added to make the total volume up to 6 μl. Tubes were then mixed gently and incubated at 25°C for 15 minutes before transferring 3 μl of the reaction mix into 50 μl of ready to use OneShot® TOP10 chemically competent cells (F⁻ *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K*rps*L (StrR) *end*A1 *nup*G).

2.3.3   QuikChange® site-directed mutagenesis

PCR reactions were set up using pDONR207 entry clones containing the sequence-verified UBE2D2 ORF and two complementary primers designed to include the required base substitutions in the centre of each primer. Following thermal cycling (see Table 2.7), reactions were transferred into 1.5 ml Eppendorf tube and 1 μl *Dpn*I was added to the reactions before incubating at 37°C for 1 hour. Immediately after 1 hour, 2 μl of the *Dpn*I-digested reaction were transformed into 45 μl XL-10-Gold Ultracompetent cells (*Tet*ʳ Δ(*mcrA*)183 Δ(*mcrCB-hsdSMR-mrr*)173 *end*A1 *supE*44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac Hte* [F´ *proAB lacI*ᵠ*ZDM15* Tn*10*(Tet^r) Amy Cam^r) containing 2 μl β-mercaptoethanol.

Table 2.7 **Typical QuikChange® mutagenesis reaction setup: (A)** Reactions indicated in the table were set up on ice. Negative controls were set up without polymerase to control for incomplete digestion of the template DNA by *Dpn*I. **(B)** Reactions were cycled as indicated. Cycling parameters were occasionally altered to suit primers T_m values.

(A)

	Ingredient	Volume	Final conc.
1	DNA template	x	50 ng
2	Forward primer (10 mM)	2.5 µl	0.5 mM
3	Reverse primer (10 mM)	2.5 µl	0.5 mM
4	dNTPs (32 mM total)	1.0 µl	0.96 mM
6	Pfu Ultra buffer	5.0 µl	1x
7	Pfu Ultra DNA pol.	1.0 µl	-
8	dH$_2$O	(38-x) µl	-
	Total	50 µl	-

(B)

Segment	Cycles	Temperatures	Duration
1	1	95°C	30 s
		95°C	30 s
		50°C	1 min
2	15	68°C	2 min per kb

2.3.4   Transformation of cloning reactions into chemically competent cells

Products from mutagenesis reactions were mixed with appropriate chemically competent cells by gentle pipetting. Cells were then incubated on ice for 30 minutes before heat-shocking at 42°C for 45 seconds and subsequent incubation on ice for 2 minutes. 250 µl of SOC (2% (w/v) bacto-tryptone, 0.5% (w/v) yeast extract, 8.56 mM NaCl, 2.5 mM KCl, 10 mM MgCl$_2$ and 20 mM glucose) media was added to the transformation reaction before incubating the tubes horizontally at 37°C, 220 rpm for 1 hour. After incubation, 200-250 µl of the mixture was plated onto 2xTY* agar plates containing the appropriate antibiotic as stated in Table 2.8. Plates were then incubated at 37°C for 16 hours to allow colony growth.

* 2xTY recipe: 16 g tryptone, 10 g yeast extract, 5 g NaCl made up with 1 L dH$_2$O and autoclaved. Media were allowed to cool before adding antibiotic.

Table 2.8   **Description of bacterial plasmid and *E. coli* strains:** When plasmids were transformed into *E. coli* strains carrying their own antibiotic selection, both antibiotics were included in 2xTY agar plates.

	Plasmid name	**Selection**	**Conc.**	**Description**
1	pDONR207	Gentamicin	10 µg/ml	Entry vector
2	pDONR223	Spectinomycin	50 µg/ml	Entry vector
3	pEGFP-N2	Kanamycin	50 µg/ml	N-terminal GFP-tag
4	pG-cherry-mGR	Kanamycin	50 µg/ml	mCherry-tag vector
5	pETM-11	Kanamycin	50 µg/ml	His-tag vector
6	pET151/D TOPO®	Ampicillin	100 µg/ml	His-tag vector
	Bacterial strain	**Selection**	**Conc.**	**Description**
1	α-select cells	-		Routine propagation
2	OneShot® TOP10	-		High-copy plasmid propagation
3	XL-10-Gold Ultracompetent cells	Chloramphenicol	25 µg/ml	For extremely demanding cloning
4	OneShot® *ccd*B	-		*ccd*B vector propagation
5	BL21 Star™ (DE3)	-		Expression host
6	Rosetta™	Chloramphenicol	25 µg/ml	Expression host
7	B834 cells	-		Expression host

2.3.5   Diagnostic bacterial colony PCR (BC-PCR)

BC-PCR was performed using vector-specific primers on colonies from the transformation plates to confirm the presence of required inserts. A mastermix (Table 2.9) was prepared and briefly vortexed. Then, two sets of PCR plates/tubes were prepared; one with 20 µl of dH$_2$O and the other with 10 µl of PCR mastermix. Selected colonies on the transformation plates were picked with sterile toothpicks and dipped into 10 µl of PCR mastermix and then resuspended in 20 µl of dH$_2$O. Reactions were then cycled in a thermocycler as described in the Table 2.9B. PCR products were then run on a 1% (w/v) agarose gel for size verification.

Table 2.9   **Reaction mixture and cycling parameters for BC-PCR:** Reagents listed **(A)** were set up on ice and PCR reactions were cycled as indicated **(B)**. Cycling parameters conditions were occasionally altered to suit primer T_m values.

(A)

	Ingredient	Volume (per sample)	Final conc.
1	Forward primer (10 mM)	0.7 µl	0.7 mM
2	Reverse primer (10 mM)	0.7 µl	0.7 mM
3	dNTPs (32 mM total)	0.25 µl	0.8 mM
4	10 x Taq pol. Buffer	1 µl	1x
5	MgCl$_2$ (50 mM)	0.4 µl	2 mM
6	Taq pol.	0.05 µl	-
7	dH$_2$O	6.9 µl	-
	Total	10 µl	-

(B)

Segment	Cycles	Temperatures	Duration
1	1	95°C	5 min
2	35	95°C	1 min
		55°C	1 min
		72°C	1 min per kb
3	1	72°C	5 min

2.3.6   DNA amplification, purification and glycerol stocks

Following size verification, inoculated dH$_2$O solutions corresponding to the correct DNA band size were transferred into 10 ml 2xTY broth containing the appropriate antibiotic and incubated at 37°C, 200 rpm for 16 hours. After incubation, 200 µl of the cultures were aliquoted into 1.5 ml Eppendorf tubes and 80 µl of 80% (v/v) glycerol were added to make a glycerol stock which was then stored at -80°C for future used. Meanwhile, the rest of the culture were spun at top speed in a bench top microfuge for 20 minutes at 4°C, the supernatant discarded and the pellet was used for plasmid isolation by mini-, midi- or maxiprep kits.

2.3.6.1 Miniprep

250 µl of cell resuspension solution (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 100 µg/ml RNase A) was added to the cell pellet and mixed by vortexing. The suspension was then transferred to a sterile 1.5 ml Eppendorf tube and 250 µl of cell lysis solution (0.2 M NaOH, 1% SDS) was added. Tubes were incubated at RT for 5 minutes. 10 µl of alkaline protease solution was added to the suspension and they were then incubated for another 5 minutes. Upon addition of 350 µl of neutralisation solution (1.32 M potassium acetate pH 4.8), the reaction was spun at 13 500 rpm for 10 minutes at 20°C. The pellet was discarded and lysate was transferred into a spin column attached with a 2 ml collection tube. They were centrifuged for 1 minute to bind the DNA to the spin column membrane before washing twice with 750 µl of column wash solution (80 mM potassium acetate, 8.3 mM Tris-HCl pH 7.5, 40 µM EDTA) containing ethanol. Finally 50 µl of nuclease free water was added and the DNA was eluted by centrifugation at 13 500 rpm for 5 minutes.

2.3.6.2 Midi- and maxiprep

Plasmid isolation was performed in a large scale for low copy plasmids or when high concentrations of cleaner plasmid were required. To perform midi- or maxiprep, larger

culture volume was required. 25-100 ml culture should be sufficient for midiprep while for maxiprep, around 100-500 ml culture was required. To prepare the large-scale culture, starter culture was prepared by picking a single colony on 2xTY agar (freshly streaked from glycerol stock) or from the inoculated dH$_2$O (from BC-PCR section 2.3.5) into 2.0 ml 2xTY containing appropriate antibiotic and grown at 37°C, 200 rpm for 5-8 hours. 25 µl from the starter culture were diluted into 25 ml fresh 2xTY with antibiotic and incubated for a further 12-16 hours (overnight) at 37°C, 220 rpm. For maxiprep, up to 500 ml fresh culture could be inoculated from a 500 µl starter culture.

On the following day, the overnight cultures were centrifuged at 4 300 rpm for 20 minutes at 4°C, the supernatant discarded and pellet was resuspended with 4 ml buffer P1 (50 mM Tris-Cl 8.0, 10 mM EDTA, 100 µg/ml RNase A). 4 ml of buffer P2 (200 mM NaOH, 1% SDS) was then added and gently mixed to lyse the cells and were left at RT for 5 minutes before 4 ml of chilled buffer P3 (3 M potassium acetate pH 5.5) was added. If doing maxiprep, 10 ml of buffer P1, P2 and P3 were used instead of 4 ml. The suspension were mixed gently before poured into QIAfilter cartridges and incubated for further 10 minutes. Meanwhile, QIAGEN-tips (midiprep: QIAGEN-tip 100, maxiprep: QIAGEN-tip 500) were equilibrated with 4 ml (10 ml if doing maxiprep) of buffer QBT (750 mM NaCl, 50 mM MOPS pH 7.0, 15% (v/v) isopropanol, 0.15% (v/v) Triton® X-100). After 10 minutes incubation, the lysed suspension in QIAfilter cartridges were injected into the equilibrated QIAGEN-tips to bind the DNA to the tips' resin while the buffer passed through the resin by gravity flow. QIAGEN-tips were then washed twice with 10 ml (30 ml if doing maxiprep) buffer QC (1 M NaCl, 50 mM MOPS pH 7.0, 15% (v/v) isopropanol) and DNA was eluted in 50 ml Falcon tubes with 5 ml (15 ml if doing maxiprep) buffer QF (1.25 M NaCl, 50 mM Tris-Cl pH 8.5, 15% (v/v) isopropanol). The eluted DNA was precipitated with 3.5 ml isopropanol (10.5 ml if doing maxiprep) and then centrifuged at 4°C, 4 300 rpm for 1 hour. The supernatant was carefully decanted and the pellet was washed with 2 ml 70% (v/v) ethanol (5 ml if doing maxiprep) before spun at 4°C, 13 500 rpm for 30 minutes. After discarding the ethanol, the pellet was

left to air-dry and finally resuspended in 100 μl dH$_2$O. The DNA was transferred to a new 1.5 ml Eppendorf tube and quantified with UV spectrophotometer at 260 nm. PCR (Table 2.10) and agarose gel electrophoresis can be performed to confirm correct product size before sending them for sequencing.

Table 2.10 **Reaction mixture and cycling parameters for mini-, midi- or maxiprep-PCR:** Reagents were set up in thin-walled PCR tubes on ice as indicated in **(A)** and were cycled as in **(B)**. Cycling conditions were altered to suit primers T_m.

(A)		Ingredient	Volume	Final conc.
	1	Mini-/midi-/maxiprep DNA (diluted to 10 ng/μl)	0.4 μl	40 pg/μl
	2	Forward primer (10 mM)	0.7 μl	0.7 mM
	3	Reverse primer (10 mM)	0.7 μl	0.7 mM
	4	dNTPs (32 mM total)	0.25 μl	0.8 mM
	5	10 x Taq pol. buffer	1 μl	1x
	6	MgCl$_2$ (50 mM)	0.4 μl	2 mM
	7	BIOTaq pol.	0.05 μl	-
	8	dH$_2$O	6.5 μl	-
		Total	10 μl	-

(B)	Segment	Cycles	Temperatures	Duration
	1	1	95°C	5 min
			95°C	1 min
	2	30	55°C	1 min
			72°C	1 min per kb
	3	1	72°C	3 min

2.3.7 Sequencing

All sequencing of DNA constructs was performed by automated fluorescent DNA sequencing by GATC BioTech (London, UK) to ascertain the correct ORF insert.

2.4   Protein expression methods

2.4.1   Reagents

β-mercaptoethanol,   N,N,N',N'-Tetramethylethylenediamine   (TEMED),   ammonium

persulphate (APS) and mammalian protease inhibitor cocktail were from Sigma Aldrich

(Poole, UK). BenchMark Prestained Protein Ladder was from Invitrogen (Paisley, UK).

Protogel electrophoresis buffers (ProtoGel® 30% Acrylamide/Bisacrylamide solution

(37.5:1 w/v ratio), ProtoGel resolving buffer and ProtoGel stacking buffer were from

GeneFlow (Staffordshire, UK). Glycine and Tris-Base were from Fisher (Loughborough,

UK). Rosetta™2(DE3)pLysS Single™ Competent Cells (F⁻$ompT$ $hsdS_B(r_B^- m_B^-)$ gal

dcm (DE3) pLysSRARE2 (Cam^R) were from Merck Chemicals Ltd. (Nottingham, UK).

Isopropyl-β-D-thio-galactopyranoside (IPTG) was from Melford Laboratories Ltd. (Ipswich,

UK). Ni-NTA agarose was from Qiagen (Crawley, UK). All chromatography columns were

from GE Healthcare (Buckhinghamshire, UK) and chromatography studies were performed

using the ÄKTA™ purification platform of GE Healthcare (Buckhinghamshire, UK).

2.4.2   Protein expression and purification

2.4.2.1 Small-scale expression test

Rosetta cells were transformed with clones containing genes of interest as in Section 2.3.4. A

single colony from the transformation plate was picked and incubated in 1.5 ml 2xTY

containing appropriate antibiotic at 37°C, 200 rpm for 16 hours. On the following day, 500

μl of the culture was diluted into 10 ml media in 50 ml Falcon tubes and grown until the

OD_{600} reached 0.8 at which point the incubator was cooled down to 18°C. 0.2 mM IPTG was

added to the culture to induce protein expression (uninduced cultures without IPTG were

also prepared as a negative control). The culture was incubated at 18°C, 200 rpm overnight.

The next day, bacterial cells were sedimented by centrifugation at 4 300 rpm and 4°C for 15

minutes and resuspended in 1 ml lysis buffer A (all buffer recipes were listed in Table 2.11).

After addition of protease inhibitors (10 μl 1.0 M PMSF, 1 μl pepstatin and 1 μl leupeptin),

cells were disrupted by incubating the sample with 1 mg/ml lysozyme for 30 minutes. 1 μg/ml DNase was sometimes added at this stage if lysates were too viscous. Lysates were centrifuged at 13 500 rpm for 30 minutes and the supernatant was transferred to a fresh 1.5 ml Eppendorf. 10 μl of 50% (w/v) slurry Ni-NTA resin was added to the supernatant which was gently mixed on ice for 15 minutes. The resin was pelleted by centrifugation at 3 000 rpm for 1 minute, the supernatant was decanted off and the resin were washed several times with lysis buffer A to remove loosely-bound proteins. Finally 1 ml of lysis buffer B was added to collect the protein tagged by Ni-NTA resin by centrifuging in the same manner as in washing step. Elution was repeated two times and tagged proteins were analysed by 1D SDS-PAGE.

2.4.2.2 Large-scale protein production

Single colonies were picked from transformation plate and grown overnight in 10 ml 2xTY with antibiotics at 37°C, 200 rpm. 1 L of 2xTY was then inoculated with the overnight cultures and left at 30°C, 200 rpm until the OD_{600} reached 0.6. Protein expression was performed by incubating the culture with 0.2 mM IPTG at 18°C, 200 rpm overnight. Pellet was collected by centrifugation at 8 000 rpm, 4°C for 15 minutes before resuspended in lysis buffer A and stored at -80°C for at least 2 hours. Cells were then lysed by French press in the presence of DNAse and a protease inhibitor cocktail. Soluble and insoluble components were separated by centrifugation at 18 000 rpm, 4°C for 30 minutes. The soluble solution was loaded directly into the Ni^{2+} affinity column to bind the His-tagged proteins. The His-tag was then cleaved by overnight incubation with TEV protease (5 μg protease for 1 mg target protein) at RT. To remove the His-tag, uncleaved proteins, TEV protease and other contaminants, the sample was again injected to the Ni^{2+} column in which the His-tag, uncleaved proteins and contaminants bound to the column while the tag-free protein eluted. This eluted proteins from Ni^{2+} column were then loaded into ion exchange chromatography (IEC) in order to separate the proteins according to their pI hence any contaminants that were

not distinguished by Ni^{2+} column can be removed. IEC eluates were buffer-exchanged into storage buffer using PD10 desalting column equilibrated with the storage buffer and concentrated to a final concentration of approximately 0.1 mM to 0.5 mM using a vivaspin concentrator (MWCO 5 000 kDa). Finally, the concentrated protein was aliquoted into 1.5 ml Eppendorf tubes (500 µl per tube) before being flash-frozen in liquid nitrogen and stored at -80°C

Table 2.11 **Buffer recipes:** All buffers were made up to 1 L with dH_2O before being filter-sterilised and degassed prior to use.

	Buffer	Ingredients	pH
1	Lysis buffer A	20 mM Na_2HPO_4, 500 mM NaCl, 25 mM imidazole	7.4
2	Lysis buffer B	20 mM Na_2HPO_4, 500 mM NaCl, 500 mM imidazole	7.4
3	Q buffer A	20 mM Tris-base, 20 mM NaCl, 2 mM DTT	8.0
4	Q buffer B	20 mM Tris-base, 1 M NaCl, 2 mM DTT	8.0
5	SP buffer A	20 mM MES, 10 mM NaCl, 2 mM DTT	6.5
6	SP buffer B	20 mM MES, 1 M NaCl, 2 mM DTT	6.5
7	NMR buffer (storage buffer)	20 mM Tris-base, 50 mM NaCl, 2 mM DTT	7.4
8	Gel filtration buffer	20 mM Tris-base, 150 mM NaCl, 2mM DTT	7.4
9	ITC buffer	20 mM Tris-base, 20 mM NaCl, 1 mM TCEP, 0.02% (w/v) NaN_3	7.4

2.4.2.3 (^{1}H^{15}N) labelled protein expression

Overexpression of (^{1}H^{15}N) labelled protein was accomplished by growing a single colony from Rosetta transformants for 6-8 hours in 500 µl 2xTY with appropriate antibiotics at 37°C, 200 rpm. 10 ml 2M9 minimal media (Table 2.12) was then inoculated with 100 µl of 2xTY preculture and grown overnight at 37°C, 200 rpm. Overnight cultures were then transferred into 1 L of 2M9 minimal media and incubated at 30°C, 200 rpm until the OD_{600} = 0.6 at which stage proteins were expressed and purified using the same protocol as described for unlabelled proteins.

2.4.3   SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Gel plates from the Mini-Protean® electrophoresis system (Bio-Rad) were assembled. Resolving gel (see Table 2.13 for acrylamide concentration) were prepared and a 4% (w/v) acrylamide stacking gel was added once the resolving gel was polymerised. The gel was then submerged in 1x electrophoresis running buffer (50 mM Tris-HCl, 0.1% (w/v) SDS and 380 mM glycine). Meanwhile, protein sample were diluted in 1x sample buffer (0.04 mM Tris pH 6.8, 2% (w/v) SDS, 5% (v/v) glycerol, 2 mM β-mercaptoethanol, 0.01% (w/v) bromophenol blue) and heated at 95°C in a heat block for 5 minutes to denature the protein. Using a Hamilton syringe, proteins were loaded into wells and run at 200 V for 38 minutes alongside a protein molecular weight ladder. Gels were then removed and soaked in the Coomassie-Blue stain (50% (v/v) methanol, 7% (v/v) acetic acid, 0.1% (w/v) Coomasie Blue R250) on a rocker for at least 1 hour at RT. The dye was removed using de-stain solvent (5% (v/v) methanol, 85% (v/v) dH$_2$O, 10% (v/v) acetic acid) to visualise and evaluate the protein band.

Table 2.12   Recipe for 2M9 minimal media: Solution A was made up with 975 ml dH$_2$O, adjusted to pH 7.2-7.4 and autoclaved. It was allowed to cool before addition of Solution B. Meanwhile, Solution B was made up in 25 ml dH$_2$O, filter-sterilised and added to autoclaved Solution A to make up 1 L 2M9 media.

2M9 minimal media		
Ingredients		**Amount per L**
A	Na$_2$HPO$_4$	14.6 g
	KH$_2$PO$_4$	5.4 g
	^{15}NH$_4$Cl	1.0 g
B	Glucose	4.0 g
	Thiamine hydrochloride	20 mg
	5 g/L MnSO$_4$	0.5 ml
	37.5 g/L CaCl$_2$	0.5 ml
	1 g/L FeCl$_2$	0.5 ml

Table 2.13   Recipe for SDS-PAGE resolving and stacking gel: Table shows the differing acrylamide concentration used for resolving gels depending on the expected band sizes and the fixed 4% (w/v) acrylamide for stacking gels.

Reagent	Resolving gel (ml)		
	8%	**10%**	**12%**
Protogel acrylamide solution	2.67	3.33	4.0
Resolving buffer	2.6	2.6	2.6
dH$_2$O	4.62	3.96	3.29
TEMED	0.01	0.01	0.01
10% APS	0.1	0.1	0.1
Total Volume	10.0	10.0	10.0
Reagent	Stacking gel (ml)		
	4%		
Protogel acrylamide solution	0.65		
Stacking buffer	1.25		
dH$_2$O	3.05		
TEMED	0.005		
10% APS	0.025		
Total Volume	4.98		

2.5 **Biophysical procedures**

2.5.1 Reagents

PD10 desalting column was from Amersham (Buckinghamshire, UK). Wizard: Cryo™ and Wizard™ Screens I and II were from Emerald BioSystems (Washington, USA), ProPlex™ was from Molecular Dimensions (Suffolk, UK), Crystal Screen was from Hampton Research (California, USA). ITC system was from MircoCal Inc. (Buckinghamshire, UK).

2.5.2 Gel filtration chromatography

The protein eluted from the ion exchange column (Section 2.4.2.2) or from -80°C freezer was passed through a PD10 column which had been equilibrated with gel filtration buffer (Table 2.11). 500 μl of individual proteins were injected onto the gel filtration column (Superdex 75 10/300 GL) which was run at 1 ml/min. The step was repeated to the premix of both proteins in 1:1 followed by varying ratio to assess complex formation.

2.5.3 Isothermal titration calorimetry

All proteins were buffer-exchanged into ITC buffer (Table 2.11) using a PD10 column which had been equilibrated with ITC buffer and concentrated to 0.1 mM (protein A) and 2.1 mM (protein B). All protein samples and buffers used in this experiment were degassed for 5 minutes. ITC cell chamber was washed three times with the degassed buffer before pipetting 300 μl of 0.1 mM protein A into the chamber. Meanwhile, 80 μl of 2.1 mM protein B (or at least ten-fold from the concentration of protein A) was loaded to the syringe. The injection of sample from the syringe to the cell was performed at 180s time intervals, 3 μl per injection and the cell temperature was set at 25°C.

2.5.4 Nuclear magnetic resonance

Protein in NMR buffer was diluted or concentrated to 0.1-0.5 mM. The sample was prepared for NMR by adding 5% (v/v) D_2O to 600 μl of protein sample. Then, the sample was placed

in a clean standard NMR tube (5 mm diameter) by a long Pasteur pipette. Care was taken not to introduce any air into the NMR tube. All NMR spectra were collected at 25°C using a Bruker Avance 600 or 800 MHz solution-state spectrometer. For the chemical shift perturbation experiments, (^{1}H^{15}N)-HSQC 2D spectra were recorded on ^{15}N-labelled protein alone and titration with different ratios of its unlabelled partners (1:0.25, 1:0.5, 1:0.75, 1:1, 1:1.5, 1:2). For all (^{1}H^{15}N)-HSQC spectra, 2 048 points with a spectral width of 8 012 Hz in the direct dimension and 512 points with a spectral width of 2 200 Hz in the indirect dimension were recorded. The number of scans varied between 16 and 64 during the titration and the relaxation delays were set to 1s. All NMR spectra were processed and analysed using Bruker TopSpin 3.1.

2.5.5 Crystallisation trial

Two sets of 500 μl of OTUB1:UBE2D2 complex were prepared in 10 mg/ml and 20 mg/ml alongside 10 mg/ml free protein. 80 μl of each condition (Wizard, ProPlex, Crystal and Cryo) were pipetted into the reservoir of the MRC 96-well sitting drop vapour diffusion crystallisation plate (Molecular Dimensions). 0.5 μl of each protein samples were pipetted into drop reservoir followed by 0.5 μl from each condition. Plates were sealed with crystal sealing film and incubated at RT and the other set at 4°C for about two weeks.

2.6   Cell biology

2.6.1   Reagents

All cell culture reagents were obtained from Invitrogen unless otherwise stated. Plasticware was obtained from Corning Inc. (New York, USA) and chemicals were from Sigma unless otherwise stated.

2.6.2   Cell culture

HeLa cells were cultured in a humidified 5% (w/v) CO_2 atmosphere at 37°C in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) foetal bovine serum, 1% (v/v) non-essential amino acids (NEAA), and 1% (v/v) penicillin/streptomycin sulphate. Cells were maintained by growing in 75 cm flasks. Every three to four days when the cells reach ~80% confluency, the media was removed and cells were washed with 30°C PBS before being trypsinised with 2 ml trypsin for 2 minutes in 37°C (5% (w/v) CO_2) incubator. Cells were then split into 1:5 to 1:10 dilution into a new flask with fresh DMEM.

2.6.3   Cell transfection and fixation

80% confluent cells were trypsinised as above and 0.5×10^5 cells (counted by hemocytometer) were seeded onto coverslips in 12-well plate containing 1 ml of DMEM. They were incubated at 37°C (5% (w/v) CO_2) for 24 hours until cells reached 50%-80% confluency. On the next day, 1.5 µl Genejuice was added to 50 µl DMEM media with no additional supplements and mixtures were vortexed gently and mixed at RT for 5 minutes before adding 0.5 µg of DNA. They were mixed by gentle pipetting and incubated at RT for 15 minutes before being pipetted into 12-well plates in a drop-wise manner to ensure the whole coverslip was covered. They were then mixed by gently rocking at 37°C (5% (w/v) CO_2) for 24 hours. Optionally, transfection mixture was removed after 4 hours incubation and replaced with complete DMEM. On the next day, the DMEM was removed and cells were rinsed twice with 1 ml of 30°C PBS before fixing with 300 µl of 4% (v/v)

paraformaldehyde. The 12-well plates were then placed on a rocker for 15 minutes to gently mix. Coverslips were then rinsed twice with PBS and dH$_2$O and the edges of coverslips were blotted. 10 µl of Vectashield® HardSet™ mounting reagent containing DAPI for nuclear staining was put onto slides and the coverslips were mounted immediately with the cells facing downwards to the Vectashield®. They were left to air-dry for 1 hour before storing in the dark at 4°C. Cells were visualised using an Olympus Model IX81 microscope system. All images were taken with a Hamamatsu C4742-80 camera.

2.7 Bioinformatic method

Four public databases namely the Human Protein Reference Database (HPRD) (Keshava Prasad *et al.*, 2009), Molecular Interaction database (MINT) (Chatr-aryamontri *et al.*, 2007), IntAct (Aranda *et al.*, 2010) and BioGRID (Breitkreutz *et al.*, 2008) were mined by Russell Hyde (University of Liverpool) to provide overlapping yet complementary datasets to detail as many literature-reported interactions as possible. The data were integrated to produce human 'interactome' network containing both binary protein-protein interactions observed and more ambiguous protein interactions derived from co-complexes detection techniques. To analyse or predict an interaction data, computational filters written by Jonathan Woodsmith (University of Liverpool) were utilised. Visualisation of protein-protein interaction data were generated using the open source bioinformatics software platform Cytoscape. All cytoscape figures were generated using Cytoscape version 2.8.2.

Chapter Three:
YEAST TWO-HYBRID SCREENING

3.1   Introduction

To provide a better understanding of the order and specificity within the human ubiquitination process, it is crucial to perform a detailed analysis of protein interaction preferences within this system. Previous large-scale Y2H screens performed in our lab generated a high-density map of human E2:E3 RING interactions and identified a surprising interaction involving a specific subset of E2 proteins (UBE2Ds and UBE2Es) and a deubiquitinating enzyme (DUB) called OTUB1 (Figure 3.1A) (Markson *et al.*, 2009). This data raised an interesting question as to why OTUB1 should associate with specific E2 proteins. Classically, DUB proteins are well-known for cleaving or trimming ubiquitin chains from the substrates and although DUB proteins had been found to interact with some E3-RING proteins, no association with E2-ubiquitin conjugating enzymes had been observed.

During the course of this project, a large-scale systematic mass-spectrometry analysis of human DUB proteins was performed in which interactions between OTUB1 and UBE2D2 and UBE2N proteins were observed (Figure 3.1B) (Sowa *et al.*, 2009). Functional investigation of these non-canonical DUB:E2 complexes revealed a role in DNA damage response and p53 signalling (Nakada *et al.*, 2010; Sun *et al.*, 2012). Although these reports have addressed two functional roles of E2:OTUB1 complexes, little attention was given to define the molecular mechanism of E2:OTUB1 interactions.

Therefore three key aims of this project were:

1) To establish if OTUB1 was unique among DUB enzymes in being able to bind E2 proteins;

2) To verify that OTUB1 binds E2 proteins via direct binary interactions and finally;

3) To provide insight into the molecular mechanism or specificity of interactions within OTUB1:E2 complexes.

(A)   (B)
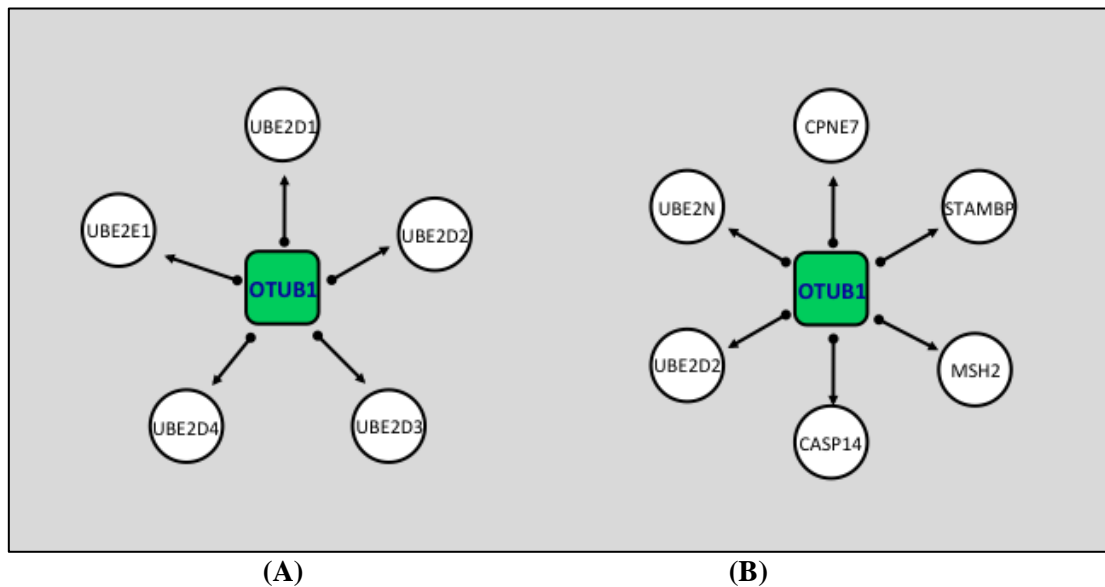
Figure 3.1   **Known E2:DUB interaction:** **(A)** Library screen shows OTUB1 interacted with E2 from D and E subfamilies (Markson *et al.*, 2009). **(B)** Mass spectrometry analysis defined that UBE2D2 and UBE2N are among OTUB1 interactors (Sowa *et al.*, 2009).

3.2   Yeast strains and vectors

The Y2H method has been extensively used to study protein-protein interactions as it can be performed in a high-throughput format and is known to be able to detect even transient or weak binary protein interactions. In these experiments, PJ69-4A (*MATa trp1-901 leu2-3, 112 ura3-52, his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*) was used as the bait strain and PJ69-4α, a suitable mating partner with identical genotype to PJ69-4A was used as the prey strain (James *et al.*, 1996). All Y2H assays performed in this study employed a mating strategy to generate diploid yeast containing both bait and prey constructs as this procedure has been shown to be more efficient than co-transfection protocols which tend to be affected by variability in transfection efficiency (Garcia-Cuellar *et al.*, 2009).

The PJ69-4A contains three promoters: GAL1, GAL2 and GAL7 that are all induced by the GAL4 transcription activator (Bram *et al.*, 1986). The three reporter genes controlled by these promoters are *HIS3* (GAL1), *ADE2* (GAL2) and *lacZ* (GAL7). The *HIS3* gene in this strain was initially reported to be "leaky" as some growth was observed on media lacking histidine (James *et al.*, 1996). Therefore to increase selection stringency, the *HIS3* product is inactivated by the competitive inhibitor 3-aminotriazole (3-AT). The concentration of this competitive inhibitor can be varied to increase screen stringency as required, providing a form of 'tunable rheostat' that can be used to eliminate non-specific background interactions (James *et al.*, 1996) or to order relative strength of interactions between different partners (Lehner and Sanderson, 2004). Previous experiments in our laboratory (Lehner *et al.*, 2004; Lehner and Sanderson, 2004) has shown that 2.5 mM 3-AT can effectively eliminate false positive interactions, so this concentration was used for all further experiments described in this study. In contrast, the *ADE2* reporter gene is a more stringent reporter than the *HIS3* reporter (James *et al.*, 1996) showing very low background or non-specific interactions. The *lacZ* reporter gene encodes β-galactosidase, an enzyme that cleaves the disaccharide lactose into glucose and galactose (Lederberg, 1948). Yeast colonies can be assayed for β-

galactosidase expression by using a chemical called X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) that is cleaved into galactose and a blue insoluble product. In our experience the *lacZ* reporter in the PJ69-4 strains is less reliable than the two growth reporters. As such, in this study we independently monitored growth on *-HIS* (+3AT) or *-ADE* conditions. Although growth on both independent reporters is considered ideal, previous studies from the Sanderson lab show that reproducible positive interactions observed on *-HIS* selection alone were sufficient to allow the reliable prediction of interactions which are confirmed by subsequent *in vitro* assays and mutagenesis studies (Markson *et al.*, 2009).

In this study, the pGBAD-B and pACTBD-B vectors were used to construct sets of bait and prey E2 clones respectively. For DUB proteins, baits were inserted in pGBDU-GW and preys were expressed in the pACTBE-B vector. Meanwhile for generating bait and prey $^{\Delta N}$OTUB1 clones, pGBAE-B and pACTBE-B were used. The pGBAD-B/pACTBD-B vectors were designed for cloning DNA sequences that incorporate an in-frame termination codon at the end of an ORF while pGBAE-B/pACTBE-B vectors have a one base pair insertion at the recombination site separated by the *att*B flanking regions to allow frame-shift and stop codon insertion after homologous recombination (Semple *et al.*, 2005). Unlike conventional Y2H vectors which encode either BD or AD domains, the pGBAD-B/pACTBD-B and pACTBE-B vectors encode both BD and AD domains separated by an in-frame linker containing the *att*B1 and *att*B2 Gateway® recombination sequences flanking a *Bam*HI restriction site which are vital in selection of positive transformants following *in vivo* homologous recombination (Figure 3.2). In contrast, pGBDU-GW was generated by insertion of a Gateway® Rf-conversion cassette into the pGBDU-C1 vector (James *et al.*, 1996). This vector contains only the BD domain hence serves as a bait plasmid and *att*R1 and *att*R2 recombination sites downstream the BD domain that make it possible to insert any ORF with *att*L1 and *att*L2 sites by LR recombination cloning. The principle behind the Gateway® cloning system is discussed in detail in Section 3.3.2.1.
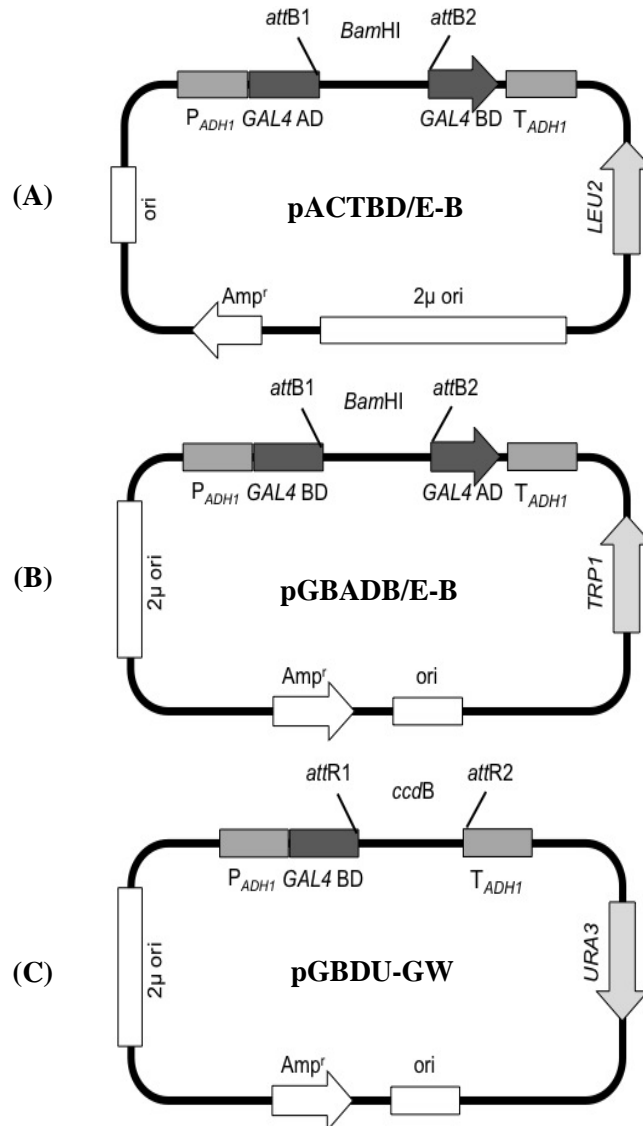
Figure 3.2   Y2H vectors used in this study: (A) pACTBD/E-B prey vectors based on pACT2 carrying *LEU2* gene to allow selection on media lacking leucine. **(B)** pGBAD/E-B bait vectors based on pGBD-C1 depends on *TRP1* as selection marker for colony growth on SD-W plate. **(C)** pGBDU-GW vector based on pGBT9 conferred with *URA3* gene to allow selection on media lacking uracil.

3.3   Construction of Y2H bait and prey clones

3.3.1   Existing constructs

3.3.1.1 E2 bait and prey

Y2H binary interactions involving human E2:E3-RING, E3-RING:E3-RING and E3-RING:DUB proteins have been systematically conducted in large-scale Y2H studies in our lab. Therefore, enormous resources of Y2H constructs were already available at the start of this study. Previously, a collection of 39 human E2 clones (Table 3.1) were obtained from Mammalian Gene Collection (Collins *et al.*, 2002) or from Matchmaker cDNA libraries (Clontech). These ORFs were PCR-amplified with primers designed in accordance with RefSeq annotations for the full-length protein-coding region, including the endogenous stop codon. Each primer contained forward or reverse flanking sequences to facilitate transfer of the PCR products into the Y2H pGBAD-B bait and pACTBD-B prey vectors by *in vivo* gap repair reactions. Positively transfected Mat-a bait and Mat-α prey yeast cells were selected for growth on media lacking tryptophan (SD-W) and leucine (SD-L) respectively, as the pGBAD-B and pACTBD-B vectors contain the *TRP1* and *LEU2* gene that allows for growth on this medium. From 39 unique human genes encoding E2 proteins, a complete set consisted of 44 bait clones (including duplicate UBE2L3, UBE2W, UEVLD and UBE2Z clones and a splice variant of UBE2I) were successfully generated, although only 35 prey clones (excluding UBE2J2, UBE2Q2, UBE2DNL and TSG101) were available for use in this study (Table 3.1).

Table 3.1 **The ubiquitin conjugating enzyme (E2s):** The names and ID of the 39 human E2 and ubiquitin E2 variant (UEV) proteins involved in this experiment. Some of the construct are only available in bait as can be seen by legend: ◆=bait;　■=prey.

	GeneID	Gene Name	Alternate name	Bait (◆) Prey (■)
1	7319	UBE2A	UBC2; HHR6A; RAD6A	◆　■
2	7320	UBE2B	HR6B; UBC2; HHR6B; RAD6B; E2-17kDa	◆　■
3	11065	UBE2C	UBCH10; dJ447F3.2	◆　■
4	7321	UBE2D1	SFT; UBCH5; UBC4/5; UBCH5A; E2(17)KB1	◆　■
5	7322	UBE2D2	UBC4; PUBC1; UBC4/5; UBCH5B; E2(17)KB2	◆　■
6	7323	UBE2D3	UBC4/5; UBCH5C; MGC5416; MGC43926; E2(17)KB3	◆　■
7	51619	UBE2D4	HBUCE1; FLJ32004	◆　■
8	7324	UBE2E1	UBCH6	◆　■
9	7325	UBE2E2	UBCH8; FLJ25157	◆　■
10	10477	UBE2E3	UBCH9; UbcM2	◆　■
11	140739	UBE2F	NCE2; MGC18120	◆　■
12	7326	UBE2G1	UBC7; E217K; UBE2G	◆　■
13	7327	UBE2G2	UBC7	◆　■
14	7328	UBE2H	UBC8; UBCH; UBCH2; E2-20K	◆　■
15	7329	UBE2I	P18; UBC9; C358B7.1	◆　■
16	51465	UBE2J1	UBC6; Ubc6p; CGI-76; NCUBE1; HSPC153; HSPC205; NCUBE-1; HSU93243; MGC12555	◆　■
17	118424	UBE2J2	NCUBE2; NCUBE-2; PRO2121BAIT	◆
18	3093	UBE2K	LIG; HIP2; HYPG; UBC1; E2-25K; DKFZp564C1216; DKFZp686J24237	◆　■
19	7332	UBE2L3	E2-F1; L-UBC; UBCH7; UbcM4	◆　■
20	9246	UBE2L6	RIG-B; UBCH8; MGC40331	◆　■
21	9040	UBE2M	UBC12; hUbc12; UBC-RS2	◆　■
22	7334	UBE2N	UBC13; MGC8489; UbcH-ben; MGC131857	◆　■
23	63893	UBE2O	E2-230K; FLJ12878; KIAA1734	◆　■
24	55585	UBE2Q1	GTAP; UBE2Q; NICE-5; PRO3094	◆　■
25	92912	UBE2Q2	DKFZp762C143	◆
26	997	UBE2R1	UBC3; UBCH3; CDC34; E2-CDC34	◆　■
27	54926	UBE2R2	UBC3B; CDC34B; FLJ20419; MGC10481	◆　■
28	27338	UBE2S	EPF5; E2EPF; E2-EPF	◆　■
29	29089	UBE2T	PIG50; HSPC150	◆　■
30	148581	UBE2U	MGC35130; RP4-636O23.1	◆　■
31	7335	UBE2V1	CIR1; UEV1; CROC1; UBE2V; UEV-1; UEV1A; CROC-1	◆　■
32	7336	UBE2V2	MMS2; UEV2; EDPF1; UEV-2; DDVIT1; EDAF-1; EDPF-1; DDVit-1	◆　■
33	55284	UBE2W	hUBC-16; FLJ11011	◆　■
34	65264	UBE2Z	USE1; HOYS7; FLJ13855	◆　■
35	100131816	UBE2DNL	MGC42638	◆
36	55293	UEVLD	ATTP; UEV3; FLJ11068	◆　■
37	64400	AKTIP	FT1; FTS	◆　■
38	57448	BIRC6	BRUCE; APOLLON; FLJ13726; FLJ13786; KIAA1289	◆　■
39	7251	TSG101	TSG10; VPS23	◆

3.3.1.2 DUB bait

The DUB bait set was generated in the Sanderson lab by Sebastian Hayes (Hayes, 2009). The ORFs were collected from various sources mainly via Human ORFeome v1.1 (Rual *et al.*, 2004), IMAGE clones, and from the lab of Mike Clague and Sylvie Urbe (University of Liverpool) and cDNA libraries (Clontech). Only the ORFeome cDNA clones were already in Gateway® format, while other clones were PCR-amplified using Gateway® primers and put through BP reactions to create pDONR223 constructs with *att*L sites. All the ORFs were cloned by LR reaction into pGBDU-GW bait vectors and then transfected into Mat-a bait yeast cells. Positive clones were selected by growing the cells on media lacking uracil (SD-U). From the original list of 92 putative human DUBs, seven were excluded either because they were thought to be pseudogenes, or because they contained massively large ORFs with incomplete sequence information. Some clones were also eliminated at different stages of clone generation, resulting in a final collection of 65 human DUB proteins which were successfully cloned into the pGBDU-GW vector. Five were disqualified for Y2H screening for being autoactivated in diploid. The final collection of DUB baits tested in this study included 69 ORFs representing a total of 60 unique DUBs and nine duplicates including additional clones for the USP5, BAP1, USP15, USP20, USP25, USP30, STAMBP and STAMBPL1 proteins. These are listed in Table 3.2.

3.3.2   Generation of a new Y2H DUB prey clone set

3.3.2.1 Application of the Gateway® cloning strategy to create DUBs in pDONR223

The Gateway® system (Invitrogen) was used to generate clone libraries in our lab because of its effective one-step cloning method which facilitates easy shuttling of ORFs from entry vectors into a wide range of expression plasmids (illustrated in Figure 3.3) (Landy, 1989; Hartley *et al.*, 2000). DUB ORFs were PCR-amplified with primers specifically designed to contain the appropriate Gateway® recombination sequences. Forward primers were designed to contain 18-25 bases of ORF-specific sequence with an additional 5' sequence containing

the *att*B1 recombination sequence (ACA AGT TTG TAC AAA AAA GCA GGC T). Reverse primers were similarly designed to contain 18-25 bases of ORF-specific sequence with a 3' extension containing the *att*B2 sequence (AC CAC TTT GTA CAA GAA AGC TGG GT). All primers were designed to contain ORF specific start and stop codons. Once a PCR product has been generated, it can be incubated *in vitro* with the pDONR223 donor vector containing the *att*P1 and *att*P2 sites flanking the *ccd*B gene and BP clonase mix containing the integrase and integration host factor (IHF) proteins. The BP recombination reaction results in the *ccd*B gene getting flipped out of the pDONR223 vector and being replaced by the DUB ORF, which is now flanked by the newly recombined *att*L1 and *att*L2 sequences. It is important to note that the *att*B1 site only recombines with the *att*P1 site (similarly for *att*B2 and *att*P2*)* and therefore the correct 5' to 3' orientation is maintained. DH5α strain *E. coli* were transformed with the BP reaction mixture and only pDONR223 vectors that had successfully undergone recombination were able to grow while pDONR223 with *ccd*B gene will not grow. The *ccd*B protein interferes with the *E. coli* DNA gyrase (Bernard and Couturier, 1992), thereby inhibiting growth of most *E. coli* strains carrying this gene. Positive colonies were selected on spectinomycin selective agar as pDONR223 confers resistance to this antibiotic (Rual *et al.*, 2004).

From these pDONR223 entry clones, ORFs can be shuttled to other expression vectors known as destination vectors by a similar recombination reaction. Entry clones (containing *att*L1-ORF-*att*L2) were incubated with a destination vector (*att*R1-*ccd*B-*att*R2) with a different antibiotic selection to the entry vector, and an LR clonase containing Int, IHF and bacteriophage excisionase. *E. coli* were transformed with the reaction mixture and transformants were selected for growth on appropriate antibiotic media. Subsequently, the product of LR reaction (termed 'expression clones') can be used for expression studies in a range of biological systems.
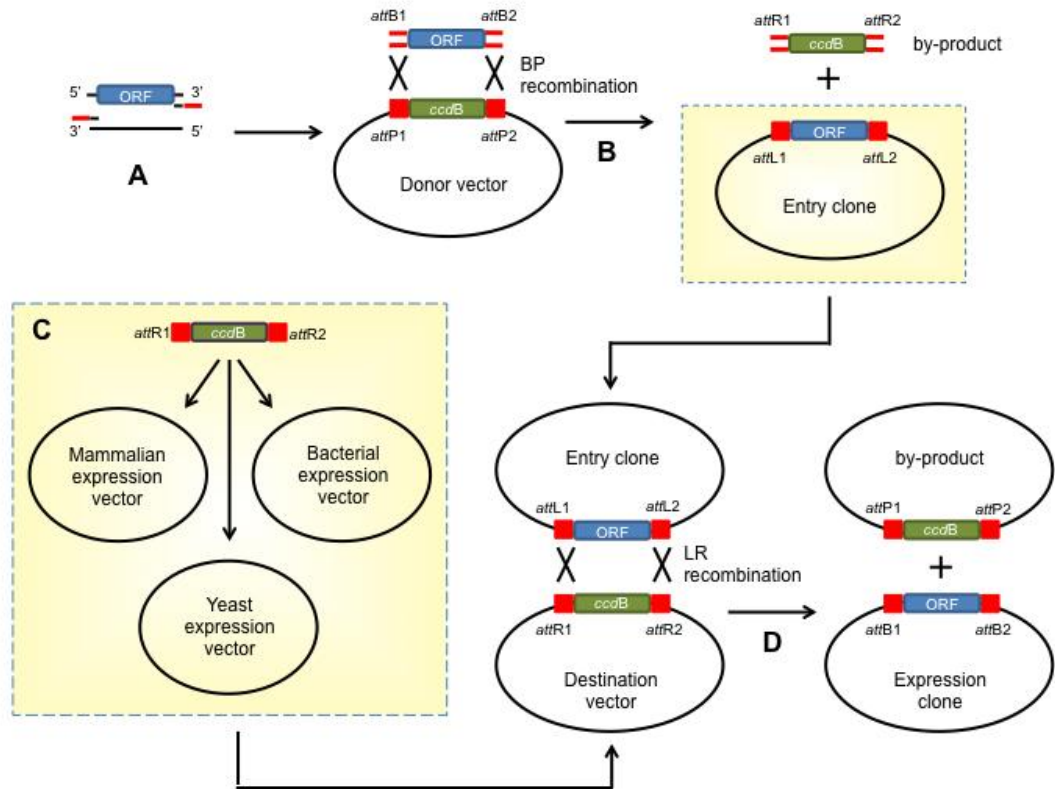
Figure 3.3 **The Gateway® system: (A)** ORFs are initially amplified with Gateway® primers with 5' and 3' overhangs. **(B)** The amplified ORF now flanked with *att*B1 and *att*B2 sites can be recombined with the *att*P1 and *att*P2 sites on a donor vector (pDONR223) to produce an entry clone flanked by *att*L1 and *att*L2 sites in what is called a BP reaction. **(C)** Generation of Gateway®-compatible destination vector by insertion of Gateway® conversion cassette (*att*R1-*ccd*B-*att*R2) into any existing mammalian, yeast or bacterial expression vector. **(D)** Entry clone in LR recombination with destination vectors containing the *att*R1 and *att*R2 sites.

3.3.2.2 PCR introduction of yeast and Gateway® sites to the DUB ORFs

In order to transfer DUB ORFs into the pACTBE-B prey vector, primers that could amplify ORF out of pDONR223 and allow homologous recombination with a linearised yeast cloning vector *in vivo* were designed. Here, blue indicates bases present in the pACTBE-B sequence only, red in pDONR223 sequence only and green present in both sequences;

pDONR GR F2:	5' GAATTCACAAGTTTGTACAAAAAAGCAGGCATG 3'
pDONR GR R1:	5' GTCGACCACTTTGTACAAGAAAGCTGGG 3'

Using this strategy, a total of 47 PCR products were successfully obtained. All of these products were cloned into the pACTBE-B vector simultaneously by *in vivo* gap repair cloning in Mat-α cells.

3.3.2.3 In vivo homologous recombination (gap repair)

DUB prey clones were generated in Mat-α yeast as described above. Transformation/gap repair reactions were then plated on SD-L, low adenine media to select positive transformants. Vectors containing an insert with an in-frame stop codon grow very weakly and accumulate red pigment containing low amounts of adenine. Red colonies were selected and yeast colony PCR was performed to confirm the correct ORF size. Even though all 47 gave a number of red colonies, only 27 ORFs produced a single insert band of the predicted size, the rest either produced multiple bands, inaccurate band size or no detectable YC-PCR product.

3.3.2.4 Autoactivation tests

Each Y2H clone generated in this study was tested for autoactivation to avoid false positive results when performing Y2H mating. Autoactivators are haploid Mat-α clones containing AD-protein fusion that can activate the transcription of reporter genes *HIS3*, *ADE2* and/or *lacZ* without the presence of an interacting bait protein (and *vice versa*). In principle, the reporter genes should only be activated upon mating the two yeast strains and forming

diploid cells. To assess whether any of the Y2H prey clones could induce reporter gene transcription within haploid Mat-α yeast, individual colonies were grown on SD-L (permissive), SD-LH(3-AT) and SD-LA (both non-permissive) agar plates. Haploid yeast which grow on SD-LH(3-AT) and/or SD-LA plates must be autoactivating the *HIS3* and/or *ADE2* reporter genes independently, therefore these were excluded from further studies.

Autoactivation of Y2H-inducible reporter genes is a common artefact of the Y2H system. There are three classes of autoactivators:

1) Genuine transcription factors that contain a *bona fide* AD and consequently will likely score as autoactivators when fused to DB,

2) Proteins that are not transcription factors in their natural context but can behave as autoactivators because they contain a cryptic AD (cognate autoactivators), and

3) Non-transcription factor proteins that contain one or more cryptic ADs that are only functional as truncated fragments and not when expressed in the context of full-length proteins (*de novo* autoactivators).

In this experiment, four independent clones were tested for each ORF in order to select a representative clone that does not autoactivate either *HIS3* or *ADE2* reporter genes. The final collection of DUB preys tested in this study included 30 ORFs representing a total of 27 unique DUBs, including duplicate clones of USP30, USP39 and USP46 (Table 3.2).

Table 3.2   **Y2H deubiquitinating enzymes (DUBs) clone collection:** The names and ID of the 61 human deubiquitinating proteins involved in this experiment. Some of the construct are only available in bait as indicated: ◆=bait;   ■=prey.

	GeneID	Gene Name	Alternate name	Bait (◆) Prey (■)
1	7345	UCHL1	PARK5; PGP95; PGP9.5; Uch-L1; PGP 9.5	◆ ■
2	8314	BAP1	UCHL2; hucep-6; FLJ35406; FLJ37180; HUCEP-13; KIAA0272; DKFZp686N04275	◆ ■
3	7347	UCHL3	UCH-L3	◆ ■
4	51377	UCHL5	UCH37; CGI-70; INO80R; UCH-L5	◆ ■
5	7398	USP1	UBP	◆ ■
6	9099	USP2a	USP9; UBP41	◆ ■
7	7375	USP4	UNP; Unph; MGC149848; MGC149849	◆
8	8078	USP5	ISOT	◆ ■
9	9098	USP6	HRP1; TRE2; TRE17; Tre-2; USP6-short	◆
10	7874	USP7	TEF1; HAUSP	◆
11	9101	USP8	UBPY; HumORF8; FLJ34456; KIAA0055; MGC129718	◆
12	8237	USP11	UHX1	◆
13	219333	USP12	UBH1; USP12L1	◆ ■
14	8975	USP13	ISOT3; IsoT-3	◆
15	9097	USP14	TGT	◆ ■
16	9958	USP15	UNPH4; KIAA0529; MGC74854; MGC131982; MGC149838	◆
17	10600	USP16	MSTP039, UBP-M	◆ ■
18	10869	USP19	ZMYND9	◆
19	10868	USP20	LSFR3A, VDU2	◆ ■
20	27005	USP21	RP11-297K8.3, USP16, USP23	◆ ■
21	29761	USP25	USP21	◆ ■
22	83844	USP26	MGC120066; MGC120067; MGC120068	◆
23	57646	USP28	KIAA1515	◆
24	84749	USP30	FLJ40511, MGC10702	◆ ■
25	84669	USP32	USP10; NY-REN-60	◆
26	23032	USP33	VDU1; KIAA1097; MGC16868	◆ ■
27	57602	USP36	DUB1	◆
28	84640	USP38	FLJ35970; HP43.8KD; KIAA1891	◆ ■
29	10713	USP39	SAD1; CGI-21; HSPC332; SNRNP65; MGC75069	◆ ■
30	373856	USP41		◆
31	84101	USP44	FLJ14528; DKFZp434D0127	◆ ■
32	85015	USP45	MGC14793	◆
33	64854	USP46	FLJ11850; FLJ12552; FLJ14283; FLJ39393	◆ ■
34	84196	USP48	USP31; RAP1GA1; MGC14879; MGC132556; DKFZp762M1713	◆
35	25862	USP49	MGC20741	◆
36	9924	USP52	PAN2	◆ ■
37	159195	USP54	C10orf29; FLJ37318; bA137L10.3; bA137L10.4	◆
38	55611	OTUB1	OTB1; OTU1; HSPC263; MGC4584; FLJ20113; FLJ40710; MGC111158	◆ ■
39	78990	OTUB2	OTB2; OTU2; MGC3102; FLJ21916; C14orf137	◆ ■
40	55432	YOD1	DUBA8; OTUD2; PRO0907; DKFZp451J1719; RP11-164O23.1	◆
41	54726	OTUD4	HIN1; DUBA6; HSHIN1; KIAA1046; DKFZp434I0721	◆
42	55593	OTUD5	DUBA; MGC104871; DKFZp761A052	◆
43	139562	OTUD6A	DUBA2; HSHIN6; FLJ25831	◆
44	51633	OTUD6B	DUBA5; CGI-77	◆
45	56957	OTUD7B	ZA20D1; CEZANNE	◆
46	80124	VCPIP1	DUBA3; VCIP135; FLJ23132; FLJ60694;	◆

			KIAA1850; DKFZp686G038	
47	7128	TNFAIP3	A20; OTUD7C; TNFA1P2; MGC104522; MGC138687; MGC138688	◆ ■
48	79184	BRCC3	C6.1A; BRCC36; CXorf53; RP11-143H17.2	◆ ■
49	10987	COPS5	CSN5; JAB1; SGN5; MOV-34; MGC3149	◆ ■
50	10980	COPS6	CSN6; MOV34-34KD	◆ ■
51	8667	EIF3H	EIF3S3; eIF3-p40; MGC102958; eIF3-gamma	◆ ■
52	8665	EIF3F	EIF3S5; eIF3-p47	◆ ■
53	5713	PSMD7	P40; S12; Rpn8; MOV34	◆ ■
54	10213	PSMD14	PAD1; POH1; RPN11	◆
55	10617	STAMBP	AMSH; MGC126516; MGC126518	◆
56	57559	STAMBPL1	AMSH-FP; AMSH-LP; ALMalpha; FLJ31524; KIAA1373; bA399O19.2	◆
57	92552	ATXN3L	MJDL; FLJ59638; MGC168806; MGC168807	◆
58	9929	JOSD1	KIAA0063; dJ508I15.2	◆
59	126119	JOSD2	SBBI54; FLJ29018	◆
60		FLJ14891		◆

3.3.3   Pooling and deconvolution strategy used to screen all possible E2:DUB interactions

Pooling-deconvolution strategies are frequently used to dramatically decrease the effort required to perform large-scale Y2H screens. In essence, this strategy allows simultaneous screening of multiple prey constructs with each bait clone, thus enabling a greater number of possible interactions to be screened in fewer experiments to provide greater screen coverage.

In this study, prey pools were prepared by combining up to 8 different Mat-α ORF prey clones (Figure 3.4). YPAD plates were set up with specific Mat-a haploid bait cells harbouring the bait constructs in a 96-well format. The suspension of pooled preys was then pipetted on top of these spots and they were allowed to mate for 24 hours. Diploid yeast were transferred by velvet replication onto SD-WL (for E2 bait:DUB preys) and SD-UL (for DUB bait:E2 preys), which were grown for 48 hours before transferred to triple selection SD-WLH(3-AT), SD-WLA (or SD-ULH(3-AT) or SD-ULA) plates. All screens were repeated twice and only positive interactions that were observed in both screens were recorded as true positive Y2H interactions. Pools containing positive hits were then selected for deconvolution where individual prey clones were tested against the observed positive bait partner. Following seven days of incubation on triple selection plates, images were taken and interactions were scored as being weak (+), medium (++), or strong (+++) activators of reporter gene expression as indicated by the extent of yeast colony growth on selective plates (Figure 3.5).
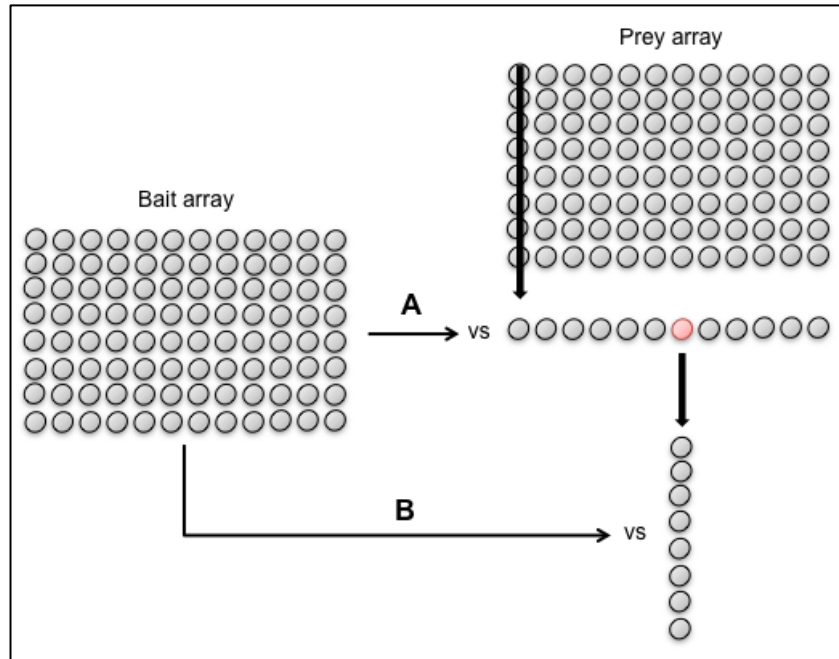
Figure 3.4   Pooling and deconvolution strategy: (A) Pooled mating assays were performed by picking yeast colonies from each vertical row and combining them together in one tube. Hence for the full 96-well plate, 12 pools were generated representing each column. Consequently, these pools were mated against a specific bait clone as shown in the figure. **(B)** Pools with positive interaction profiles were then deconvoluted and mated with the same bait clone which gave positive results in initial pooled screens.
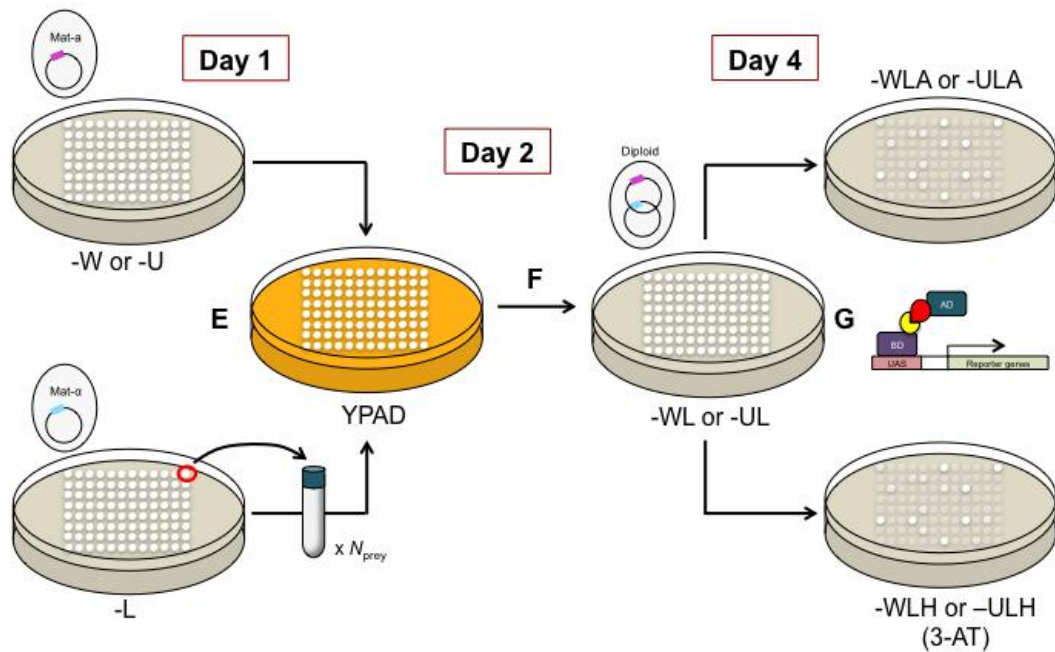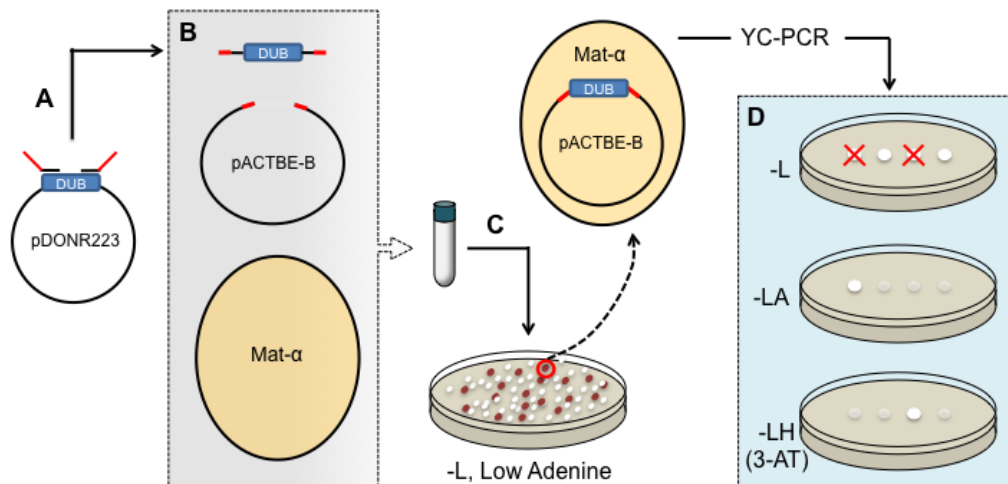
Figure 3.5 Schematic representation of Y2H protocol: **(A)** PCR-amplification of DUBs in pDONR223 vector incorporating flanking sequences with both Gateway® and pACTBE-B recombination sequences. **(B)** *In vivo* homologous recombination in Mat-α cells. **(C)** Positive transformants are identified as pink colonies on SD-L, low adenine media plate. **(D)** Autoactivation tests are performed following size-verification by YC-PCR. **(E)** Baits and preys are spotted in a 96-array format on YPAD rich medium and incubated for 24h. **(F)** Diploid yeasts are transferred by velvet-replication from YPAD plates onto SD-UL or SD-WL plates after 48h incubation. **(G)** Diploid yeast are replicated onto triple dropout selection plates (-ULH(3-AT) and -ULA or -WLH(3-AT) and –WLA) to positively select interacting proteins.

3.4 Binary Y2H screen result

3.4.1 Reconfirmation of previously described interaction

In this experiment, 30 DUB preys pooled into 12 separate groups were screened against 44 E2 baits. Using this approach only 528 matings were needed to test 1 320 potential binary interactions. From these screens, 24 positive interactions were detected for E2 baits, 14 of which were observed in DUB pool 6 containing USP2a, OTUB1 and USP46n (Figure 3.6). Apart from these interactions, some weak interactions were seen between UBE2Q1:DUB pool 4, AKTIP:DUB pool 9, UBE2F:DUB pool 12 and UBE2N:DUB pool 12. DUB pool 10 and 12 also showed weak interactions with UBE2U, UBE2DNL and TSG101. Meanwhile, positive interactions involving UBE2G2 were ignored as they were suspected to be false-positive due to low level autoactivation with this clone. Also for UBE2H and UBE2K, even though they have no record of autoactivation in other experiments they were found to interact with almost all preys in this assay. As a result, UBE2G2, UBE2H and UBE2K were excluded from deconvolution assay in order to avoid any false positives data.

Following deconvolution assays, 14 E2:DUB interactions were identified, most of which involved OTUB1 which interacted with UBE2N and all members of the E2 D and E subfamilies (with the exception of UBE2E2). Besides, other positive interactions were observed between UBE2U, TSG101 and UBE2DNL, which share common DUB interactions; COPS6 and EIF3F. EIF3F also interacted with AKTIP (Figure 3.7). It should be noted that the construct of EIF3F generated for this experiment contained a non-coding C561T mutation, which may not be affecting Y2H experiment but should be addressed crucially when performing any further analysis.

Y2H screens were also performed in the opposite bait/prey direction in which 38 E2 preys were pooled into 12 groups and mated against 72 individual DUB baits (Figure 3.8). In this study 2 736 binary interactions were tested in 864 pooled matings. Five DUB preys constructs (USP13, USP25, OTUD6A and both STAMBP duplicates) had previously been

shown to exhibit autoactivation in other experiments performed in our lab hence data from these clones were excluded in subsequent studies. In this orientation, less positive hits were observed with only 10 positive interactions being identified. As expected, OTUB1 shows a medium-strength positive interaction with E2 pools 2, 3, 5, 8 and 9, which contain members of the D and E subfamilies. Beside that, some weaker positive interactions were are also observed between USP2a, COPS5 and TNFAIP3 with E2 pool 1 and UBE2N:pool 7. OTUD5, COPS6 and ATXN3L also appear to share common interactions with E2 pool 10. Meanwhile, COPS6 interacted with most of the E2 prey clones indicating possible non-specific binding or low level autoactivation with this clone. Therefore, these hits were excluded from deconvolution assay. As very weak growth was observed (day 14) with EIF3F, TNFAIP3, ATXN3L, YOD1, OTUD5 and BRCC3, these interactions were recorded as possible positive interactions.

Y2H deconvolution mating of selected E2 pool preys against DUB baits resulted in 12 binary E2:DUB interactions that activated *HIS3*, *ADE2* or both reporters. OTUB1 was again the major interactor for the E2s, with obvious positive interactions with UBE2D2, UBE2E1, UBE2E2, UBE2E3 and UBE2W. Beside that, other partnerships were observed between TNFAIP3:UBE2I, TNFAIP3:UBE2U and USP2a:UBE2U (Figure 3.9).

All of these interactions with their score of either weak (+), medium (++), or strong (+++) activators of reporter expression are listed in Table 3.3.

(A)

Figure 3.6 panel A – E2 bait-pooled DUB prey interaction plates

(B)

	1	2	3	4	5	6	7	8	9	10	11	12
A	UBE2A	UBE2B	UBE2C	UBE2D1	UBE2D2	UBE2D3	UBE2D4	UBE2E1	UBE2E3	UBE2G1	UBE2G2	UBE2H
B	UBE2I	UBE2L3	UBE2L3	UBE2L6	UBE2M	UBE2N	UBE2V1	UBE2V2	UBE2S	UBE2R1	UBE2K	UBE2R2
C	UBE2J1	UBE2J2	UBE2W	UBE2E2	UEVLD	UEVLD	UBE2W	UBE2Z	AKTIP	UBE2O	UBE2Q1	UBE2T
D	UBE2I	UBE2F	BIRC6	UBE2U	UBE2DNL	UBE2Q2	UBE2Z	TSG101				
E												
F												
G												
H												

Figure 3.6   E2 bait-pooled DUB prey interactions: (A) E2 baits mated against the DUB pooled preys on matrix format. Yeast diploids were allowed to grow for two days and then scored for growth on SD-ULA (stringent) and SD-ULH(3-AT) (less stringent) after seven days. **(B)** Grid positions of the E2 baits assayed in Figure 3.6(A).

Legend: ○ = Strong;   ○ = Medium;   ○ = Weak;   ○ = Auto-activated

(A)

-WLH(3-AT)   -WLA

(B)
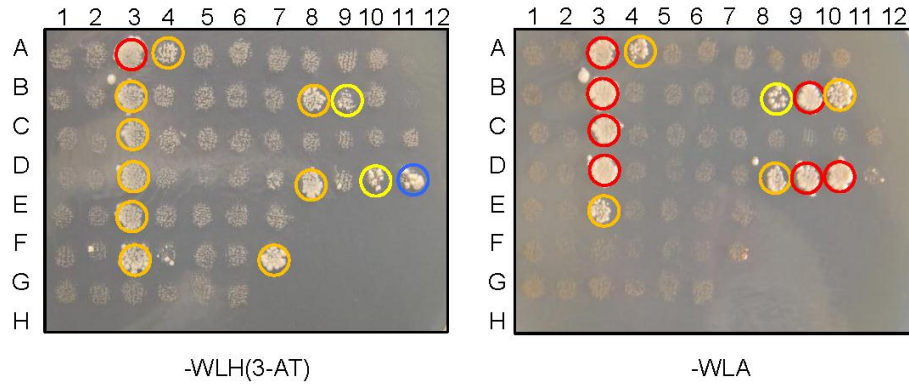
	1	2	3	4	5	6	7	8	9	10	11	12
A	UBE2D1 USP2a	UBE2N USP2a	UBE2D1 OTUB1	UBE2N OTUB1	UBE2D1 USP46n	UBE2N USP46n	UBE2Q1 BAP1	UBE2U USP16	TSG101 USP16	UBE2DNL USP16		
B	UBE2D2 USP2a	UBE2V1 USP2a	UBE2D2 OTUB1	UBE2V1 OTUB1	UBE2D2 USP46n	UBE2V1 USP46n	UBE2Q1 USP44	UBE2U COPS6	TSG101 COPS6	UBE2DNL COPS6		
C	UBE2D3 USP2a	UBE2E2 USP2a	UBE2D3 OTUB1	UBE2E2 OTUB1	UBE2D3 USP46n	UBE2E2 USP46n	UBE2Q1 USP30n	UBE2U USP30	TSG101 USP30	UBE2DNL USP30	UBE2F USP30	
D	UBE2D4 USP2a	UEVLD USP2a	UBE2D4 OTUB1	UEVLD OTUB1	UBE2D4 USP46n	UEVLD USP46n	UBE2Q1 EIF3H	UBE2U EIF3F	TSG101 EIF3F	UBE2DNL EIF3F	UBE2N EIF3F	
E	UBE2E1 USP2a	UBE2W USP2a	UBE2E1 OTUB1	UBE2W OTUB1	UBE2E1 USP46n	UBE2W USP46n	AKTIP USP14					
F	UBE2E3 USP2a	UBE2U USP2a	UBE2E3 OTUB1	UBE2U OTUB1	UBE2E3 USP46n	UBE2U USP46n	AKTIP COPS5					
G	UBE2M USP2a	UBE2Q1 USP2a	UBE2M OTUB1	UBE2Q1 OTUB1	UBE2M USP46n	UBE2Q1 USP46n						
H												

Plate 1

Figure 3.7   Pooled DUB preys deconvoluted in singe mating: (A) E2 baits mated against the DUB preys on matrix format. Yeast diploids were allowed to grow for two days at 30°C and then scored for growth on SD-WLA (stringent) and SD-WLH(3-AT) (less stringent) after seven days. **(B)** Plate 1 showing grid positions of the selected 18 E2 baits and 30 DUB preys mated in Figure 3.7(A); red font denotes the E2s whilst blue indicates the DUBs.

Legend: ⬤ = Strong;   ⬤ = Medium;   ⬤ = Weak;   ⬤ = Auto-activated

(A)

(B)

	1	2	3	4	5	6	7	8	9	10	11	12
A	UCHL1	UCHL3	UCHL5	BAP1	USP1	USP2a	USP3	USP4	USP5	USP6	USP7	USP8
B	USP11	USP12	USP13	USP14	USP15	USP16	USP19	USP20	USP20	USP21	USP25	USP26
C	USP28	USP30	USP30	USP32	USP33	USP36	USP38	USP39	USP41B	USP44	USP45	USP46
D	USP48B	USP49	USP52	USP54	OTUB1	OTUB2	OTUD7A	VCPIP1	OTUD4	YOD1	TNFAIP3	OTUD6B
E	OTUD6A	OTUD5	BRCC3	COPS5	COPS6	EIF3H	EIF3F	PSMD7	PSMD14	STAMBP	STAMBPL1	
F	ATXN3L	JOSD1	JOSD2	UBE2N	FLJ14891	USP5n	BAP1n	USP15n	STAMBP			
G												
H												

Figure 3.8   Pooled E2 prey-DUB bait interactions: (A) DUB baits mated against the E2 pooled preys. Yeast diploids were allowed to grow for two days and then scored for growth on SD-ULA and SD-ULH(3-AT) after seven days. **(B)** Grid positions of the DUB baits assayed in Figure 3.8(A).

Legend: 🔴= Strong;   🟠= Medium;   🟡 = Weak;   🔵= Auto-activated

(A)

(B)

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	COPS6 UBE2A	EIF3F UBE2A	TNFAIP3 UBE2A			USP2a UBE2A	OTUB1 UBE2B	OTUB1 UBE2C	OTUB1 UBE2D1	OTUB1 UBE2D2	BRCC3 UBE2A	
C	COPS6 UBE2I	EIF3F UBE2I	TNFAIP3 UBE2I			USP2a UBE2I	OTUB1 UBE2L3	OTUB1 UBE2L3	OTUB1 UBE2L6	OTUB1 UBE2M	BRCC3 UBE2I	
D	COPS6 UBE2J1	EIF3F UBE2J1	TNFAIP3 UBE2J1			USP2a UBE2J1	OTUB1 UBE2W	OTUB1 UBE2E2	OTUB1 UBE2W	OTUB1 UBE2Z	BRCC3 UBE2J1	
E	COPS6 UBE2G1	EIF3F UBE2G1	TNFAIP3 UBE2G1			ATXN3L UBE2G1	YOD1 UBE2G1	OTUD5 UBE2G1	UBE2N UBE2D4	OTUB1 UBE2E3	BRCC3 UBE2G1	
F	COPS6 UBE2R1	EIF3F UBE2R1	TNFAIP3 UBE2R1			ATXN3L UBE2R1	YOD1 UBE2R1	OTUD5 UBE2R1	UBE2N UBE2V1	OTUB1 UBE2S	BRCC3 UBE2R1	
G	COPS6 UBE2I	EIF3F UBE2I	TNFAIP3 UBE2I			ATXN3L UBE2I	YOD1 UBE2I	OTUD5 UBE2I	UBE2N UBE2O	OTUB1 UBE2T	BRCC3 UBE2I	
H	COPS6 UBE2U	EIF3F UBE2U	TNFAIP3 UBE2U			USP2A UBE2U	OTUB1 UBE2W	OTUB1 UBE2E1	OTUB1 UBE2V2	OTUB1 UBE2Q1	BRCC3 UBE2U	

Plate 2

Figure 3.9 **Pooled E2 preys deconvoluted in singe mating: (A)** DUB baits mated against the E2 preys on matrix format. Yeast diploids were allowed to grow for two days at 30°C and then scored for growth on SD-ULA (stringent) and SD-ULH(3-AT) (less stringent) after seven days. **(B)** Plate 2 showing grid positions of the selected 10 DUB baits and 38 E2 preys mated in Figure 3.9(A); blue font denotes the DUBs whilst red indicates the E2s.

Legend: ⭕ = Strong;   🟠 = Medium;   🟡 = Weak;   🔵 = Auto-activated

Table 3.3 **E2:DUB interaction summary:** List of all interactions tested on deconvolution assay. Interactions are scored as being weak (+), medium (++), or strong (+++) activators of reporter expression evaluated by the size and intensity of the yeast colony growth.

Plate / Well	Bait GeneID	Bait Gene Name	Prey Gene Name	Prey GeneID	*HIS3*	*ADE2*
Plate 1: E2 bait vs. DUB prey						
A1	7321	UBE2D1	USP2a	9099		
A2	7334	UBE2N	USP2a	9099		
A3	7321	UBE2D1	OTUB1	55611	+++	+++
A4	7334	UBE2N	OTUB1	55611	++	++
A5	7321	UBE2D1	USP46n	64854		
A6	7334	UBE2N	USP46n	64854		
A7	55585	UBE2Q1	BAP1	8314		
A8	148581	UBE2U	USP16	10600		
A9	7251	TSG101	USP16	10600		
A10	100131816	UBE2DNL	USP16	10600		
B1	7322	UDE2D2	USP2a	9099		
B2	7335	UBE2V1	USP2a	9099		
B3	7322	UDE2D2	OTUB1	55611	++	+++
B4	7335	UBE2V1	OTUB1	55611		
B5	7322	UDE2D2	USP46n	64854		
B6	7335	UBE2V1	USP46n	64854		
B7	55585	UBE2Q1	USP44	84101		
B8	148581	UBE2U	COPS6	10980	++	+
B9	7251	TSG101	COPS6	10980	+	+++
B10	100131816	UBE2DNL	COPS6	10980		++
C1	7323	UBE2D3	USP2a	9099		
C2	7325	UBE2E2	USP2a	9099		
C3	7323	UBE2D3	OTUB1	55611	++	+++
C4	7325	UBE2E2	OTUB1	55611		
C5	7323	UBE2D3	USP46n	64854		
C6	7325	UBE2E2	USP46n	64854		
C7	55585	UBE2Q1	USP30n	84749		
C8	148581	UBE2U	USP30	84749		
C9	7251	TSG101	USP30	84749		
C10	100131816	UBE2DNL	USP30	84749		
C11	140379	UBE2F	USP30	84749		
D1	51619	UBE2D4	USP2a	9099		
D2	55293	UEVLD	USP2a	9099		
D3	51619	UBE2D4	OTUB1	55611	++	+++
D4	55293	UEVLD	OTUB1	55611		
D5	51619	UBE2D4	USP46n	64854		
D6	55293	UEVLD	USP46n	64854		
D7	55585	UBE2Q1	EIF3F	8665		
D8	148581	UBE2U	EIF3F	8665		
D9	7251	TSG101	EIF3F	8665	++	++
D10	100131816	UBE2DNL	EIF3F	8665		+++
D11	7334	UBE2N	EIF3F	8665	+	+++
E1	7324	UBE2E1	USP2a	9099		
E2	55284	UBE2W	USP2a	9099		
E3	7324	UBE2E1	OTUB1	55611	++	++
E4	55284	UBE2W	OTUB1	55611		
E5	7324	UBE2E1	USP46n	64854		
E6	55284	UBE2W	USP46n	64854		
E7	64400	AKTIP	USP14	9097		
F1	10477	UBE2E3	USP2a	9099		
F2	148581	UBE2U	USP2a	9099		
F3	10477	UBE2E3	OTUB1	55611	++	

F4	148581	UBE2U	OTUB1	55611		
F5	10477	UBE2E3	USP46n	64854		
F6	148581	UBE2U	USP46n	64854		
F7	64400	AKTIP	COPS5	10987	++	
G1	9040	UBE2M	USP2a	9099		
G2	55585	UBE2Q1	USP2a	9099		
G3	9040	UBE2M	OTUB1	55611		
G4	55585	UBE2Q1	OTUB1	55611		
G5	9040	UBE2M	USP46n	64854		
G6	55585	UBE2Q1	USP46n	64854		
Plate 2: E2 prey vs. DUB bait						
B1	10980	COPS6	UBE2A	7319		
B2	8665	EIF3F	UBE2A	7319		
B3	7128	TNFAIP3	UBE2A	7319		
B6	9099	USP2a	UBE2A	7319		
B7	55611	OTUB1	UBE2B	7320		
B8	55611	OTUB1	UBE2C	11065		
B9	55611	OTUB1	UBE2D1	7321		
B10	55611	OTUB1	UBE2D2	7322	+++	
B11	79184	BRCC3	UBE2A	7319		
C1	10980	COPS6	UBE2I	7329		
C2	8665	EIF3F	UBE2I	7329		
C3	7128	TNFAIP3	UBE2I	7329		
C6	9099	USP2a	UBE2I	7329		
C7	55611	OTUB1	UBE2L3	7332		
C8	55611	OTUB1	UBE2L3	7332		
C9	55611	OTUB1	UBE2L6	9246		
C10	55611	OTUB1	UBE2M	9040		
C11	79184	BRCC3	UBE2I	7329		
D1	10980	COPS6	UBE2J1	51465		
D2	8665	EIF3F	UBE2J1	51465		
D3	7128	TNFAIP3	UBE2J1	51465		
D6	9099	USP2a	UBE2J1	51465		
D7	55611	OTUB1	UBE2W	55284		
D8	55611	OTUB1	UBE2E2	7325	+++	
D9	55611	OTUB1	UBE2W	55284		
D10	55611	OTUB1	UBE2Z	65264		
D11	79184	BRCC3	UBE2J1	51465		
E1	10980	COPS6	UBE2G1	7326		
E2	8665	EIF3F	UBE2G1	7326		
E5	7128	TNFAIP3	UBE2G1	7326		
E6	92552	ATXN3L	UBE2G1	7326		
E7	55432	YOD1	UBE2G1	7326		
E8	55593	OTUD5	UBE2G1	7326		
E9	7334	UBE2N	UBE2D4	51619		
E10	55611	OTUB1	UBE2E3	10477	+++	
E11	79184	BRCC3	UBE2G1	7326		
F1	10980	COPS6	UBE2R1	997		
F2	8665	EIF3F	UBE2R1	997		
F3	7128	TNFAIP3	UBE2R1	997		
F6	92552	ATXN3L	UBE2R1	997		
F7	55432	YOD1	UBE2R1	997		
F8	55593	OTUD5	UBE2R1	997		
F9	7334	UBE2N	UBE2V1	7335	+++	++
F10	55611	OTUB1	UBE2S	27338		
F11	79184	BRCC3	UBE2R1	997		
G1	10980	COPS6	UBE2I	7329		
G2	8665	EIF3F	UBE2I	7329		
G3	7128	TNFAIP3	UBE2I	7329	++	+
G6	92552	ATXN3L	UBE2I	7329		
G7	55432	YOD1	UBE2I	7329		
G8	55593	OTUD5	UBE2I	7329	+	++

G9	7334	UBE2N	UBE2O	63893		
G10	55611	OTUB1	UBE2T	29089		
G11	79184	BRCC3	UBE2I	7329		
H1	10980	COPS6	UBE2U	148581		+
H2	8665	EIF3F	UBE2U	148581		+
H3	7128	TNFAIP3	UBE2U	148581	++	++
H6	9099	USP2a	UBE2U	148581	++	++
H7	55611	OTUB1	UBE2W	55284	++	
H8	55611	OTUB1	UBE2E1	7324	+++	
H9	55611	OTUB1	UBE2V2	7336		
H10	55611	OTUB1	UBE2Q1	55585		
H11	79184	BRCC3	UBE2U	148581		

These results demonstrate a strong correlation with data from previous Y2H library screens and mass spectrometry studies (Markson *et al.*, 2009; Sowa *et al.*, 2009) regarding the interaction between OTUB1 and E2 proteins from the D and E subfamilies.

An interesting partnership is observed between COPS6 and TSG101, as both of these proteins are involved in interactions with p53 (Li *et al.*, 2001; Bech-Otschir *et al.*, 2001). TSG101 is an E2 belonging to UEV (ubiquitin E2 variant) domain members which shows significant sequence similarity to E2 enzymes. However, they are unable to catalyse ubiquitin transfer as they lack the active site cysteine that forms the transient thioester bond with the C-terminus of ubiquitin (Koonin and Abagyan, 1997; Ponting *et al.*, 1997). In relationship with p53, TSG101 participates with E3 ligase MDM2 in an autoregulatory loop that modulates the cellular levels of both proteins and of p53. Meanwhile, COPS6 (also known as CSN6) is one of the eight subunits that make up the COP9-signalosome, a highly conserved protein complex that functions as an important regulator in multiple signalling pathways (Wei *et al.*, 2008). Recently, another DUB component of the COP9 signalosome, COPS5 (or CSN5), has been shown to regulate p53 function (Zhang *et al.*, 2008), and p53 has also been shown to bind the native COP9 signalosome with high affinity through COPS5 (Bech-Otschir *et al.*, 2001). The observed interaction between COPS6 and TSG101 could therefore represent a mechanism by which p53 activity or stability could be regulated. In this experiment, COPS6 also interacted with UBE2DNL, a pseudogene with UBE2D N-terminal like region but very few interactions or literature reports have been recorded for UBE2DNL

hence the characteristic of this interaction could not be predicted.

In the human E2:DUB known interaction network (Figure 3.10), UBE2I was interacting with UCHL1 (Caballero *et al.*, 2002) but the interaction was not observed in our experiment. The function of UBE2I:UCHL1 interaction is unknown but UBE2I has been associated with protein trafficking by involvement in SUMO-modification which facilitates transport of modified proteins into the nucleus (Okuma *et al.*, 1999; Hoege *et al.*, 2002). Instead, in our experiment (Figure 3.9) UBE2I may has weak interaction with TNFAIP3 (also known as A20), functionally known as an inhibitor of cell death and chronic inflammation that downregulates NF-κB activation via the tumour necrosis factor receptor (TNFR)-associated pathway (Wertz *et al.*, 2004). TNFAIP3 is a very interesting protein because it is the only known DUB that also has E3 ligase activity mediated by one of its C-terminal zinc-finger domains that can promote the conjugation of Lys48-linked ubiquitin chains and proteasomal degradation (Wertz *et al.*, 2004). In order to disrupt interactions between E2:E3 enzymes in TNFR and the TLR4/IL-1R pathways, TNFAIP3 together with the regulatory molecule TAX1BP1 has been shown to interact with UBE2N and UBE2D3, the E2 involved in this pathway, thus triggering their ubiquitination and proteasome-dependent degradation (Shembade *et al.*, 2010). Interestingly, TNFAIP3 shares similarity with OTUB1 as both have OTU domains and both are known to be immunoregulatory DUBs (Sun, 2008). In this experiment, TNFAIP3, as well as USP2a, also interacted with UBE2U. However, no interaction data for UBE2U conjugation activity have been corresponded in the literature at present, which may be related to its restricted expression pattern in the urogenital tract (van Wijk *et al.*, 2009).
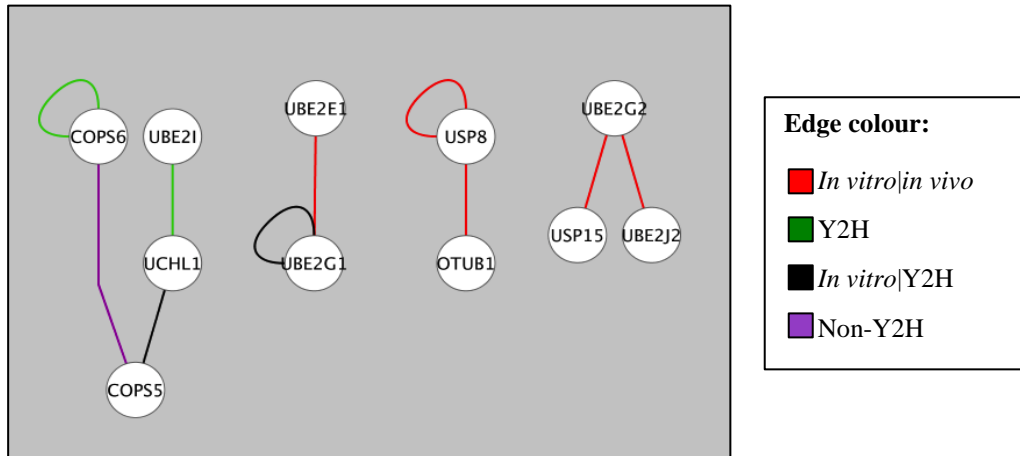
Figure 3.10   Known interactions between E2 ubiquitin conjugating enzymes and DUBs in humans: This data was extracted from our in-house database of known interactions mined from the MINT, HPRD, IntAct and BioGRID.

Interactions observed in this study do not corresponded to known interactions found in human databases (Figure 3.10). Nevertheless, it offers novel candidates to be investigated in future interaction studies. From these studies, it appears that E2:DUB pairs are very uncommon in cells because from a total of 4 056 binary interaction tested, only 23 positive hits were identified. However, this low score of positive results may also due to the restricted number of DUBs tested in our Y2H screens. As our final DUB bait and prey clone sets are not totally comprehensive clone sets, it is possible that several potential interactions may have been missed. From the 23 hits, 12 are involving OTUB1 and more importantly, only OTUB1 gave a confident and reproducible results in both bait and prey direction while other interactions were only detected in one of the bait prey orientation. As a result, OTUB1 and its partners were selected for further investigation using complementary biophysical and structural methods.

3.4.2   OTUB1 binds a subset of E2 conjugating enzymes

OTUB1 is a DUB belong to the OTU-family and its deubiquitination activity was firstly detected by the ability to cleave a tetraubiquitin substrate *in vitro* (Balakirev *et al.*, 2003). Its unexpected function was first observed when in transgenic mice transduced to express this gene, it promoted rather than inhibited the Lys48-linked self-ubiquitination and degradation of its endogenous interactor, GRAIL an E3-RING protein (Soares *et al.*, 2004). In more recent studies, the atypical characteristics of OTUB1 which enable it to prevent ubiquitin attachment rather than mediating the cleavage or removal of bound ubiquitin has been revealed and will be discussed more comprehensively in Chapter 6.

In these experiments, OTUB1 shows a clear binding preference for E2 proteins from the D and E subfamilies and UBE2N, Interestingly, the E2-conjugating D family is widely known as the most promiscuous (Brzovic and Klevit, 2006) E3-RING binding partner (Markson *et al.*, 2009), which may be consistent with their role as house-keeping E2s within the ubiquitin system. The D family share a high level of primary sequence similarity with E family members; however, the latter contain an approximate 60aa N-terminal extension not found in D family E2 proteins. These extensions are thought to be influential in determining specificity as E family members are less promiscuous than UBE2D proteins. Meanwhile, in forming ubiquitin chains, UBE2Ds, UBE2Es and UBE2N show a diversity of preferences. D and E families are more promiscuous and can catalyse the formation of multiple linkage chain types (Kirkpatrick *et al.*, 2006; Kim *et al.*, 2007; Jin *et al.*, 2008). In addition, promiscuous UBE2Ds preferentially promote the formation of Lys11-, Lys48- and Lys63-linked chains *in vitro* with three different E3s and show evidence of mixed and branched chains (Kim *et al.*, 2007). A dimeric complex composed of UBE2N and UBE2V1 was determined to only form a Lys63 linkage specificity (VanDemark *et al.*, 2001; Eddins *et al.*, 2006).

During this stage of analysis, the binding interface of E2:OTUB1 was not yet revealed, therefore our key aim was to investigate the molecular nature of interactions between OTUB1 and different E2 partners based on these three hypotheses:

1) OTUB1 binds E2 proteins at a different site from the well characterised E2-E3 binding surface;

2) OTUB1 and E3 compete with each other for binding to E2 partners;

3) OTUB1, E2 and E3-RING proteins exist in a single multiprotein complex.

As the structural model of the UBE2D2:CNOT4 complex (CNOT4 is an E3 RING ligase that interacts specifically with UBE2D2) had been solved by NMR and docking approaches (Dominguez *et al.*, 2004), we reasoned that this information and a similar approach would help to distinguish between the above options.

3.5   Analysis of a truncated form of OTUB1 lacking the first 39 N-terminal amino acids

Initial NMR analysis was performed to identify the binding interface of E2s and wild type (WT) OTUB1 complexes. Specific chemical shift patterns were observed in these studies suggesting specific points of contact between these two proteins (discussed further in Chapter 4). However, the full range of amino acids involved in these interactions could not be determined due to a very dense central region of the spectrum caused by an intrinsically disordered region within the OTUB1 protein. These regions (termed *'spaghetti'* because they usually exist as unorganised, flexible short linear peptide motifs) do not form a well-defined three-dimensional structure and sometimes even appear to be totally unfolded in their native state. As these types of regions can affect protein solubility and crystallisability which were an important part of our proposed plans, the probability disorder for OTUB1 was calculated using RONN (Yang *et al.*, 2005), which decides the likelihood of disorder based on alignments to a group of sequences of known folding state (Figure 3.11).
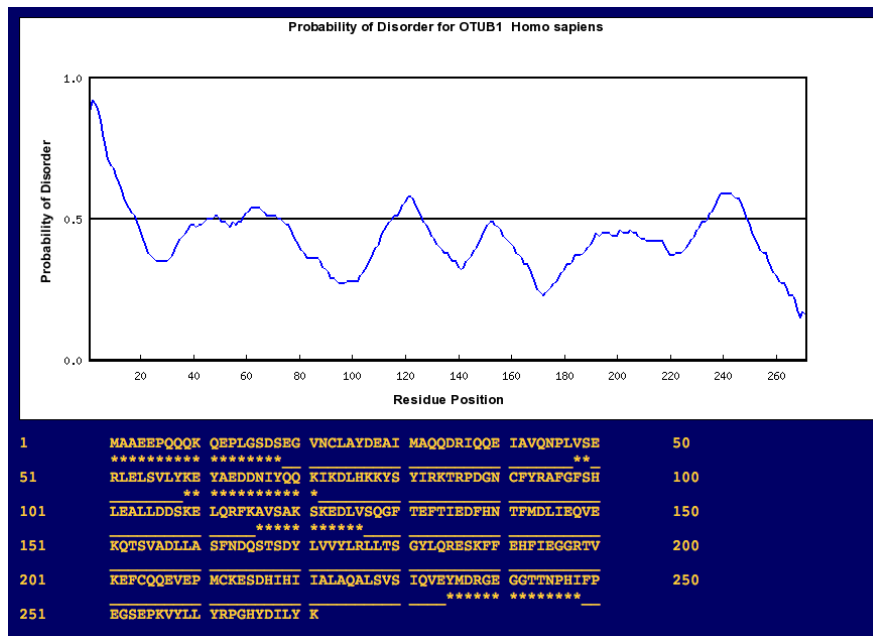
Figure 3.11   Probability of disorder for OTUB1:   The amino acids 1-18, 48-49, 59-71, 116-126, 235-248 are above probability cut-off of 0.5 and considered to be unfolded.

RONN predicted a strong area of disorder at the N-terminus of OTUB1, which does not include the catalytic site. In addition, in structural studies performed on the OTUB1 protein, the N-terminal region was removed to produce OTUB1 (residue 40-271 only) crystal which successfully diffracted to 1.7Å resolution (Edelmann *et al.*, 2009). Therefore, assuming that the N-terminal region would be problematic for structural studies we generated the truncated form of OTUB1 which lacked the N-terminal (1-39) amino acids.

It was possible that removal of the N-terminal region of OTUB1 could make OTUB1 more like OTUB2, as the main difference between the two proteins is the presence of the N-terminal extension in OTUB1 (Figure 3.12). Significantly, no interaction was observed between OTUB2 and any E2 proteins in our screens. In order establish if N-terminally truncated OTUB1 could still interact with defined E2 partners, further Y2H screens were performed using newly generated $^{\Delta N}$OTUB1 bait and prey constructs.
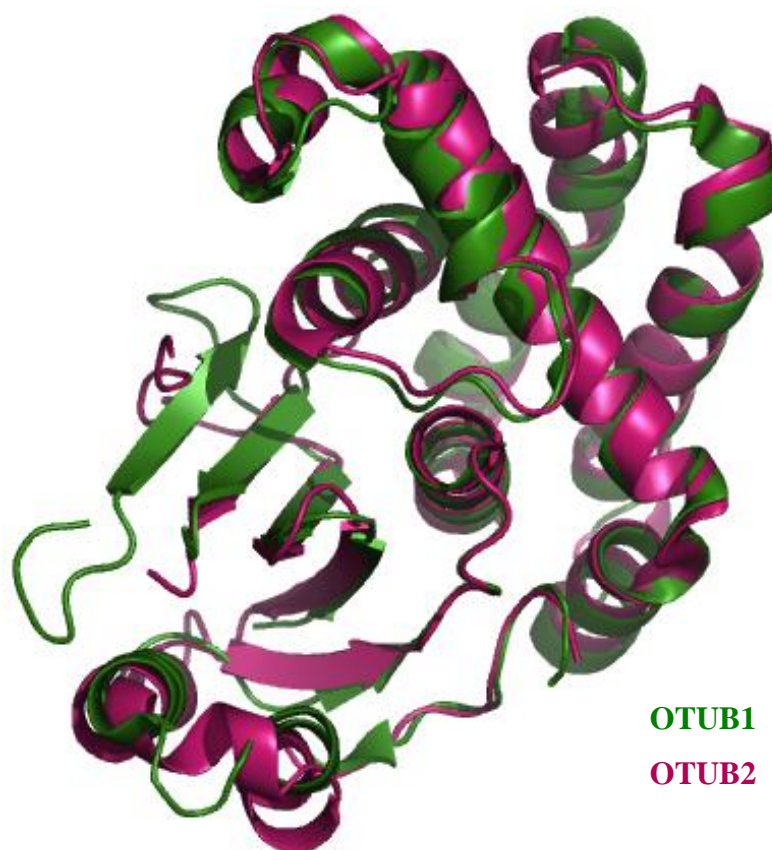
Figure 3.12   **Superposition of OTUB1 and OTUB2:** A very similar display with identical folds of OTUB1 (40-271) without N-terminal (PDB ID: 2ZFY) and OTUB2 (PDB ID: 1TFF). Picture generated using MacPyMOL v1.3.

3.5.1   Generating ^{ΔN}OTUB1

To generate ^{ΔN}OTUB1 Y2H bait and prey constructs, PCR was performed on full-length OTUB1 in pDONR223 using a gene-specific forward primer instead of vector-specific primer as used previously. The gene-specific primer was designed to contain a sequence of OTUB1 starting from base 118 (blue font) following the yeast/Gateway® sequence (red font):

5' GAATTCACAAGTTTGTACAAAAAAGCAGGCTGGATGGAGATTGCTGTGC 3'.

This primer would anneal in such a way as to amplify from base 118 onwards hence 117 bases representing the first 39 amino acids were removed. The same reverse primer pDONR223 GR R1 was used in conjunction with the newly designed forward primer. The product of this PCR is N-terminal truncated OTUB1 (aa 40-271) with overhang gap repair sequence at the 5' and 3'. Consequently, gap repair transformation was performed to using this PCR product to generate bait clones in pGBAE-B/Mat-a cells and equivalent prey clones in pACTBE-B/Mat-α. Following YC-PCR analysis of gap repair colonies, verified clones were tested for autoactivation.

3.5.2   ^{ΔN}OTUB1 slightly reduced interaction with E2 binding partners

Y2H results show that removal of N-terminal did not totally abolished interactions with E2 proteins (Figure 3.13). However, the interaction profile between ^{ΔN}OTUB1 and its binding partners were greatly reduced especially in the ^{ΔN}OTUB1 prey orientation, which might be due to a stearic effect. The stearic effects arise from the fact that each atom within a molecule occupies a certain amount of space. In some cases, if atoms are brought too close together, this may affect the molecule's preferred conformation and reactivity. A truncation is commonly disfavoured and notoriously known to cause stearic interaction with the fusion domain resulting in a reduced ability to bind to known interaction partners. The ^{ΔN}OTUB1 Y2H screen were analysed and the candidates for structural analysis were selected upon evaluating Y2H data.

(A)

(B)

	1	2	3	4	5	6	7	8	9	10	11	12
A	UBE2A	UBE2B	UBE2C	UBE 2D1	UBE2D2	UBE2D3	UBE 2D4	UBE 2E1	UBE 2E3	UBE 2G1	UBE 2G2	UBE2H
B	UBE2I	UBE2L3	UBE2L3	UBE2L6	UBE2M	UBE2N	UBE2V1	UBE2V2	UBE2S	UBE2R1	UBE2K	UBE2R2
C	UBE2J1	UBE2J2	UBE2W	UBE 2E2	UEVLD	UEVLD	UBE2W	UBE2Z	AKTIP	UBE2O	UBE2Q1	UBE2T
D	UBE2I	UBE2F	BIRC6	UBE2U	UBE2DNL	UBE2Q2	UBE2Z	TSG101				

(C)

(D)

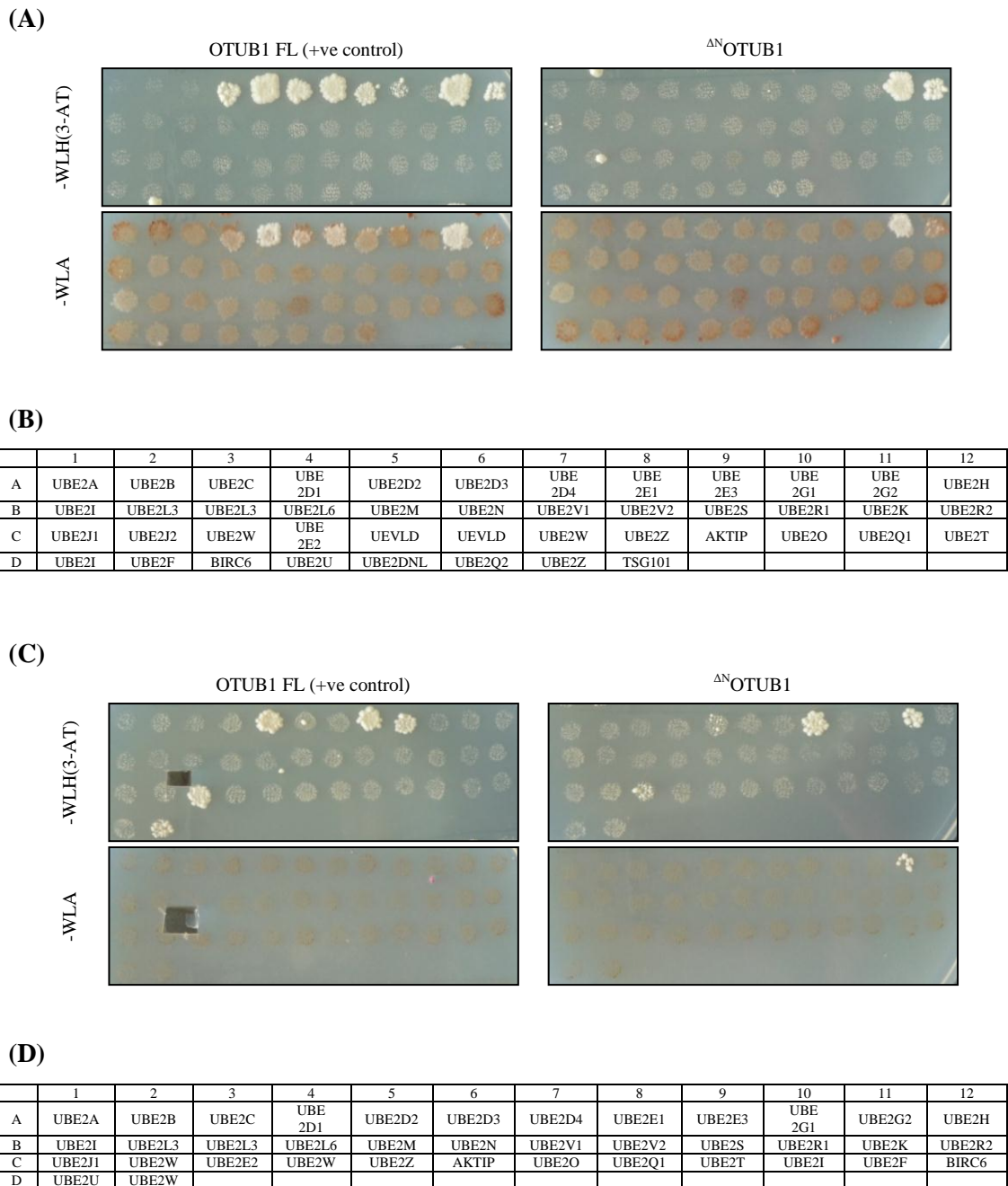|	1	2	3	4	5	6	7	8	9	10	11	12
A	UBE2A	UBE2B	UBE2C	UBE 2D1	UBE2D2	UBE2D3	UBE2D4	UBE2E1	UBE2E3	UBE 2G1	UBE2G2	UBE2H
B	UBE2I	UBE2L3	UBE2L3	UBE2L6	UBE2M	UBE2N	UBE2V1	UBE2V2	UBE2S	UBE2R1	UBE2K	UBE2R2
C	UBE2J1	UBE2W	UBE2E2	UBE2W	UBE2Z	AKTIP	UBE2O	UBE2Q1	UBE2T	UBE2I	UBE2F	BIRC6
D	UBE2U	UBE2W										

Figure 3.13 **ᐧᴺOTUB1:E2 interactions: (A)** ᐧᴺOTUB1 prey mated against E2 baits on SD-WLA and SD-WLH(3-AT) at day seven. **(B)** Grid positions of the E2 baits assayed in Figure 3.13(A). **(C)** ᐧᴺOTUB1 bait mated against E2 preys on SD-WLA and SD-WLH(3-AT) at day 7. **(D)** Grid positions of the E2 preys assayed in Figure 3.13(C).

3.6   Selection of candidates for biophysical analysis

OTUB1 in prey orientation interacts with the whole set of UBE2Ds and UBE2Es (except UBE2E2) and UBE2N in bait vectors, activating both *HIS3* and *ADE2* reporter genes. This observation is expected because previous experiments in our lab have shown that E2 expressed as bait fusions always show better interaction profiles compared to when they are expressed as prey clones. Surprisingly, in this orientation all E2 interactions were lost with $^{\Delta N}$OTUB1, possibly due to the stearic effects. However, when OTUB1 was expressed as a bait construct, interactions were only observed with UBE2D2, UBE2E1, UBE2E2 and UBE2E3 and only on the less stringent *HIS3* reporter. Interestingly, this pattern was maintained in screening with $^{\Delta N}$OTUB1 where only UBE2E3 interaction was abolished and the rest were preserved. The reason for OTUB1 bait selective preferences towards UBE2D2, UBE2E1, UBE2E2 and UBE2E3 are not well understood. Considering the Y2H data validation by biophysical analysis will be performed on $^{\Delta N}$OTUB1, therefore UBE2D2, UBE2E1 and UBE2E2 were selected for further experimentation as these were the only E2 proteins to maintained interaction positive interaction profiles with the truncated version of OTUB1. Table 3.4 on the next page shows the summary of OTUB1:E2 interaction.

Table 3.4   **OTUB1:E2 interaction summary:** Interaction between OTUB1 and E2s. Red denotes strong interaction, orange represents medium strength and yellow shows the observed interaction is weak, evaluated by the size and intensity of yeast colony growth.

			2D1	2D2	2D3	2D4	2E1	2E2	2E3	2N
			BAIT							
PREY	Deconvolution assay	*HIS3*	red	orange	orange	orange	orange		orange	orange
		ADE2	red	red	red	red	orange			orange
	OTUB FL ctrl	*HIS3*	orange	red	red	red	orange		yellow	
		ADE2	yellow	orange			yellow			
	$^{\Delta N}$OTUB1	*HIS3*								
		ADE2								
			PREY							
BAIT	Deconvolution assay	*HIS3*		red			red	red	red	
		ADE2								
	OTUB FL ctrl	*HIS3*		red			red	orange	orange	
		ADE2								
	$^{\Delta N}$OTUB1	*HIS3*		yellow			orange	yellow		
		ADE2								

Chapter Four:
BIOPHYSICAL EVIDENCE OF BINARY E2:OTUB1 COMPLEXES FORMATION

4.1   Introduction

Although the Y2H system is a valuable tool for discovering and analysing protein interactions, data from these studies should only be used as a prediction to inform further experimental or hypothesis driven studies. Therefore, in an attempt to validate our primary Y2H data and provide greater insight into the molecular mechanism or biophysical properties of different OTUB1:E2 complexes, genes of interest (OTUB1, UBE2D2, UBE2E1 and UBE2E2) were cloned into the pETM-11 polyhistidine-tag vector by conventional cloning methods in order to generate large quantities of purified proteins for use in gel filtration chromatography, ITC and NMR studies.

The primary aims of this part of the project were:

1) To express and purify sufficient quantities of recombinant OTUB1 and selected E2 proteins in order to generate sufficient amounts of each protein to facilitate biophysical and structural analysis of protein complexes.

2) To verify the observed Y2H interactions in multiple assay systems.

3) To provide more detailed structural information relating to the molecular organisation of OTUB1, E2 and E3-RING proteins in different protein complexes.

4.2   Producing affinity tagged proteins

Both biophysical and structural techniques require substantial quantities of pure protein. In this study, an inducible bacterial expression system was used to express recombinant proteins containing a cleavable in-frame 6xHis affinity tag. In brief, His-tagged proteins

were first purified by Ni^{2+} affinity chromatography before performing Q/SP ion exchange chromatography to further increase the purity of target proteins.

4.2.1 pETM-11 system

Target genes were initially cloned into the PETM-11 vector under inducible control of the T7/*lac* promoter (Figure 4.1 shows the map of pETM-11). This vector was chosen as it generally provides high expression yields while only adding four amino acids (Gly-Ala-Met-Ala) to the N-terminus of the target protein, with a TEV protease cleavage site prior to the start of the inserted protein-coding sequence.

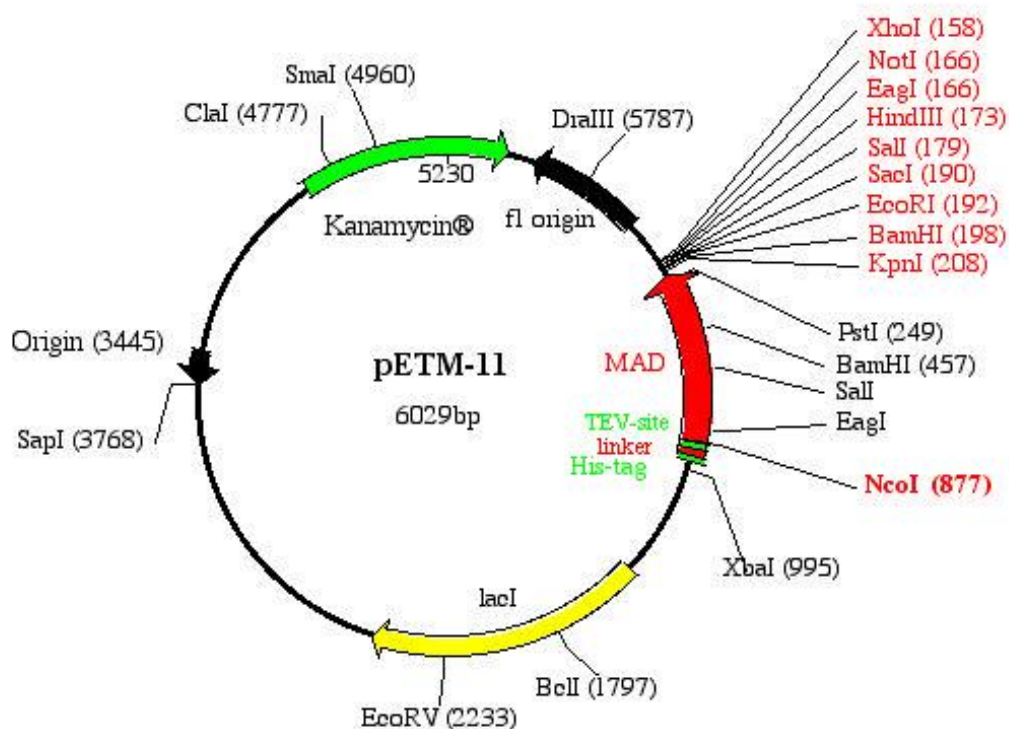**Figure 4.1   Map of pETM-11 vector:** The conserved multiple cloning site (MCS) contains an *Nco*I restriction site which has an in frame methionine codon (ATG) after which protein coding sequences were cloned, a protein called MAD which will be excised and replaced with genes encoding proteins of interest, and a number of restriction sites downstream the MAD protein for double digestion alongside *Nco*I.

4.2.2   Construction of His-tag protein of interest

*4.2.2.1 Introduction of Nco*I *and Hin*dIII *restriction sites into the N- and C-termini of protein coding inserts*

Sequences for the forward and reverse primers designed for each expression construct are shown in Table 4.1. Target genes were PCR-amplified from the pDONR entry vector and PCR products were digested with *Nco*I and *Hin*dIII prior to ligation into pETM-11 digested with the same restriction enzymes.

Table 4.1   *Nco*I and *Hin*dIII restriction site sequences incorporated in forward and reverse primers. Primers were designed to contain *Nco*I (red) and *Hin*dIII (blue) sequences to allow PCR of the targeted genes of interest while incorporating restriction sites for digestion protocol.

	Forward primer	**Reverse primer**
OTUB1	5' CAG GGC G**CC ATG G**CC GCG GCG GAG GAA CCT CAG 3'	5' CCC **AAG CTT** GGG CTA TTT GTA GAG GAT ATC GTA GTG TCC AGG CCG 3'
UBE2D2	5' CAG GGC G**CC ATG G**CC GCT CTG AAG AGA ATC C 3'	5' CCC **AAG CTT** GGG CTA CAT CGC ATA CTT CTG AGT CCA TTC CCG 3'
UBE2E1	5' CAG GGC G**CC ATG G**CC TCG GAT GAC GAT TCG AGG 3'	5' CCC **AAG CTT** GGG CTA TGT AGC GTA TCT CTT GGT CCA CTG TCT G 3'
UBE2E2	5' CAG GGC G**CC ATG G**CC TCC ACT GAG GCA CAA AG 3'	5' CCC **AAG CTT** GGG CTA TGT GGC GTA CCG CTT GGT CCA C 3'
ᐃNOTUB1	5' CAG GGC G**CC ATG G**CC GAG ATT GCT GTG CAG AAC CC 3'	5' CCC **AAG CTT** GGG CTA TTT GTA GAG GAT ATC GTA GTG TCC AGG CCG 3'

4.2.2.2 Ligation and transformation.

Despite the fact that the ligation step is particularly problematic, DNA fragments were successfully inserted into the vector by optimising insert/vector ratios, enzyme and buffer concentrations and also ligation time. As an alternative cloning strategy, both TOPO® (Invitrogen) and In-Fusion® (Clontech) cloning methods were also tried. Ligation products were transformed into α-select *E. coli* cells and PCR-verification was performed to selected positive colonies for insert sequencing. As pETM-11 is a low copy number plasmid,

maxipreps were performed to obtain clones, which were then transformed into Rosetta™(DE3)pLysS host strains, a BL21 derivative suitable for expression of target genes cloned into the pETM-11 vector (Figure 4.2).
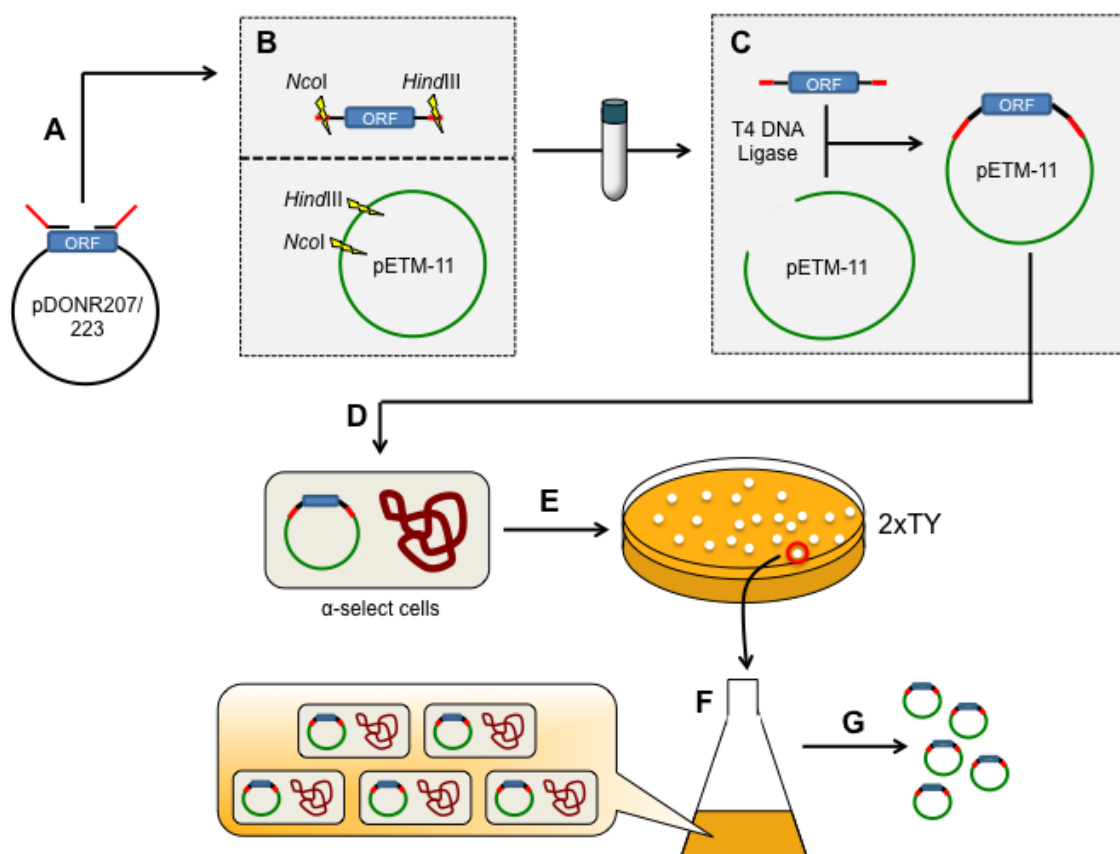
Figure 4.2 **Generation of pETM-11 clones**: **(A)** Target genes were amplified from sequence-verified templates in pDONR207 or pDONR223 vectors using forward and reverse primers containing *Nco*I and *Hin*dIII cleavage sites respectively. **(B)** Digestion of PCR products and pETM-11 plasmid using *Nco*I and *Hin*dIII restriction enzymes to produce compatible sticky ends on both PCR product and plasmid. **(C)** Ligation of PCR product into pETM-11 vector by T4 DNA ligase. **(D)** Transformation of ligated product into α-select cells. **(E)** Transfected clones were selected following colony growth on 2xTY agar containing kanamycin. **(F)** Growth of large-scale cultures. **(G)** Maxiprep to isolate the desired construct ready for transformation into Rosetta cells for expression studies.

4.3   Protein expression and purification

4.3.1   Small-scale expression test

Small-scale pilot experiments were performed to confirm the effectiveness of the pETM-11 expression system. In each case, protein purification was performed using Ni-NTA agarose (Qiagen®) and target proteins were eluted using lysis buffer B (20 mM Na_2HPO_4, 500 mM NaCl, 500 mM imidazole). A high concentration of imidazole, a histidine competitor is required because in excess amount it can displace the histidines from Ni^{2+} co-ordination thereby freeing the His-tagged protein. The result in Figure 4.3 suggests the presence of recombinant proteins of sizes commensurate with all proteins of interest (predicted molecular weight of each construct is shown in Table 4.2). As expected, expressed proteins migrate slightly slower than the predicted molecular weight due to of the addition of the ~3 kDa 6xHis-tag and a TEV cleavage site (MKHHHHHHPMSDYDIPTTENLYFQ), which was uncleaved in this pilot study. Very faint bands can be seen in uninduced samples suggesting that protein expression responded well to IPTG induction. Also, the Ni^{2+} affinity beads appear to clean up the sample quite appreciably.
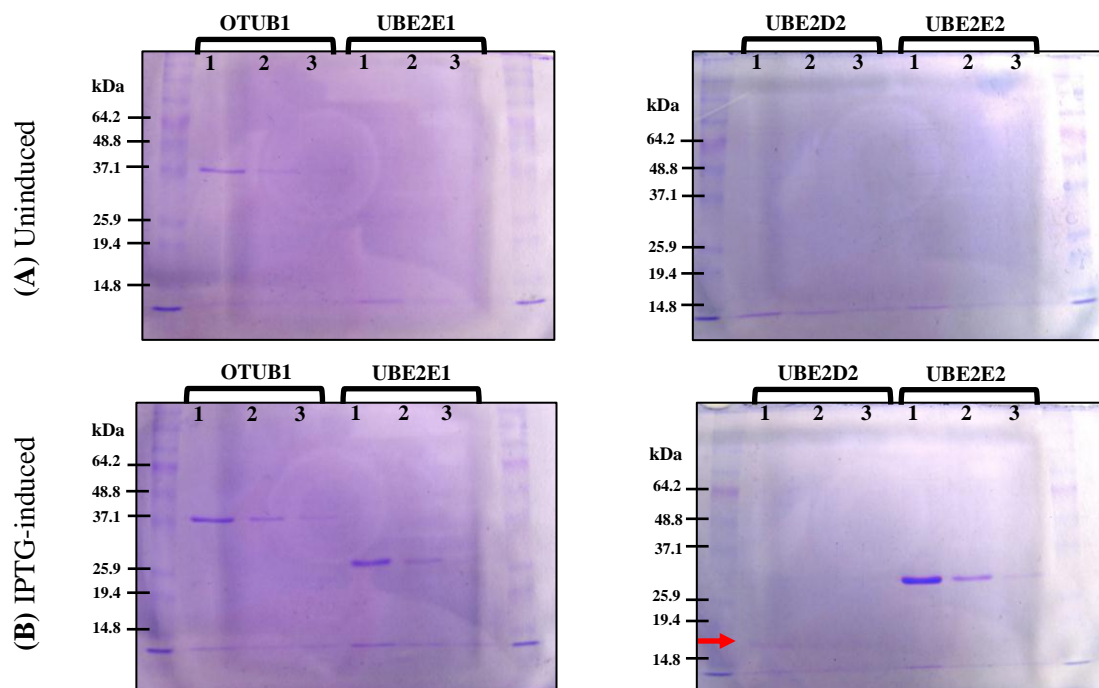
Figure 4.3 **Coomassie blue SDS-PAGE gels of uninduced and IPTG-induced samples from Ni²⁺ column elution:** IPTG-induced samples (B) gave more visible bands compared to uninduced (A). Number 1, 2 and 3 in each protein representing the elution from thricely washed Ni²⁺ agarose beads with lysis buffer B (20 mM Na₂HPO₄, 500 mM NaCl, 500 mM imidazole). All bands show sizes higher than the proteins' molecular weight (OTUB1 31.3 kDa, UBE2E1 21.4 kDa, UBE2E2 22.3 kDa, UBE2D2 16.8 kDa) due to the addition of the 6xHis-tag and a TEV cleavage site. For IPTG-induced UBE2D2 the bands are very faint but noticeable nonetheless, indicated by the red arrow.

4.3.2 Large-scale protein production

4.3.2.1 Selection of host strain and induction temperature

Many factors can influence the expression of proteins in *E. coli*, including the bacterial strain, the vector expression system, growth medium and incubation temperature (Berrow *et al.*, 2006; Gräslund *et al.*, 2008). To optimise protein expression, each construct was expressed in either Rosetta or B834(DE3) cells. In each case, cells were grown until cultures reached an OD_{600} of 0.6 before being induced with 0.2 mM IPTG at either 18°C or 30°C. As both Rosetta and B834 cells lack an outer membrane protease which can degrade proteins during purification (Grodberg and Dunn, 1988), it was expected that expressed proteins should be more stable in these strains. In addition, Rosetta cells also contain tRNA codons rarely used in *E.coli*, which can dramatically increase expression of human proteins (Brinkmann *et al.*, 1989). Analysis of lysates showed that OTUB1 expressed in Rosetta cells at 18°C gave the highest amount of soluble protein meanwhile E2 proteins did not show either host or temperature dependent changes in protein expression profiles. Therefore, all clones were subsequently expressed in Rosetta cells with induction of protein expression being performed at 18°C.

4.3.2.2 Two-step purification process

Lysates prepared from 5 L cultures of IPTG-induced Rosetta cells were subjected to a two-step purification process. Firstly, lysates were injected through a Ni^{2+} affinity column to recover the His-tagged proteins, which were then eluted in lysis buffer B. Fractions containing eluted target proteins were pooled and incubated overnight with TEV protease to remove the His-tags. Following TEV digestion, samples were again injected through Ni^{2+} column to separate the target proteins from cleaved His-tags and residual TEV protease (Figure 4.4).
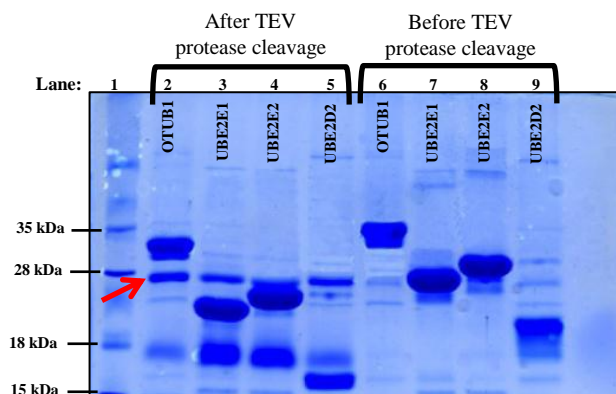
Figure 4.4 **Coomassie blue SDS-PAGE of large-scale protein purification before and after TEV protease cleavage:** Lanes 2-5 show the band of target protein after incubation with TEV protease. TEV is visible as a 27 kDa band (arrow head) accompanying each protein band in the lane. All protein bands are in correct size compared to their respective bands before TEV protease incubation (lane 6-9), which have additional ~3 kDa due to the uncleaved His-tag and TEV site

To remove histidine-rich contaminants, samples were also subjected to cationic or anionic exchange chromatography as indicated in Table 4.2. Peaks representing each protein are demonstrated in Figure 4.5A (Ni^{2+} affinity column) and Figure 4.5B (ion exchange column).

Table 4.2 **Q/SP selections for each protein:** Selection of either cation or anion exchange columns for each protein is based on their pI and buffer pH. (MW): Molecular weight, (pI) Isoelectric point.

Protein	MW (Da)	pI	Buffer pH	Net charge (in buffer)	Column
OTUB1	31284.0	4.85	8.0	-ve	Q anion
UBE2D2	16753.2	7.69	6.5	+ve	SP cation
UBE2E1	21404.18	8.77	6.5	+ve	SP cation
UBE2E2	22255.1	7.59	6.5	+ve	SP cation
$^{\Delta N}$OTUB1	26951.35	5.14	8.0	-ve	Q anion

Figure 4.5(A) and (B) Typical column peaks for (i) OTUB1, (ii) UBE2D2, (iii) UBE2E1 and (iv) UBE2E2: Ni^{2+} affinity column (A) and ion exchange column (B) show very prominent protein peaks measured by UV Absorbance at 280 n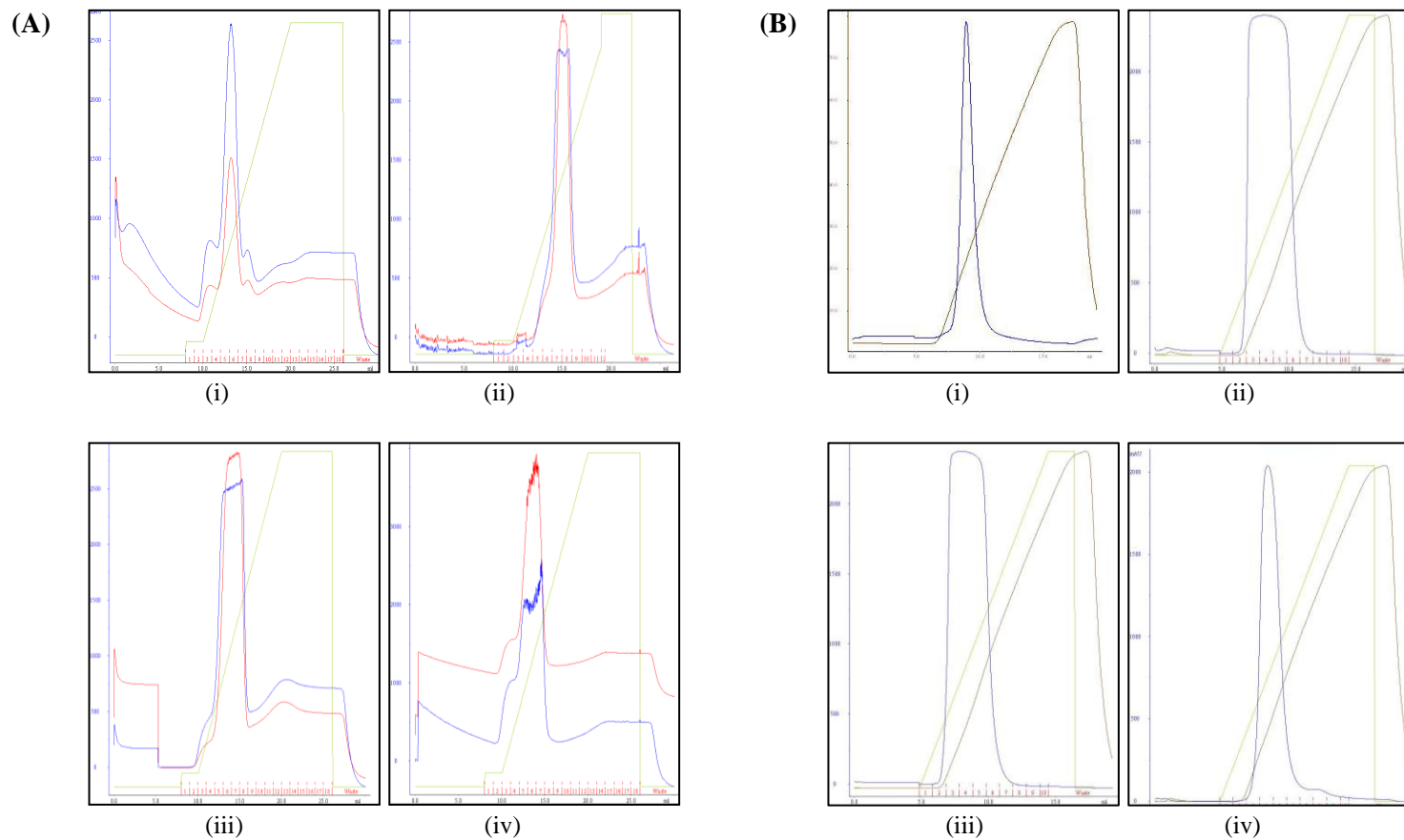m (blue lines). The red lines represent the absorbance by nucleic acids at 260 nm while the rest green and brown lines denote the gradient of increasing imidazole and salt concentrations.

4.3.3 Correct protein folding was analysed by 1D ^{1}H NMR

All purified protein samples were analysed by solution state NMR to ensure the expressed recombinant proteins are not misfolded in solution. The NMR spectrum was collected at 25°C using a Bruker 600 MHz Avance solution-state spectrometer. Samples were analysed at a concentration of 0.2 mM in Tris buffer pH 7.4, mimicking the optimal human physiological pH. Figure 4.6 suggested that all the proteins are folded properly. In general, folded proteins exhibit a range of chemical shifts due to the anisotropic magnetic fields of proximal aromatic or carbonyl groups compared to the applied magnetic field. Thus, if very little chemical shift dispersion is observed, the protein may be unfolded, or may have an unstable structure. However, as all protein spectrums show a wide range of ppm, it can be assumed that they are all appropriately folded. The term 'ppm' refers to the unit of 'parts per million' or one part per 1 000 000 parts, which is used to measure a small increments in magnetic field range in NMR spectroscopy. The magnetic field range displayed in the common NMR diagram is very small compared with the actual field strength hence it is customary to describe this small value using a pseudo units like 'ppm'.

Sharp peaks observed in the spectrum also suggest that all proteins are well-behaved in solution. In general, aggregation causes peaks to be broadened and in extreme cases, formation of protein aggregates can result in no peaks being observed as peaks tend to merge into the baseline. Sometimes low molecular weight impurities (such as imidazole) can also be observed in 1D spectrums as sharp peaks amid the broader envelope of protein resonance. As seen in Figure 4.6, all of the tested spectra have these sharp peaks at 2.6-2.7 ppm, suggesting that sample purity or buffer exchange should be improved. In this 1D NMR experiment, most of the signal overlapped heavily due to the large size of the protein. However, all proteins appeared to be folded properly and therefore suitable for use in subsequent biophysical and structural studies.
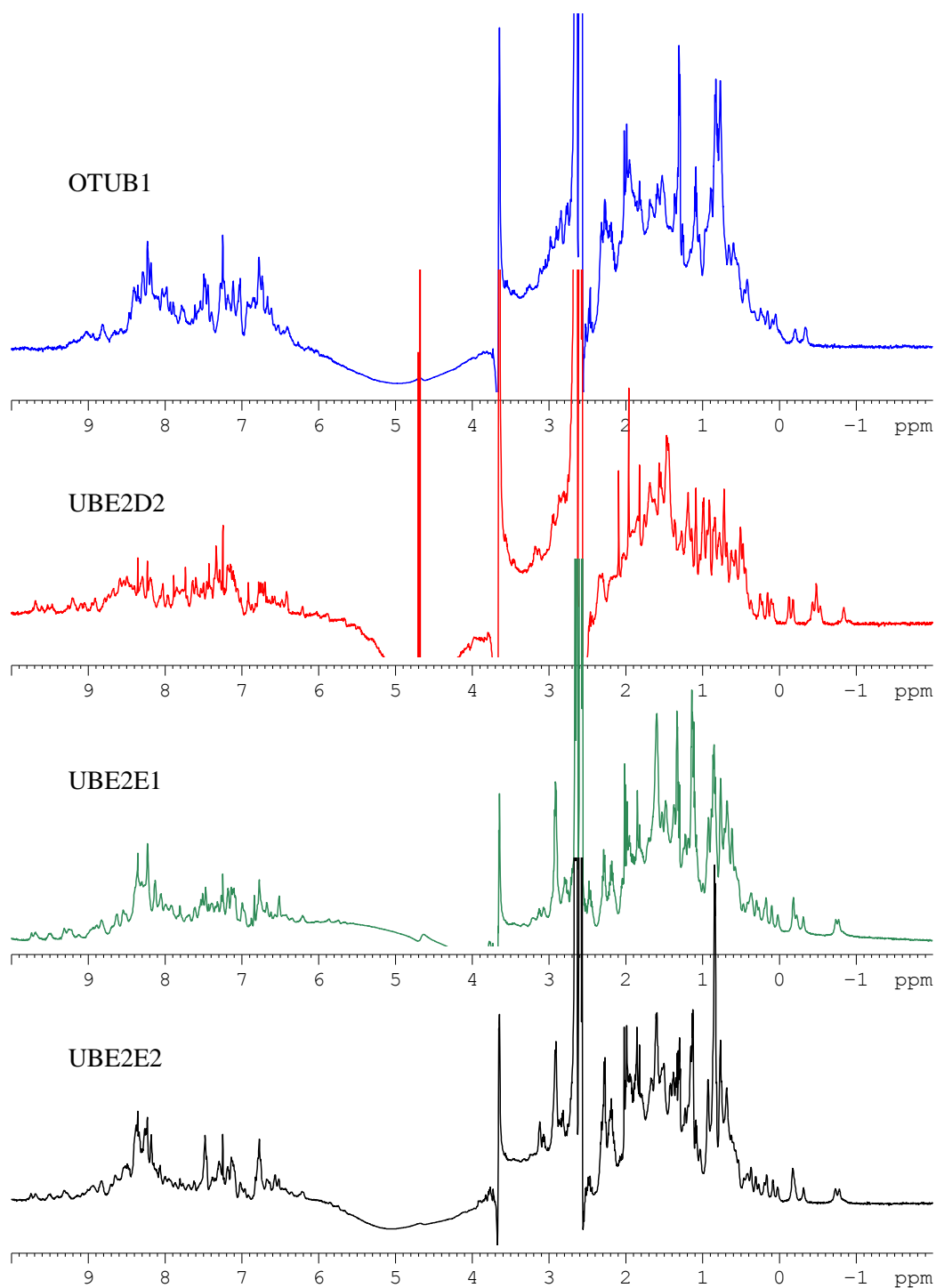
Figure 4.6   **1D-NMR spectrum:** One-dimensional spectra of proteins were acquired at 25°C in a Bruker 600 MHz Avance solution-state spectrometer.

4.4   Preliminary analysis of OTUB1 in complex with UBE2D2, UBE2E1 and UBE2E2

4.4.1   Identification of an OTUB1:UBE2E1 complex by gel filtration chromatography

As an initial approach to validate Y2H data, gel filtration (GF) chromatography was used to observe formation of complexes between predicted protein interaction partners. GF is a simple and reliable chromatographic method which can be used to separate proteins or complexes according to their size. In principle, proteins with higher molecular weights travel faster through the column as they will interact less with the matrix (Sun *et al.*, 2004). In contrast, smaller proteins tend to diffuse into the matrix hence retarding their movement through the column. Thus, proteins are eluted in order of decreasing size. Superdex™ 75 GF column (GE Healthcare) is a composite of dextran and cross-linked agarose, and has the ability to size-fractionate multiprotein complexes of molecular weight in the range of 3-70 kDa. Unlike ion exchange or affinity chromatography, molecules do not bind to the chromatography medium so buffer composition does not directly influence the resolution in GF. Tris-buffer pH 7.4 was used to mimic physiological pH and 0.15 M NaCl was included to avoid non-ionic interactions between the sample and the matrix. To observe interaction between proteins, each GF experiment involved injection of a quantified amount of each protein, both individually and in mixtures with their partner in pre-determined concentration ratios.

In this experiment, quantified amounts of OTUB1 and E2 proteins were injected separately into the column followed by premixed samples consisting two protein partners in a 3:1 E2:OTUB1 ratio. Differences in elution profiles for mixtures of protein with respect to their individual control samples are indicative of complex formation (Buechler *et al.*, 1993). GF analysis demonstrated the presence of potential complex formation between OTUB1 and UBE2E1 but no clear complexes were observed with the other E2s. As shown in Figure 4.7A, there is no significant shift in OTUB1:UBE2D2 mixture as the resolved peak from the mixture was clearly representing individual OTUB1 and UBE2D2 instead of one peak

indicating formation of a protein complex. It should be noted that the final concentration of mixed and single protein samples were set to be similar hence a decreased peak intensity observed in mixed samples was expected due to the presence of less single protein. Interestingly, mixtures of OTUB1:UBE2E1 gave rise to three unresolved peaks (Figure 4.7B); one may represents a complex as it is the first to eluted suggesting a complex bigger than OTUB1 and UBE2E1, while the other two appear to represent individual proteins, as they elute at the same time when compared with each of the corresponding single sample. Because single sample analysis shows only one resolved peak for both OTUB1 and UBE2E1, it is unlikely that the first elution peak represents a dimer of either OTUB1 or UBE2E1. Theoretically, when 31 kDa OTUB1 and 21 kDa UBE2E1 form a complex, it is estimated to be 52 kDa in size but the first peak was eluted very close to the single OTUB1 peak suggesting the size is not far from ~31 kDa. This may be explained by the fact that GF is fractionating the molecules based solely on their size, which in some cases can be biased by the protein conformation rather than simple molecular weight. In Figure 4.7C, OTUB1:UBE2E2 also does not give convincing evidence of complex formation because the peak pattern observed in the mixed sample does not show any significant shift.

The absence of complex formation especially in the case of OTUB1:UBE2D2 mixture was surprising because in Y2H experiments OTUB1 shows the strongest interaction with UBE2D2. However, considering that Y2H is binary system, possibly no other factors were needed for protein interaction meanwhile *in vitro* experiment may requires other participants apart from the two proteins to establish an interaction.
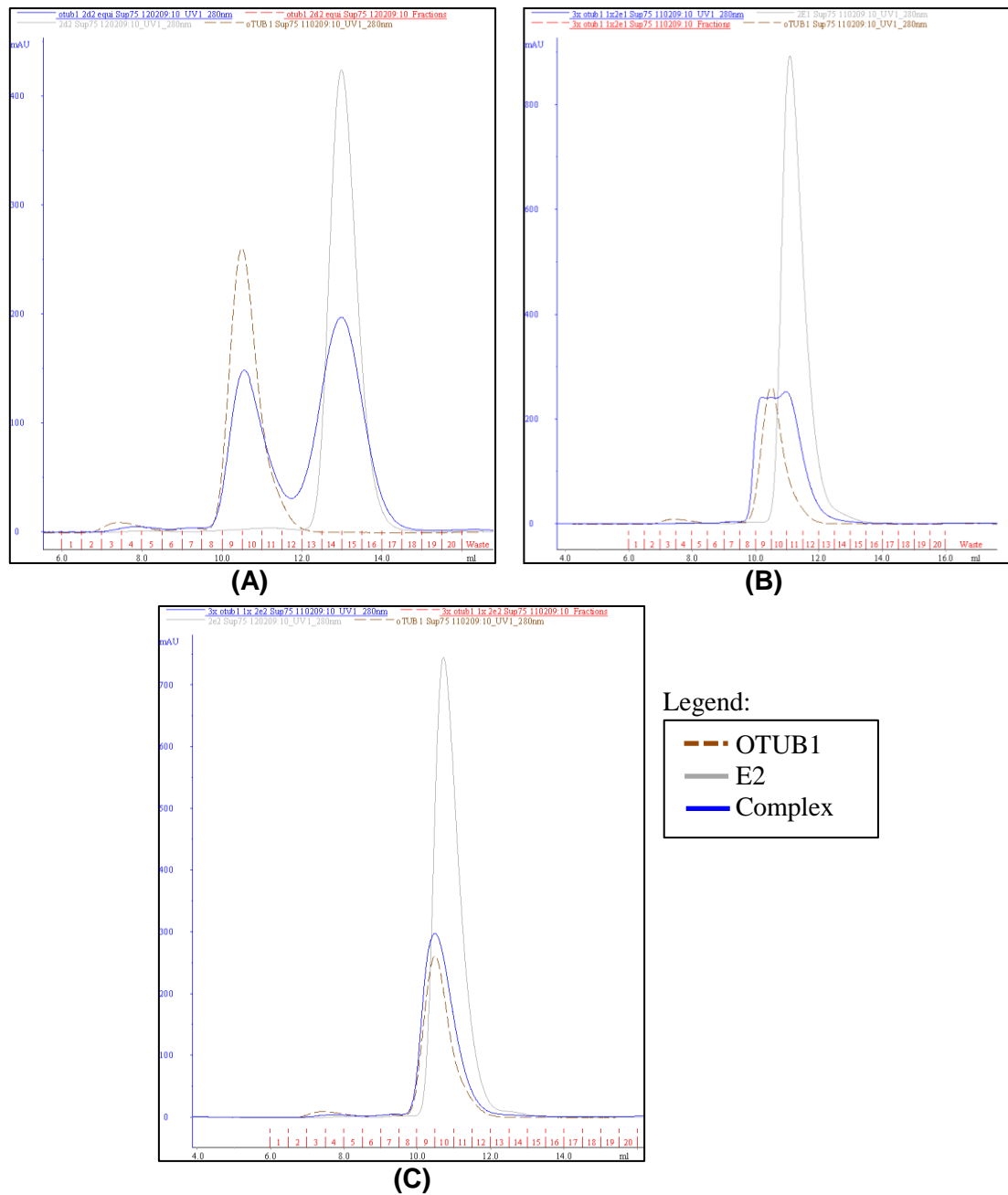
Figure 4.7   Typical gel filtration column peaks: (A) OTUB1:UBE2D2 **(B)** OTUB1:UBE2E1 **(C)** OTUB1:UBE2E2 in Tris buffer pH 7.4. The dotted brown peaks represent OTUB1 protein alone, grey lines denote distribution of E2 proteins alone and blue peaks correspond to the mixture of OTUB1 and its E2 partner.

4.4.2   ITC analysis of potential OTUB1:E2 complexes

Initial ITC and NMR experiments were performed on the OTUB1:UBE2E1 complex, as this complex showed promising complex formation in GF analysis. ITC was performed to define the binding affinity, mechanism and stoichiometry of proteins within complexes. Obtaining an ITC profile for the OTUB1:UBE2E1 complex proved to be challenging. Since the K_d for this complex is unknown, the concentration of component proteins may require optimisation. An initial binding curve was observed when 0.73 mM UBE2E1 (in syringe) was titrated into 75 μM OTUB1 (in cell) in 40 injections of 2 μl each. The ITC heat dilution curve, after subtracting background (established by titrating UBE2E1 into buffer without OTUB1) produced a final binding curve with a $K_d = 0.15$ mM, which generally indicates very weak binding (Figure 4.8). However, the curve was not considered satisfactory as the heat release was still detected when samples reached saturation, which is often a sign of buffer mismatch. The signal to noise ratio was very small suggesting a higher protein concentration was required to produce more significant peaks. Also, the small (near zero) ΔH value could suggest the reactions are unmeasurable or that a higher protein concentration is needed to increase ΔH values and enable prediction of the mechanisms of interaction. The data was fitted to a one-site binding model and although the n value of 0.00372 would not suggest a 1:1 stoichiometry, this can be adjusted by reducing the amount of injection, seeing that saturation was reached at less than 40 injections.

To generate a more reliable ITC binding curve, higher concentrations of protein were required to detect a significant heat release. However, when the concentration of UBE2E1 was increased to 0.15 mM, it produce a curve which implies that double binding activities were happening (Figure 4.9). The most likely explanation for this result is the possibility that UBE2E1 might form dimers in high concentration.

(A)
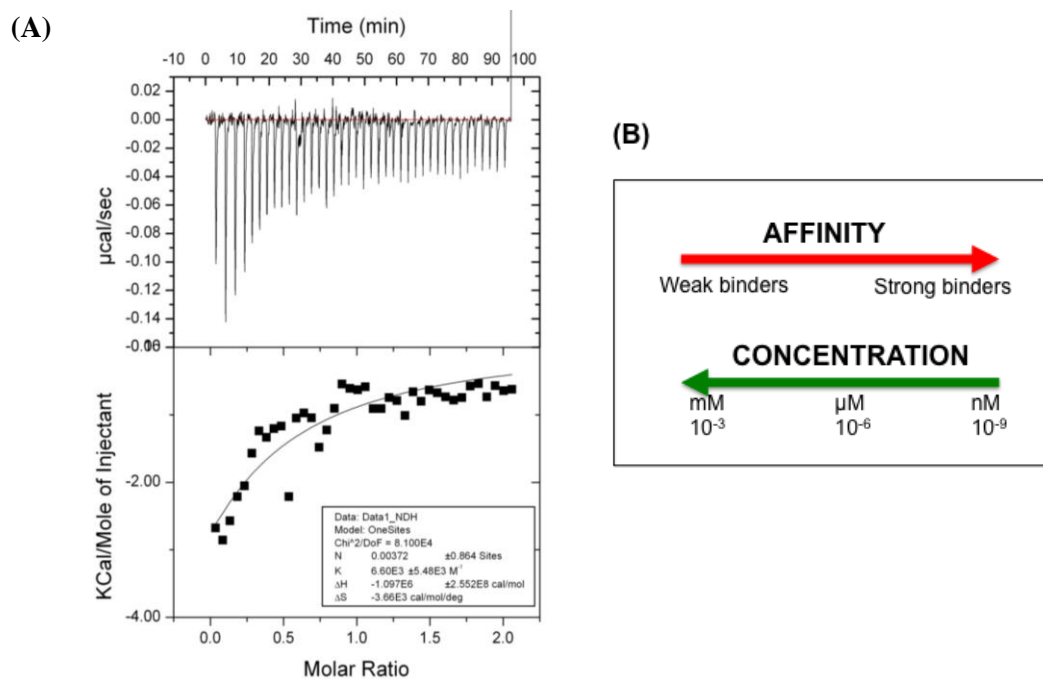
Figure 4.8   ITC titration of UBE2E1 with OTUB1: (A) Experiment was performed in phosphate buffer pH 7.4 at 25°C. The concentrations of reactants are 0.73 mM UBE2E1 (in syringe) and 75 μM OTUB1 (in cell). The solid line corresponds to theoretical curves with n = 0.00372, K_a = 6.6x10^{-3} M^{-1} and ΔH = 1.1x10^{-6} cal/mol. **(B)** Dissociation constant = $1/K_a$, lower K_d values indicate tighter binding and therefore a greater affinity and vice versa.
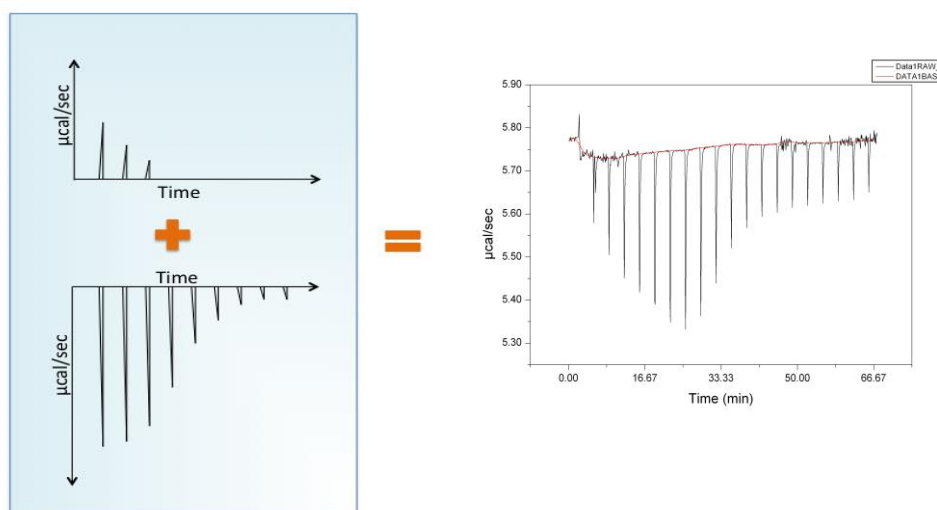
Figure 4.9   UBE2E1 may form dimers at high concentration: Beside the heat release from OTUB1:UBE2E1 binding interaction, at higher concentrations of UBE2E1 (0.15 mM) there is evidence of a different binding mode.

When catalysing polyubiquitination, there is a high possibility for some E2 enzymes to act as dimers because at least one of the E2 monomers could remain associated with the substrate while maintaining continuous additions of ubiquitin monomers as an intramolecular reaction. The sequential addition of ubiquitin will probably require multiple cycles of E2-E3 binding and release. Also, the presence of a stable E2 dimer in the E2-E3 complex might circumvent the need for complete dissociation of the E3 from the E2 (David *et al.*, 2010).

To avoid any possibility of dimerisation, injection parameters were set in such a way that the protein concentration could not be increased. However, at this stage of the study, a decision was made to generate a truncated version of OTUB1 upon looking at the 2D-NMR spectrum of ^{15}N-OTUB1 (discussed below). Hence it was decided that subsequent ITC experiments should be performed using the N-terminal truncated version of OTUB1.

4.4.3 (^{1}H^{15}N)-HSQC NMR experiment

4.4.3.1 ^{15}N-labelled protein

Isotopic-labelled proteins were produced by expression of target proteins in bacteria grown on minimal medium supplemented with ^{15}NH$_4$Cl as the sole nitrogen source and common glucose (McIntosh and Dahlquist, 1990). Basically, the minimal media contained all salts and trace elements required by the bacteria except for a carbon or nitrogen source. The ^{15}N-label is the simplest and cheapest form of label, which is mainly used to record the standard solution NMR HSQC spectrum. The HSQC spectrum is like a fingerprint of the protein because each peak corresponds to an NH of each amino acid residue. Highly purified ^{15}N-labelled protein can be prepared for ^{15}N-HSQC in which the folding of protein could be checked and the quality of the spectrum could be assessed to determine whether it is worth recording other spectra and proceeding on to more expensive labelling schemes such as ^{13}C labelling. ^{15}N-labelled protein can also be useful for titrations with other proteins with which it forms a complex. Double-labelled proteins, or ^{15}N^{13}C-labelling on the other hand were produced by growing bacteria in minimal media supplemented with ^{15}NH$_4$Cl and ^{13}C glucose (McIntosh and Dahlquist, 1990). This form of labelling enables straightforward assignment of both the backbone and side-chain ^{1}H, ^{13}C and ^{15}N atoms using triple-resonance spectra. A high proportion of these assignments are required to precisely calculate the protein structure (McIntosh and Dahlquist, 1990).

Initially, ^{15}N-labelled OTUB1 protein was successfully produced and HSQC titration analysis performed on this protein with unlabelled UBE2E1 shows a distinct chemical shift indicating direct protein interaction. Double ^{15}N^{13}C-labelled protein was produced once but gave rather a weak signal. We therefore resorted to production of ^{15}N-labelled protein to optimise OTUB1 expression.

4.4.3.2 Investigating the OTUB1:UBE2E1 complex

(^{1}H^{15}N)-HSQC spectra were recorded on ^{15}N-OTUB1 in the absence and presence of unlabelled UBE2E1. From the 271 residues that make up OTUB1 protein, about 220 peaks are visible, suggesting that the protein was in a folded form since the spectra showed signs of dispersion. However, the rest of the peaks were observed as disordered and dispersed signals which overlapped with each other and therefore could not be counted. As seen in Figure 4.10, a very dense spectrum was obvious in the region between 7.6 and 8.5 ppm on the proton chemical shift scale. Peaks in this region usually arise from residues that are in unstructured and flexible regions resulting in their brisk tumble movement. In contrast, residues held rigid by the structure tend to have very broad peaks indistinguishable from the background. This observation led to the decision to generate the truncated OTUB1 without the N-terminal region, which had previously been identified as an unstructured region (discussed in Chapter 3) in order to get a cleaner spectrum to evaluate the chemical shifts.

Chemical shifts, or the changes in a peak position and intensity in an NMR HSQC spectrum indicate a change in the environment of the corresponding residue due to the protein being in a contrasting free or bound form. The proximity of the unlabelled partner in the complex will have the effect of modifying the environment of residues at the interface of the complex. As a result, residues involved in or affected by binding will have a different chemical shift than observed in the unbound form. To observe the chemical shifts upon formation of the OTUB1:UBE2E1 complex, a 2D-HSQC spectrum was acquired for uniformly deuterated ^{15}N-OTUB1 with a titration of UBE2E1 in 1:1 ratio (Figure 4.10)
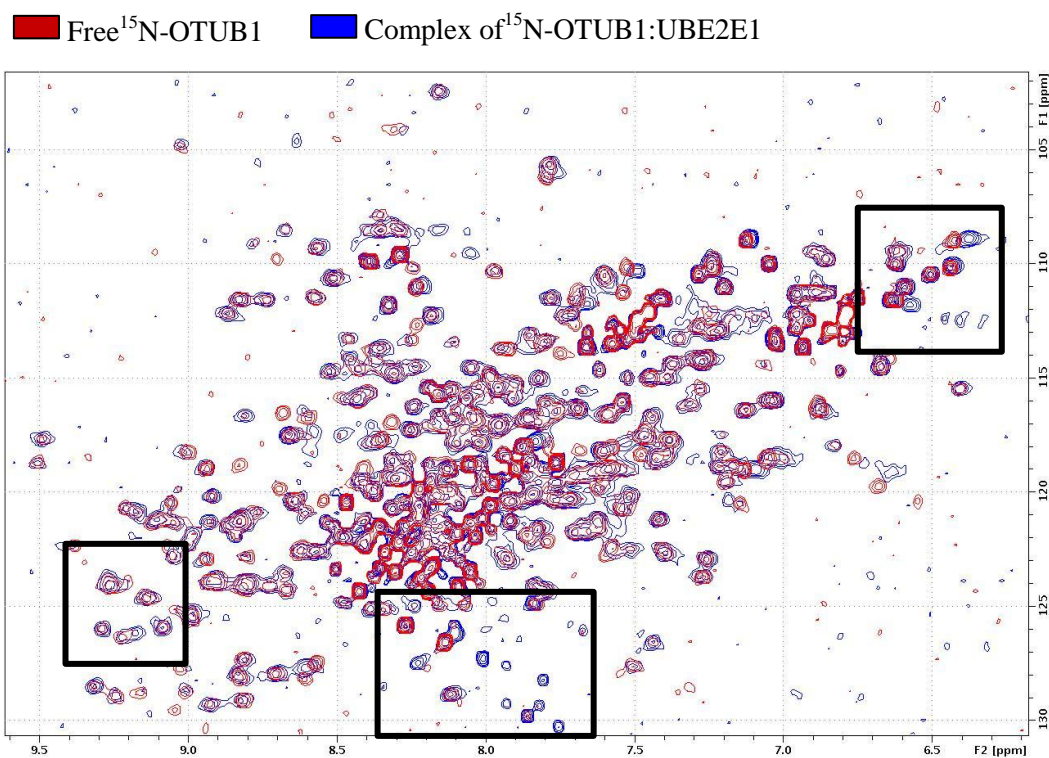
Figure 4.10   NMR titration ^{15}N-OTUB1:UBE2E1: (^{1}H^{15}N)-HSQC spectra of ^{15}N-OTUB1:UBE2E1 (0.2 mM each at pH 7.4) recorded at 25°C using Bruker 800 MHz Avance spectrometer. Red peaks were coming from the free^{15}N-OTUB1 while the complex of ^{15}N-OTUB1:UBE2E1 was represented by blue peaks. The black boxes emphasizing the area with the most noticeable chemical shifts.

Comparison of the spectra acquired for free ^{15}N-OTUB1 with the ^{15}N-OTUB1:UBE2E1 complex revealed a series of clear chemical shift perturbations induced by UBE2E1 binding. The perturbations of signals, ranged from differences in chemical shifts, increased line width, decreased peak intensity, through to complete loss of signal (Figure 4.10) suggesting specific points of contact between OTUB1 and bound UBE2E1. As yet, the perturbations could not be mapped and the exact amino acids involved in the interactions could not be determined due to the fact that amino acids within the OTUB1 spectrum have not yet been assigned and the disordered region within OTUB1 induced a dense central cluster of ambiguous peaks. As a result, a N-terminal truncated OTUB1 was generated as discussed in Chapter 3.

4.5   Analysis of the $^{\Delta N}$OTUB1:E2 complex

4.5.1 $^{\Delta N}$OTUB1 expression and purification

The $^{\Delta N}$OTUB1 protein was expressed and purified in the same manner as wild type OTUB1. For ion exchange chromatography, although the pI of $^{\Delta N}$OTUB1 was slightly increased to 5.14 compared to 4.85 for wild type OTUB1, the Q column and Tris-buffer pH 8.0 were still capable of retaining the protein in the column. Removal of the N-terminal gives a clearer NMR spectrum due to the elimination of the dense central part from disordered N-terminal (Figure 4.11).
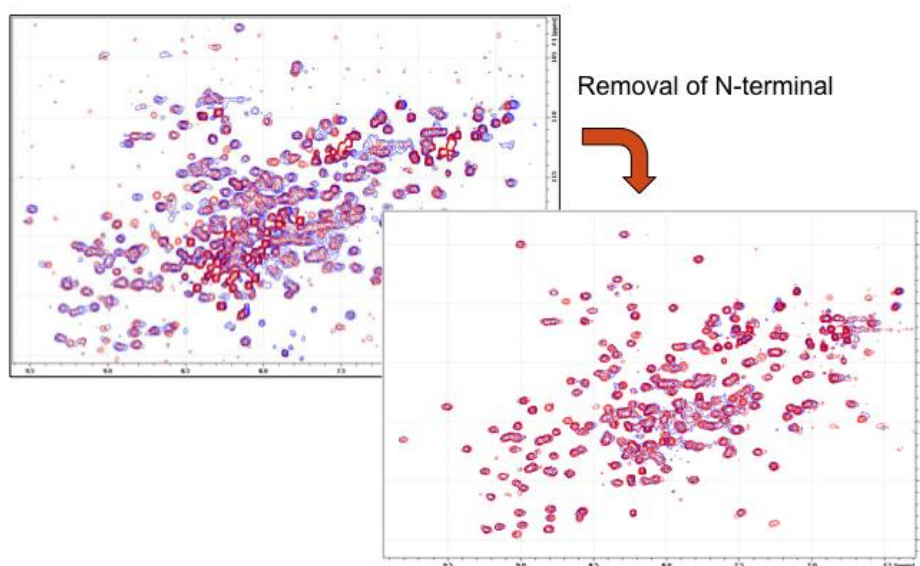
Figure 4.11   NMR spectrum of $^{\Delta N}$OTUB1: The elimination of the dense spectrum region caused by the unstructured N-terminal region of OTUB1 makes the NMR spectrum easier to evaluate.

4.5.2   Improved gel filtration chromatography data

Gel filtration chromatography experiments were repeated using $^{\Delta N}$OTUB1 in combination with either UBE2D2 or UBE2E1 proteins. Both showed evidence of complex formation when mixed in a 1:1 ratio (Figure 4.12). Premixing of $^{\Delta N}$OTUB1 and UBE2D2 in a 1:1 ratio gave rise to a significantly shifted peak, which elutes earlier than $^{\Delta N}$OTUB1 suggesting the formation of a $^{\Delta N}$OTUB1:UBE2D2 complex. A prominent small peak eluted just before the complex peak reached baseline indicating an excess of unbound UBE2D2. The same model can be seen in $^{\Delta N}$OTUB1:UBE2E1 interaction, where a significantly shifted peak corresponding to a protein complex was observed. Surprisingly, when both single proteins were run independently, $^{\Delta N}$OTUB1 (26.951 kDa) was eluted later than UBE2E1 (21.404 kDa) even though it is bigger than UBE2E1 in size. This may be an example of how gel filtration chromatography patterns are dependent on both size and shape. To get information about the stoichiometry of complexes observed in GF studies and estimate the absolute mass of complexes, results obtained by gel filtration chromatography could be further explored using the SEC MALLS (**S**ize **E**xclusion **C**hromatography–**M**ulti-**A**ngle **L**aser **L**ight **S**cattering) method, where a gel filtration column is attached to light scattering equipment (Oliva *et al.*, 2004). Unfortunately, this step could not be taken due to the limited amount of samples.

Legend:

— —· $^{\Delta N}$OTUB1   = = = E2   —— Complex

(A)   **(B)**
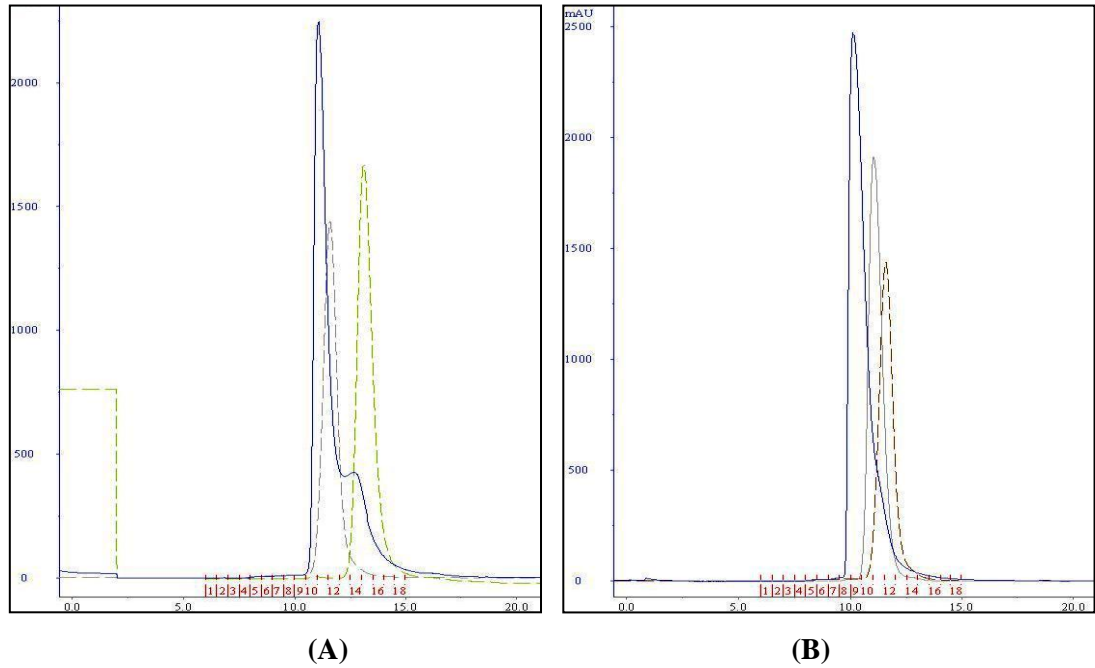
Figure 4.12   **Typical gel filtration column peak: (A)** $^{\Delta N}$OTUB1:UBE2D2, green line represents the single UBE2D2 protein, brown line represents $^{\Delta N}$OTUB1 while the blue peak represents the complex of both proteins. **(B)** $^{\Delta N}$OTUB1:UBE2E1, brown peak is $^{\Delta N}$OTUB1, the grey one is single UBE2E1 and the blue peak corresponds to the complex of both proteins. All samples were run in Superdex™ 75 GF column and eluted in Tris buffer pH 7.4.

4.5.3   Analysis of $^{\Delta N}$OTUB1:UBE2D2 and $^{\Delta N}$OTUB1:UBE2E1 thermodynamic profiles by ITC

As stated in Section 4.4.2, preliminary analysis of UBE2E1 implied the formation of dimers at higher protein concentrations. This observation was supported by the fact that titration of 0.5 mM UBE2E1 into buffer also produced a detectable heat release. For this reason, UBE2E1's role as the 'ligand' in the syringe was replaced by $^{\Delta N}$OTUB1. In this experiment, 2.1 mM $^{\Delta N}$OTUB1 (in syringe) was titrated into 0.1 mM UBE2E1 (in cell) at 25°C cell temperature. The total injection was reduced from 40 to 15, in order to achieve 1:1 stoichiometry. As can be seen in Figure 4.13, both $^{\Delta N}$OTUB1:UBE2D2 and $^{\Delta N}$OTUB1:UBE2E1 complexes appear to have affinity interaction with K_d = 3.89 μM and 16.5 μM, respectively. The smaller unit of concentration for $^{\Delta N}$OTUB1:UBE2D2 implies that UBE2D2 binds to $^{\Delta N}$OTUB1 with a slightly greater affinity than UBE2E1. Interestingly, these results are consistent with data from our initial Y2H studies where UBE2D2 always showed the strongest interaction profile. Negative heat flow and decrease in enthalpy ($\Delta H < 0$) in ITC measurements denotes that both complexes have exothermic reactions in which the heat is lost to the surroundings upon protein interaction. The one-site models and the value of $n \approx 1$ for both complexes clearly suggest that $^{\Delta N}$OTUB1:UBE2D2 and $^{\Delta N}$OTUB1:UBE2E1 both adopted a 1:1 stoichiometry. Even though the ITC profiles were greatly improved compared to the experiment with full-length OTUB1, we cannot conclude that $^{\Delta N}$OTUB1 has stronger affinity for E2s than FL OTUB1 because ITC experiments were performed using different parameters. Unfortunately, it was not possible to repeat the experiment with both versions of OTUB1 in comparison due to time restrictions and subsequent problems with protein production.
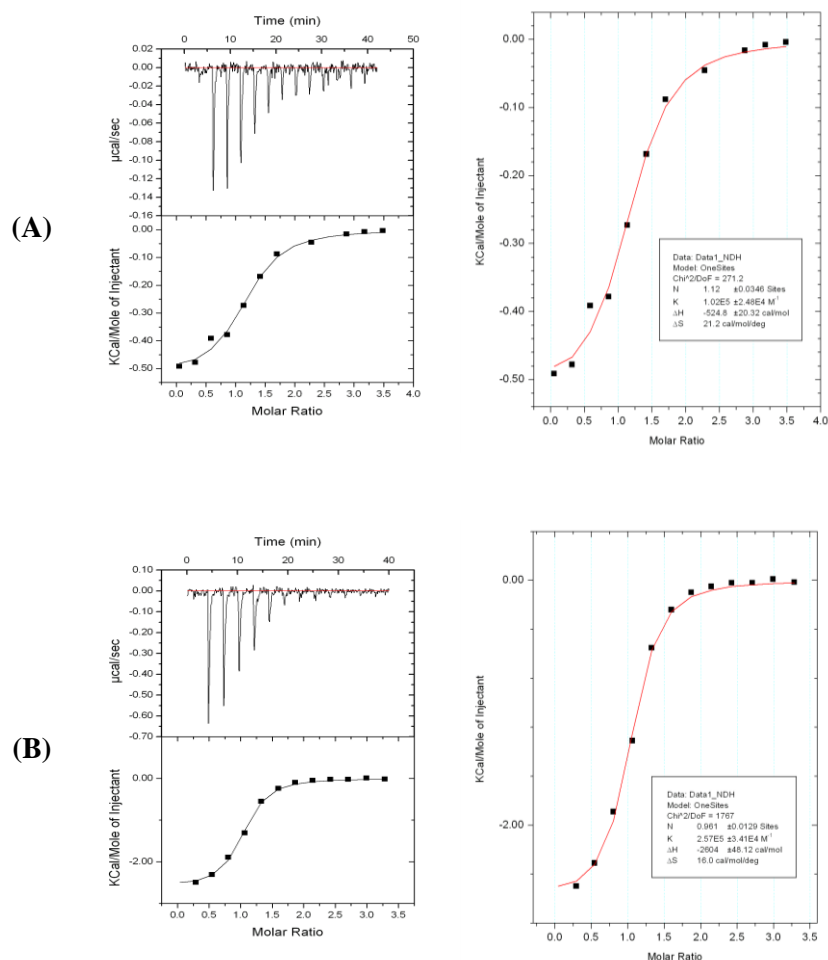
Figure 4.13 **ITC titration of UBE2E1 and UBE2D2 with $^{\Delta N}$OTUB1:** Experiments were performed in phosphate buffer pH 7.4 at 25°C using 2.1 mM $^{\Delta N}$OTUB1 (in syringe) and 0.1 mM UBE2E1 or UBE2D2 (in cell). **(A)** The solid line corresponds to the $^{\Delta N}$OTUB1:UBE2E1 binding curve with $n = 1.16$, $K_a = 6.04 \times 10^{-4}$ M^{-1} and $\Delta H = -0.576$ kcal/mol. **(B)** The solid line corresponds to the $^{\Delta N}$OTUB1:UBE2D2 binding curve with $n = 0.961$, $K_a = 2.57 \times 10^{-5}$ M^{-1} and $\Delta H = -2.604$ kcal/mol.

4.5.4 Identification of possible points of contact in $^{\Delta N}$OTUB1:UBE2D2 and $^{\Delta N}$OTUB1:UBE2E1 complexes observed by shifts in (^{1}H^{15}N)-HSQC NMR spectra

As discussed above (Section 4.4.3.2), during titration, the chemical environment of amino acid residues can change due to the protein being in the free or bound form, or a mixture of both. Figure 4.14 shows the peaks from $^{\Delta N}$OTUB1:UBE2D2 titration in free form, 1:1 ratio and 1:2 (excess of unlabelled partner) ratio. Many of the signals from the complex (green and red) were shifted and even completely disappeared when compared to the black (free forms) signals, indicating residues which may be involved in binding. The same pattern was observed in Figure 4.15 where red and blue signals show some chemical shifts suggesting points of contact in the $^{\Delta N}$OTUB1:UBE2E1 complex. To identify which amino acids correspond to shifted residues, NMR assignment on ^{15}N^{13}C-labelled protein should be performed. Also, because both complexes were large there were some overlapping signals in the ^{15}N spectrum which could only be resolved by a three-dimensional spectrum. For the $^{\Delta N}$OTUB1:UBE2D2 complex, comparison of the green signal (spectrum of 1:1 titration recorded for 1 hour) with the red signal (spectrum of 1:2 ratio recorded for 15 hours), it can be suggested that binding is not transient because the red signal did not return to the free form (black signal) over time.

The exchange rates, as monitored by the NMR resonance characteristics, can be classified as slow, intermediate and fast on the chemical shift timescale. At substoichiometric quantities of the unlabelled partner, the NMR signals resonate at a position between those from the free and fully complexed states, which reflects the mole fraction of both the free and bound form. If the two different states are interchanging rapidly, the effect on the spectra will be to produce a single peak at a position reflecting the average populations of each state; this is fast exchange. If the spin are in slow exchange each state will give rise to separate peaks. In intermediate exchange a broad peak will arise between positions representative of either state. Analysis of the exchange rates for both complexes show that the shifted and loss of signal occur as early as 1:0.25 titration. Supposedly, with further additions of the unlabelled

partner, the resonance position should move closer to that in the bound state. However, in this experiment the signals stay shifted/lost for the full duration of the titration, suggesting slow exchange and tight binding.

4.5.4.1 Are $^{\Delta N}$OTUB1 not fully active?

The conversion from free form to bound form appears to occur very early in titration (ratio 1:0.25), which is far from a 1:1 ratio. The ITC experiment suggested that the OTUB1 was not fully active because up to 2.1 mM OTUB1 was used to titrate 0.1 mM E2s (a 21-fold excess of OTUB1) in order to achieve a 1:1 stoichiometry. This is twice the concentration suggested for ITC where usually for a complex with a micromolar K_d, a ten-fold concentration of sample in the syringe compared to the cell is recommended. For this reason, the unbalanced ratio in $^{\Delta N}$OTUB1:E2 complexes maybe be due to less $^{\Delta N}$OTUB1 in an 'active form' being available for binding. In addition, the ^{15}N-E2s titrated with $^{\Delta N}$OTUB1 also gave a signal shifted/lost at 1:0.25 which remained throughout the titration, strongly supporting the hypothesis of $^{\Delta N}$OTUB1 is only partly active.
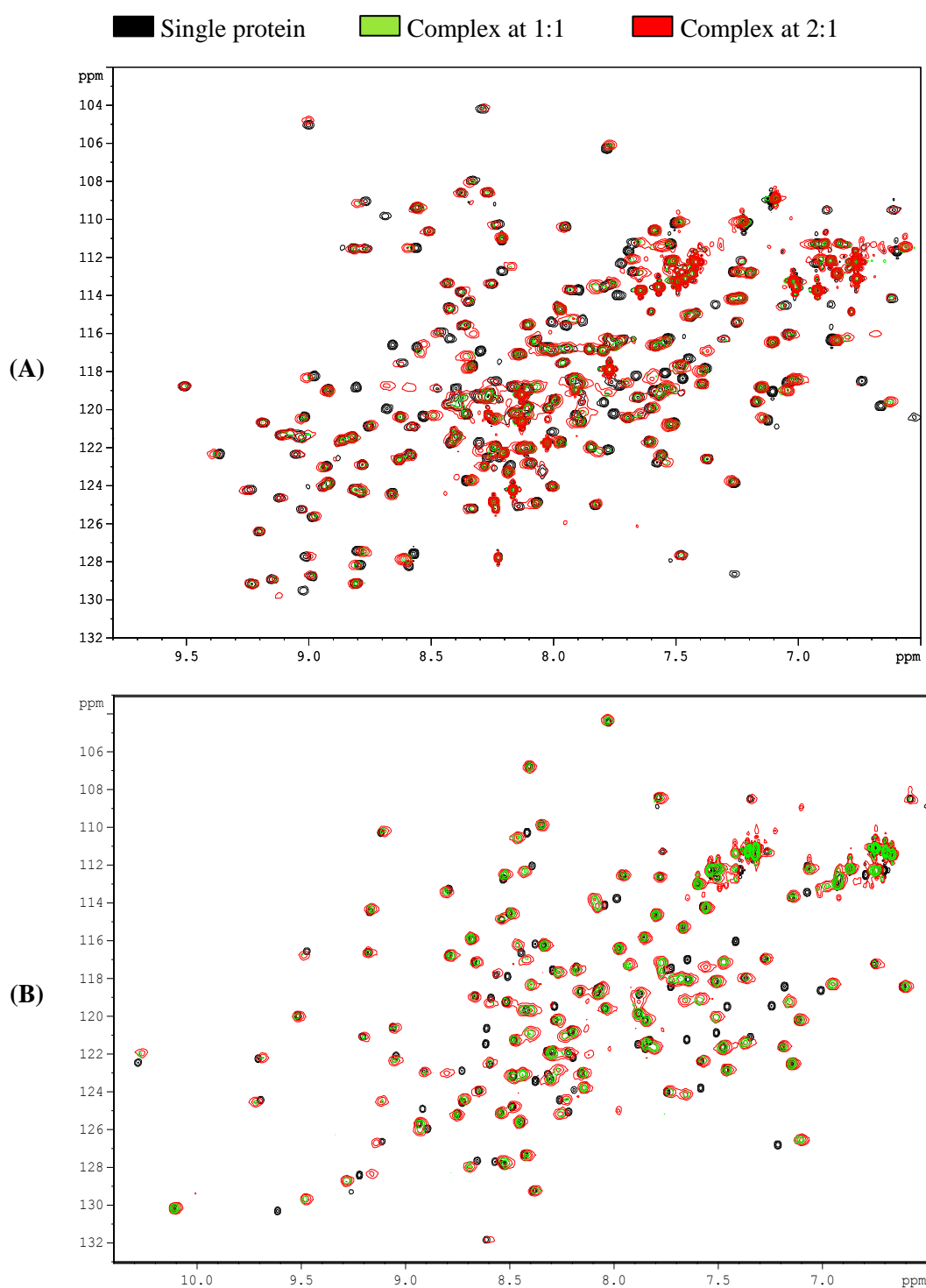
Figure 4.14   2D NMR titration of $^{\Delta N}$OTUB1:UBE2D2: **(A)** (^{1}H^{15}N)-HSQC spectra of ^{15}N-$^{\Delta N}$OTUB1:UBE2D2, and **(B)** (^{1}H^{15}N)-HSQC spectra of ^{15}N-UBE2D2:$^{\Delta N}$OTUB1 recorded at 25°C using a Bruker 800 MHz Avance spectrometer. Black signals represent 0.2 mM free proteins, green are from the complex at 1:1 ratio recorded for 1 hour and red signals represent the complex at 1:2 ratio (excess of unlabelled partner) incubated in the spectrometer for 15 hours.
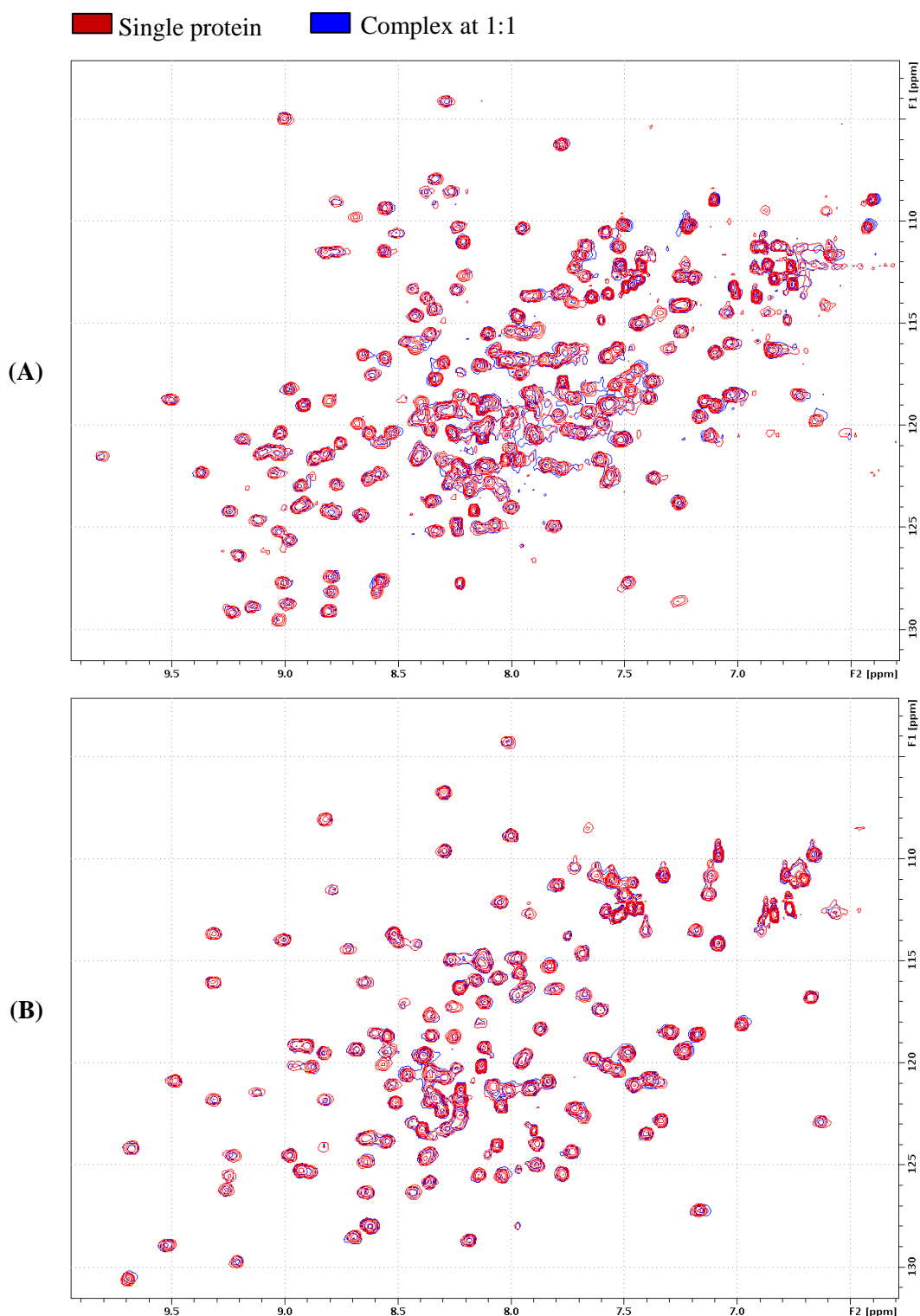
Figure 4.15   2D NMR titration ^{ΔN}OTUB1:UBE2E1: (A) (^{1}H^{15}N)-HSQC spectra of ^{15}N-^{ΔN}OTUB1:UBE2E1, and **(B)** (^{1}H^{15}N)-HSQC spectra of ^{15}N-UBE2E1:^{ΔN}OTUB1 recorded at 25°C using a Bruker 800 MHz Avance spectrometer. Red signals are derived from the 0.2 mM free proteins while blue signals represent signals from the complex at 1:1 ratio incubated in the spectrometer for 1 hour.

4.5.4.2 UBE2D2 and UBE2E1 share a similar interface in binding with ^{ΔN}OTUB1

(^{1}H^{15}N)-HSQC spectral superposition of the single ^{15}N-^{ΔN}OTUB1, ^{15}N-^{ΔN}OTUB1:UBE2D2 and ^{15}N-^{ΔN}OTUB1:UBE2E1 complexes show that ^{ΔN}OTUB1 is bound in a very similar position and orientation relative to both UBE2D2 and UBE2E1. In Figure 4.16, signals representing free ^{ΔN}OTUB1 (blue) were either shifted or lost when in complex with UBE2D2 (green) and UBE2E1 (red). Interestingly, the occurrence of chemical shifts at the same residues suggests that ^{ΔN}OTUB1 uses the same surface to bind with both E2s. However, approximately 3-4 peaks have only blue (free) and red (^{ΔN}OTUB1:UBE2E1) signals while green (^{ΔN}OTUB1:UBE2D2) was lost, and in one peak only blue (free) and green (^{ΔN}OTUB1:UBE2D2) were observed. This suggests that although both E2s bind to the same surface, there are selective amino acids in ^{ΔN}OTUB1 that only bind to either UBE2D2 or UBE2E1. The similarity of ^{ΔN}OTUB1 binding surface with both E2s are not surprising because both E2s are very closely related in sequence and tend to also have common E3-RING binding protein partners (Markson *et al.*, 2009).
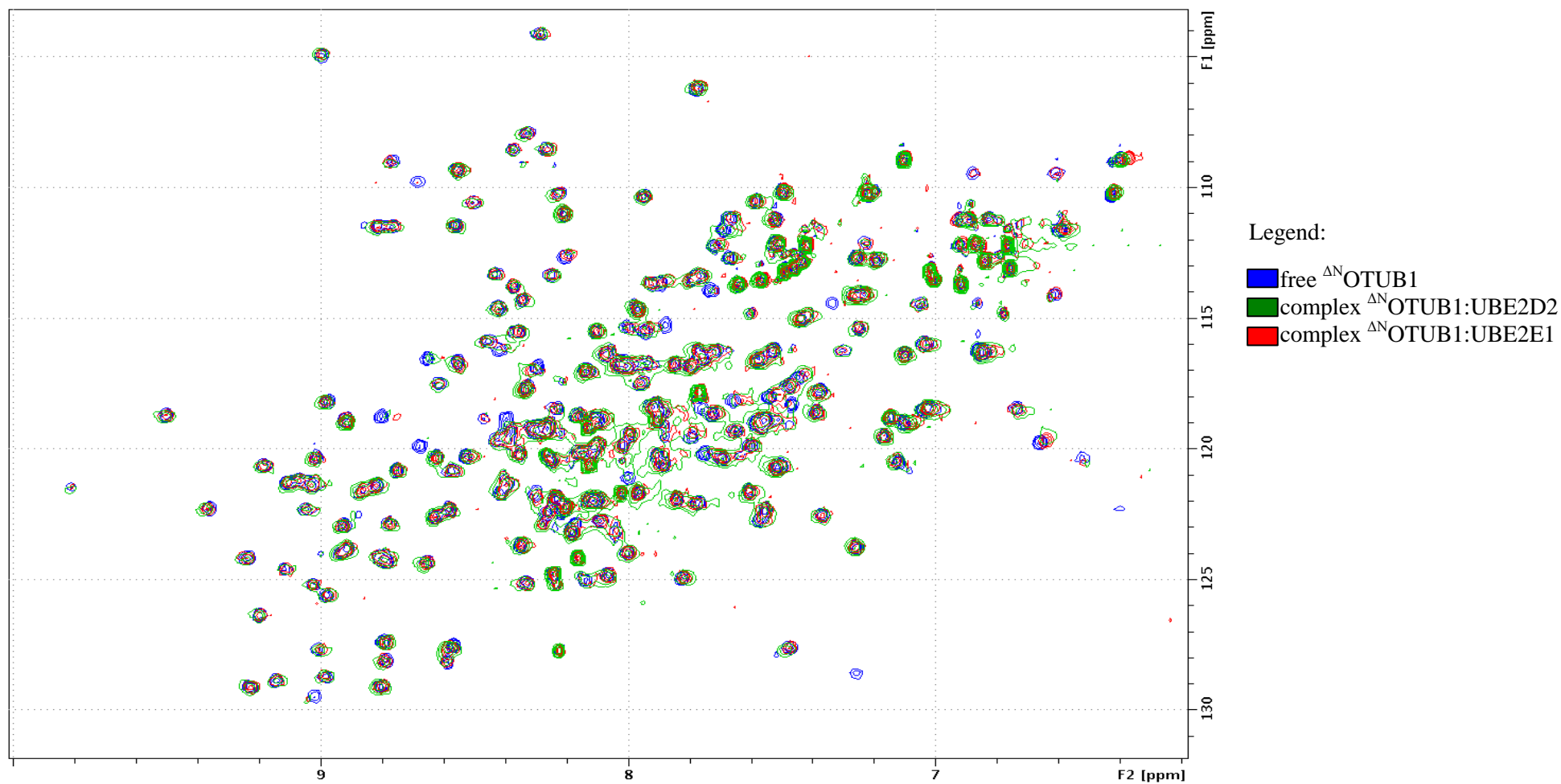
Figure 4.16   **Superposition of free ^{15}N-$^{\Delta N}$OTUB1, ^{15}N-$^{\Delta N}$OTUB1:UBE2D2 and ^{15}N-$^{\Delta N}$OTUB1:UBE2E1 2D NMR spectrum:** Chemical shifts observed in all three spectrum suggested that a similar binding interface may be used by $^{\Delta N}$OTUB1 to bind with UBE2D2 and UBE2E1.

4.5.4.3 Prediction of amino acids involved in $^{\Delta N}$OTUB1:UBE2D2 interaction

To further investigate the residues in UBE2D2 that interact with $^{\Delta N}$OTUB1, we wished to compare the spectrum of our ^{15}N-UBE2D2:$^{\Delta N}$OTUB1 complex with the spectrum of the qualitative UBE2D2 profile that is available in the PDB (Houben *et al.*, 2004) (*BRMB accession number: 6277). However, this UBE2D2 assignment data was collected at a difference temperature, pH and salt concentration (26.85°C, pH 7.0, 150 mM KCl) compared to our experiments (25°C, pH 7.4, 50 mM Tris). Therefore, all the signals may not appear at the exactly same location. However, by overlaying the spectrum of ^{15}N-UBE2D2:$^{\Delta N}$OTUB1 with the known UBE2D2 NMR assignment, chemical shifts caused by binding interaction can be predicted. This is illustrated in Figure 4.17, in which the residues of ^{15}N-UBE2D2 making contact with $^{\Delta N}$OTUB1 are represented by red peaks while green peaks correspond to the known UBE2D2 assignment. All red peaks that show distinct chemical shifts compared to the respective green peaks were considered to be the predicted points of contact. The prediction of UBE2D2 residues that are key in the binding interaction with $^{\Delta N}$OTUB1 allowed us to perform targeted mutagenesis studies (Chapter 5) to test the requirement for particular predicted contact points within OTUB1:E2 complexes. However, related residues in $^{\Delta N}$OTUB1 could not be predicted due to the absence of data for OTUB1 in the PDB.

Predicted ^{15}N-UBE2D2:$^{\Delta N}$OTUB1 binding points were mapped onto the UBE2D2 3D structure and rendered using pyMol v1.3 to define the binding surface. As can be seen in Figure 4.18, the data suggests that the 1st α-helix, L1 loop of 3rd and 4th β-sheet and the area around L2 loop connecting the 4th β-strand and H2 α-helix are potentially the binding interface for UBE2D2. When compared to the structural model of UBE2D2:CNOT4 complex (Dominguez *et al.*, 2004), this suggests that $^{\Delta N}$OTUB1 and CNOT4 (an E3 RING ligase) may utilise some of the same residues to interact with UBE2D2 (Table 4.3). At this stage of the project, the OTUB1:UBE2D2 structure was still unknown, therefore a hypothesis was made suggesting that OTUB1 and E3-RINGs may compete with each other to bind to UBE2D2. However, there are some differences in UBE2D2 binding residues

especially in the L2 loop connecting 4th β-strand and H2 α-helix, which could also suggest that DUB and E3 may bind to similar residues but in a slightly different manner, which raises the possibility that in some cases it may be possible to form DUB:E3:UBE2D2 complexes.

Figure 4.17 **Superposition of 15N-UBE2D2:ΔNOTUB1 and known UBE2D2 profile:** UBE2D2 surfaces that bind with ΔNOTUB1 were predicted by overlaying the spectrum of 15N-UBE2D2:ΔNOTUB1 (red) and known UBE2D2 NMR analysis (green).

N- |MALKRIHKELNDLAR|DPPAQ|CSAGPVGD|DM|FHWQATIM|GPNDSPYQGG|VFFLTIHF|PTDYPFKP|PKVA|FTTRIY

HP*N*INSNGSI*C*LDILRS|QWSPALTI|SKVLLSICSLLC|DPNPDDPLV|PEIARIYK|TD|REKYNRIAREWTQK|YAM -C

N-

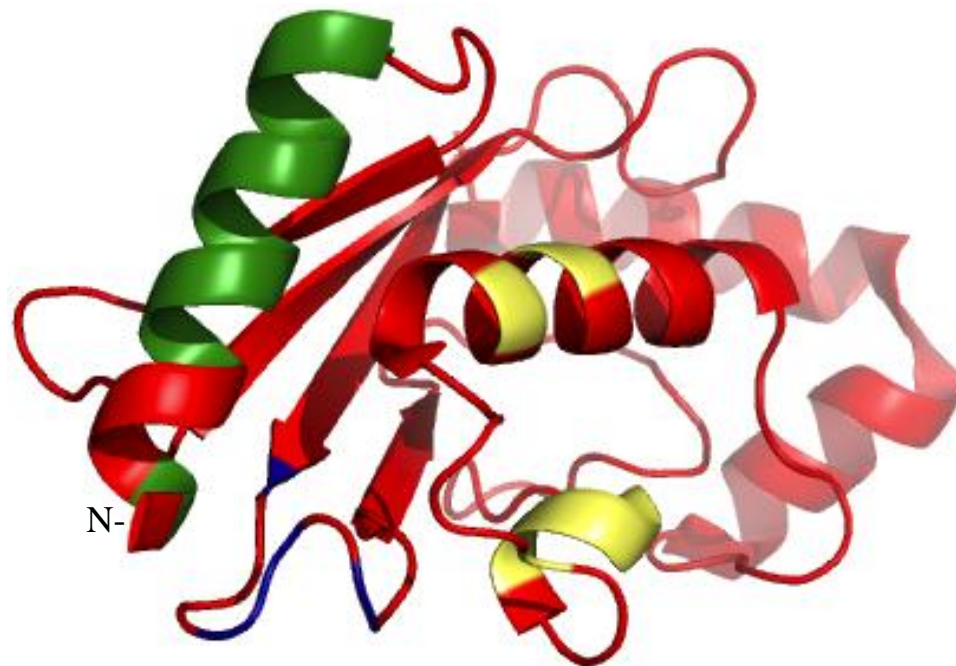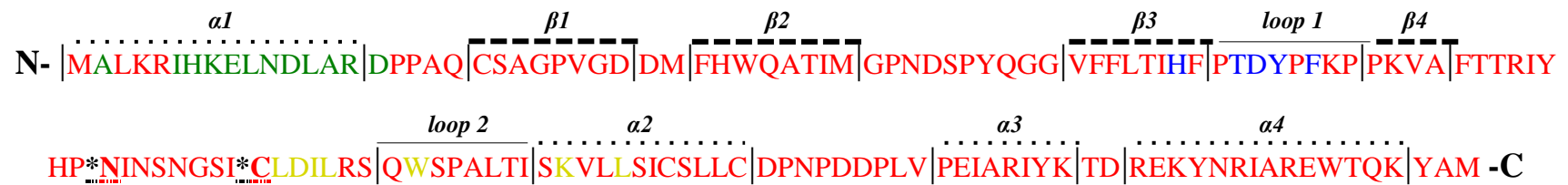**Figure 4.18   Predicted   UBE2D2 binding interface with** [AN]**OTUB1:** Structure of UBE2D2 and the predicted binding interface with [AN]OTUB1 were depicted by the area labelled green (1st α-helix), blue (L1 loop of 3rd and 4th β-sheet) and yellow (around L2 loop connecting 4th β-strand and H2 α-helix). The active site Asn77 and Cys85 are indicated by the asterisks (*). Picture generated using MacPyMOL v1.3.

Table 4.3   Comparison of ^{ΔN}OTUB1 and CNOT4 binding sites on UBE2D2: Predicted UBE2D2 amino acids involved in ^{ΔN}OTUB1 binding in comparison with CNOT4 binding site from Dominguez *et al.* (2004).

1st α-helix		L1 loop of 3rd and 4th β-sheet		L2 loop of 4th β-strand and H2 α-helix	
^{ΔN}OTUB1	CNOT4	^{ΔN}OTUB1	CNOT4	^{ΔN}OTUB1	CNOT4
Ala2	Ala2	His55		Leu86	
	Leu3	Thr58	Thr58	Asp87	
Arg5	Arg5	Asp59	Asp59	Ile88	
Ile6	Ile6	Tyr60	Tyr60	Leu89	
His7		Phe62		Trp93	Trp93
Lys8			Lys63		Ser94
Glu9	Glu9				Ala96
Leu10	Leu10				Thr98
	Asn11				Ile99
Asp12	Asp12			Lys101	Lys101
Leu13				Leu104	Leu104
Ala14					
Arg15					
Asp16					

4.5.5   Co-crystallisation trial

To verify that the interactions observed in solution-state NMR are also observed in solid-state, a crystallisation trial was performed as a complementary approach to NMR experiments. Moreover, it is widely known that proteins greater than 30 kDa can be problematic in NMR structural studies. A protein of 30 kDa like OTUB1 will have around 270 cross-peaks, which are difficult to identify via overcrowded 2D spectra. Although acquiring the spectra in a higher field or using a 3D, selective or segmental labelling strategy can alleviate these problems, the cost and technical difficulties made this an unattractive option. In addition, the most significant problem of large protein NMR studies is that of tumbling, although this can be surmounted Transverse Relaxation Optimised Spectroscopy

(TROSY) (Fernández and Wider, 2003). In contrast, crystallisation trials are not limited so much by size but by the search for exact condition for crystal growth. Unfortunately, in protein crystallisation, there are no empirical rules for finding the exact chemical condition that can crystallise a protein (Hosfield *et al.*, 2003). The highly effective approach to identify optimal crystallisation conditions is to systematically explore chemicals or conditions that have been shown to deliver crystals. A strategy known as 'sparse matrix' randomly maps a wide sector of the crystallisation conditions using many different precipitants, salts and pH. Conditions are based on successful crystallisation experiments found in databases, the most famous being the PCCD (**P**rotein-**C**omplex **C**rystallization **D**atabase), established by Radaev *et al.* (2006). Sparse matrix strategies were chosen in this experiment because they usually provide an excellent starting point to begin the search for a successful crystallisation condition (Jancarik and Kim, 1991).

In this experiment, co-crystallisation screens of the ~43.7 kDa complex of $^{\Delta N}$OTUB1:UBE2D2 were performed using four commercially available sparse matrix initial condition kits, each having 96 unique mixtures of precipitants, salts and pH. Complexes were prepared at 10 mg/ml and 20 mg/ml concentrations and they were incubated at RT and 4°C, which means that 1 536 initial conditions were screened for the $^{\Delta N}$OTUB1:UBE2D2 complex. Crystal growth was observed under several conditions, however, most examples were probably due to the formation of UBE2D2 crystals as the individual UBE2D2 (positive control) also generated crystals in these particular conditions (Fig. 4.19A). These UBE2D2 crystals were not brought forward for X-ray diffraction analysis because the structure had already been solved. Meanwhile, two positive hits that did not show any crystal growth in positive UBE2D2 controls were also observed, suggesting that these may be complex specific crystals. Unfortunately, these crystals did not achieve a diffraction-quality thought to be suitable for X-ray diffraction analysis.

The first hit of 20 mg/ml $^{\Delta N}$OTUB1:UBE2D2 complex was observed as 2D needles (whiskers-like) growing from a single nucleation center in initial condition 15% (w/v) PEG 20K,100 mM HEPES and pH 7.0 (Fig. 4.19B). Nucleation is the initial process that occurs in the formation of a crystal where a few molecules join together and become solute molecules (like an aggregate) but their combined attraction forces are stronger than other forces in the solution which usually can disrupt the formation of these aggregates. They are also called 'protocrystals' or 'pre-crystals' that become nucleation sites upon which crystals will begin to grow. To encourage the crystal formation from the conditions giving this nucleation, an optimisation of the conditions was performed where 96 conditions were prepared with a range of 13%-24% (w/v) PEG 20K, 100 mM HEPES and a pH range from 6.5-7.8, arranged in a 96-well sitting drop-vapour diffusion crystallisation plate.

The second hit was also observed in the 20 mg/ml complex conditions, in 20% (w/v) PEG 8K,100 mM HEPES pH 7.5. These crystals appeared more promising than the first because they were observed in thin-plates shape and could be grow more in third dimension form (Fig. 4.19C). In this case, there may be a need for condition optimisation to get good diffraction crystals. The conditions were optimised with a range of new condition consisting PEG 8K 15%-25% (w/v), 100 mM HEPES and pH 7.2-8.4. Unfortunately, at this stage we started to experience problems with $^{\Delta N}$OTUB1 homogeneity, which appeared to inhibit crystal formation in subsequent crystallisation trials.

(A)

(B)

(C)

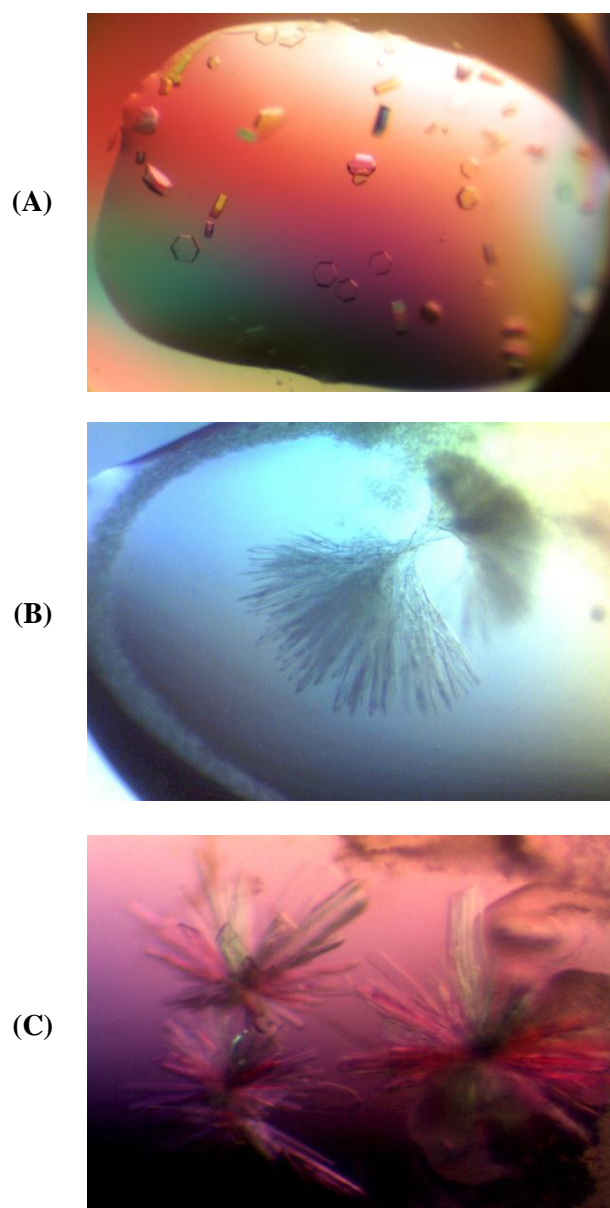**Figure 4.19   Crystal growth from initial trials: (A)** Example of a very good crystal growth of UBE2D2 (10 mg/ml) in 0.1 M Tris pH 8.0. **(B)** 2D needles obtained from $^{\Delta N}$OTUB1:UBE2D2 complex (20 mg/ml) in the condition 15% (w/v) PEG 20K, 100 mM HEPES pH 7.0 and **(C)** 3D needles obtained from $^{\Delta N}$OTUB1:UBE2D2 complex (20 mg/ml) in 20% (w/v) PEG 8K, 100 mM HEPES pH 7.5.

4.6   Progress limitation

Although the initial objectives of this chapter to express the recombinant proteins and validate the Y2H data were achieved, only limited insight could be provided into the structural information relating to the molecular organisation of OTUB1 and E2 protein complexes, which is important for understanding how DUB and E2 proteins aid the specificity and architecture of substrate ubiquitination on a systematic level. The main hindrance in this project was the failure to reproducibly produce pure $^{\Delta N}$OTUB1 protein. During the later stages of this project, purification of $^{\Delta N}$OTUB1 was routinely producing low yields of protein which appeared to exist in more than one species when analysed in 2D NMR spectra. However, no noticeable differences could be detected on SDS-PAGE gels, or by ion exchange chromatography, suggesting that both forms have the same molecular weight and pI. Much effort was invested to obtain the perfect form, however, this issue remained unresolved through to the end of the project.

Usually the reason for low protein expression relates to vector selection or host growth conditions. However, the vector and host cells were carefully chosen and validated in the initial stages of this study. Therefore, only the expression and purification protocol were re-evaluated. To begin with, the protease inhibitor PMSF was replaced by leupeptin, pepstatin and a protease inhibitor cocktail in order to see if there were any noticeable increases in the yield of purified proteins. When purifying an expressed protein, the first step is to rupture the cells resulting in the release of the total proteins, some of which might be proteases. These proteases may potentially digest proteins, including the expressed one, resulting in lower yield. PMSF as a serine protease inhibitor was chosen to inhibit targeted protein digestion considering that $^{\Delta N}$OTUB1 itself is a cysteine protease hence the use of protease inhibitor cocktail might inactivate $^{\Delta N}$OTUB1. However, other proteases may also act to reduce protein yield hence purifications were also performed with leupeptin (serine and thiol protease inhibitor), pepstatin (aspartic protease inhibitor) and protease inhibitor cocktail. Unfortunately, this had not effect on the $^{\Delta N}$OTUB1 2D NMR spectrum.

Secondly, the protein purification was performed at both RT and 4°C (constantly on ice). In our previous procedure, some of the steps for example during the incubation with TEV protease (to remove the His-tag) were carried out overnight at RT. In both temperatures, the spectra still consisted of two forms of $^{\Delta N}$OTUB1, which means that the aberrant form does not come about because of high temperature. 2D NMR screens were also performed on freshly purified protein and after being flash frozen in liquid nitrogen to confirm the formation of a second species was not simply caused by freezing.

Other unsuccessful attempts to recover $^{\Delta N}$OTUB1 into a single 'good' form include modifying growth conditions. Proteins were expressed at higher ODs to ensure sufficient cells to express target protein and by shortening the induction time to avoid expression of toxic proteins. Also, addition of glucose was tested as a way of suppressing leaky expression. Unfortunately, none of the modifications in growth conditions or purification steps solved the problem. Moreover, the problem was only shown by the $^{\Delta N}$OTUB1 protein as all E2s proteins continued to behave well throughout this project. For this reason, the vector as well as expression were also re-evaluated.

To rule out the possibility of any contamination of the $^{\Delta N}$OTUB1-pETM11 stock, the initial glycerol stock was revived and clones were isolated by maxiprep and sent for sequencing. They were then transformed into four different expression hosts namely: BL21 Star, B834 pLysS, Arctic express and Rosetta cells to re-evaluate the expression host (Figure 4.20 shows the SDS PAGE evaluation of these four hosts). Sadly, none of the hosts successfully restored the initial good spectrum of $^{\Delta N}$OTUB1.
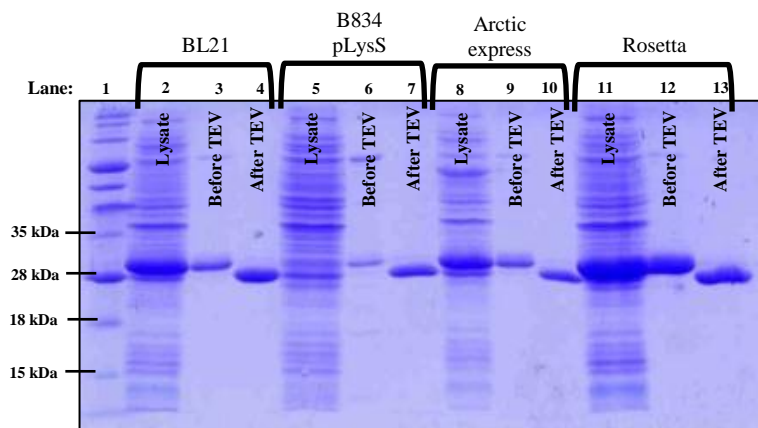
Figure 4.20   Coomassie blue 1D SDS-PAGE of newly purified $^{\Delta N}$OTUB1: Purification $^{\Delta N}$OTUB1 grown in four different expression host cells. Lanes 2,5,8,11 show the lysates after French press. Lanes 3,6,9,12 were the $^{\Delta N}$OTUB1 after Ni^{2+} purification before incubation with TEV protease and lanes 4,7,10,13 show the $^{\Delta N}$OTUB1 after TEV treatment and a second purification with ion exchange chromatography.

Re-cloning the $^{\Delta N}$OTUB1 ORF into pET151/D TOPO® His-tag vector (TOPO® Cloning were described comprehensively in Chapter 5) as an alternative to pETM-11 vector also does not give any fruitful outcome. The expression and purification of $^{\Delta N}$OTUB1 consistently gave an extra signals on 2D NMR spectrum and the population ranged from being in the minority to the majority. Figure 4.21A shows the best spectrum that was achieved where the $^{\Delta N}$OTUB1 spectrum is predominantly seen in the good form while Figure 4.21B shows a $^{\Delta N}$OTUB1 preparation dominated by the 'bad' form of the protein.

At this point, further expression studies were halted to focus on the generation and testing of NMR informed mutagenesis studies (Chapter 5).
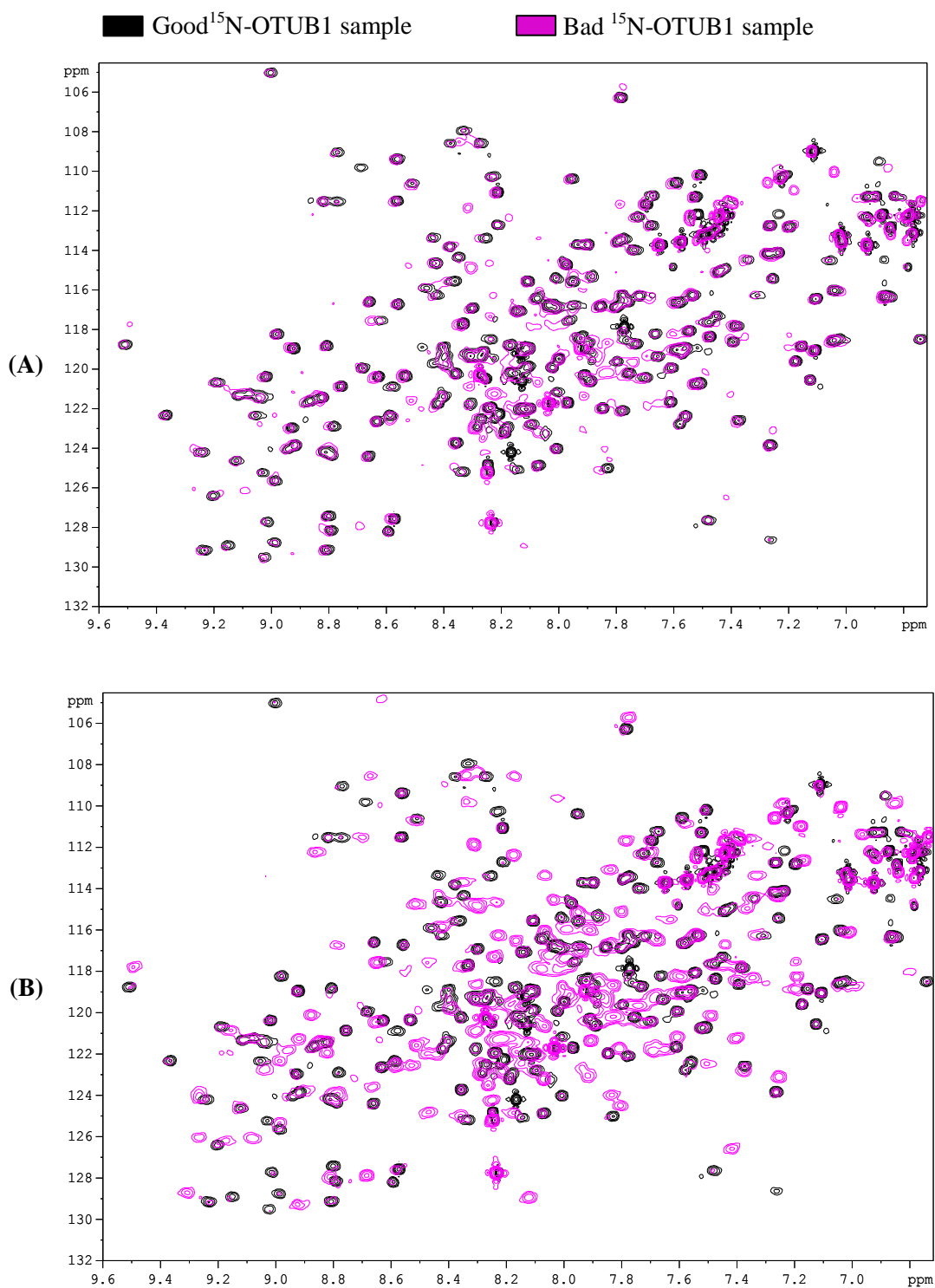
Figure 4.21 **2D NMR spectrum of double species ^{ΔN}OTUB1: (A)** and **(B)** Both are the
(^{1}H^{15}N)-HSQC spectra of ^{15}N-^{ΔN}OTUB1. Black signals were from the good sample and pink
were from suspected double species spectrum recorded at 25°C using a Bruker 800 MHz
Avance spectrometer. Purification and expression of ^{ΔN}OTUB1 continuously produced a
second form in a variety of percentages, ranging from **(A)** the minority to **(B)** the majority.

Chapter Five:
TARGETED ANALYSIS OF E2:OTUB1 COMPLEXES

5.1   Introduction

In the previous chapter, our initial goal of getting detailed structural information about OTUB1:E2 complexes could not be achieved within the time scale of this project due to repeated problems in getting pure $^{\Delta N}$OTUB1 expressed in a good conformation. Nevertheless, some useful information relating to the molecular requirements for OTUB1:E2 complex formation were obtained from NMR studies. Therefore, a decision was made to investigate the role that these predicted contact points may play in mediating OTUB1:E2 or E2:E3-RING binary interactions. When studying the molecular requirements for specific protein-protein interactions, generation of targeted binding site mutants can be a powerful method of identifying residues required for interaction or partner specificity. Indeed, similar mutagenesis approaches have been used successfully to study interactions between OTUB1 and other interaction partners (Edelmann *et al.*, 2009; Juang *et al.*, 2012).

In this chapter, the targeted E2 proteins; UBE2D2, UBE2E1 and UBE2N; were studied with either full-length OTUB1 or $^{\Delta N}$OTUB1. In addition, a panel of point mutations were introduced into UBE2D2 and the effects on interaction patterns with OTUB1, $^{\Delta N}$OTUB1 or a broad collection of human E3-RING proteins were analysed by Y2H and *in vivo* immunofluorescence techniques in order to develop a better understanding of the molecular basis of OTUB1:E2 or E2:E3-RING interactions. Y2H analysis of mutant UBE2D2s and fluorescent co-localisation tests were done in collaboration with Amy Ponsford, who was a final year undergraduate student working in the Sanderson lab partially under my supervision.

The aims of this part of the project were:

1) To generate a series of binding site-impaired UBE2D2 mutants based on predictions made from our previous NMR studies.

2) To assess the relative effects of targeted mutations on the binding of OTUB1 and E3-RING proteins.

3) To generate a model which predicts the mechanism of action of human E2:OTUB1 complexes.

5.2   Generation of E2-binding site mutants

5.2.1   Potential amino acids to be mutated

In the previous chapter, a series of key residues in UBE2D2 were identified as being potentially important for interaction with OTUB1. These residues were identified by overlaying spectrum of the ^{15}N-UBE2D2:$^{\Delta N}$OTUB1 complex and an available qualitative UBE2D2 NMR spectrum (Houben *et al.*, 2004). However, predicted binding points may not be exact due to the parameter differences used to record both spectra. Also, observed shifts may be the consequence of binding rather than actual physical points of contact.

Targeted mutagenesis was performed to generate forms of UBE2D2 with amino acid point mutations in order to determine if amino acid substitution is sufficient to disrupt interaction with OTUB1, $^{\Delta N}$OTUB1 or E3-RING proteins. If this is the case then it would implicate these particular amino acids as being necessary for binding to OTUB1, thereby validating our UBE2D2 binding interface predictions. Therefore, Lys8, Glu9 and Asp12 within the 1st α-helix and Lys101 at the L2 loop of 4th β-strand and H2 α-helix were chosen to be mutated as they are all polar, and polar amino acid side chains tend to gather on the outside of the protein where they have the potential to interact with the aqueous environment. Meanwhile, the non-polar amino acid side chains are buried on the inside to form a tightly packed hydrophobic core of atoms that are hidden from the aqueous environment (Alberts *et al.*, 2002). For this reason, we avoided mutating non-polar amino acids in order to avoid imposing unwanted conformational changes. Also, polar amino acids are more likely to be accessible on partner binding surfaces.

Acidic side chain amino acids Glu9 and Asp12 were mutated to alanine or arginine respectively. As the simplest and smallest amino acid, mutation to alanine could ensure that the mutations will not disrupt the secondary structure due to stearic hindrance (Alberts *et al.*, 2002). It was also thought to be a good idea to mutate them into basic side chain amino acids such as lysine, arginine or histidine in order to swap the acidity/basicity in order to see if the

binding interaction could be abolished rather than substituting them into another acidic amino acids, which may also mediate interactions. Therefore, Glu9 was substituted to Ala ([E9A]UBE2D2) and Asp12 was changed to Arg ([D12R]UBE2D2). Meanwhile, both Lys8 and Lys101 were mutated to alanine ([K8A]UBE2D2 and [K101A]UBE2D2). All these residues were mutated following QuikChange® site-directed mutagenesis protocol.

5.2.2   Mutagenesis strategy

QuikChange® mutagenesis was carried out as described in Chapter 2 and point mutations were made in such a way that encoded amino acids were converted to either alanine or arginine. Two complementary forward and reverse primers were designed carrying base substitutions in the centre of the primers as tabulated below:

Table 5.1   Forward and reverse UBE2D2 mutagenesis primers: The mutation conferred by these primers is described in the "Mutation" column. Bases that are mutated compared to the wild-type sequence are highlighted in maroon and pink.

	Mutation	Forward primer	Reverse primer
1	K8A	5' G AAG AGA ATC CAC **GCC** GAA TTG AAT GAT CTG 3'	5' CAG ATC ATT CAA TTC **GGC** GTG GAT TCT CTT C 3'
2	E9A	5' GA ATC CAC AAG **GCC** TTG AAT GAT CTG G 3'	5' C CAG ATC ATT CAA **GGC** CTT GTG GAT TC 3'
3	D12R	5' G GAA TTG AAT **CGC** CTG GCA CGG G 3'	5' C CCG TGC CAG **GCG** ATT CAA TTC C 3'
4	K101A	5' GCA CTA ACT ATT TCA **GCC** GTA CTC TTG TCC ATC 3'	5' GAT GGA CAA GAG TAC **GGC** TGA AAT AGT TAG TGC 3'
5	K8A,E9A	5' GA ATC CAC **GCC GCC** TTG AAT GAT CTG G 3'	5' C CAG ATC ATT CAA **GGC GGC** GTG GAT TC 3'
6	K8A,D12R	5' G AAG AGA ATC CAC **GCC** GAA TTG AAT **CGC** CTG 3'	5' CAG **GCG** ATT CAA TTC **GGC** GTG GAT TCT CTT C 3'
7	E9A,D12R	5' GA ATC CAC AAG **GCC** TTG AAT **CGC** CTG G 3'	5' C CAG **GCG** ATT CAA **GGC** GGC GTG GAT TC 3'
8	K8A,E9A, D12R	5' **C GCC** TTG AAT **CGC** CTG GCA CGG G 3'	5' C CCG TGC CAG **GCG** ATT CAA **GGC G** 3'

The mutation strategy was to perform PCR-amplification firstly using primers 1, 2, 3 and 4 on UBE2D2 in pDONR207 in order to generate a series of point UBE2D2 mutants: ^{K8A}UBE2D2, ^{E9A}UBE2D2, ^{D12R}UBE2D2 and ^{K101A}UBE2D2. PCR for double point mutations could then be performed in these single mutants using the appropriate primers. The PCR products were treated with *Dpn*I before transforming them into chemically competent cells for positive transformants selection. *Dpn*I is a restriction endonuclease that recognises a short DNA sequence (GATC), but only when the DNA is methylated. DNA replicated within bacteria, like the template used in the reactions, will be methylated and therefore recognised by *Dpn*I which will subsequently cleaved it. However, products synthesised *in vitro* will not be methylated and will not serve a substrate. They will be safely transformed into bacterial cells as an intact vector containing the UBE2D2 mutants ORF.

This experiment aimed to generate UBE2D2 as single, double, triple and quadruple mutants, however only 8 mutants were successfully generated. These included all the single mutants ^{K8A}UBE2D2, ^{E9A}UBE2D2, ^{D12R}UBE2D2 and ^{K101A}UBE2D2, and the double mutants ^{K8A,K101A}UBE2D2, ^{K8A,E9A}UBE2D2, ^{K8A,D12R}UBE2D2 and ^{E9A,D12R}UBE2D2. Generation of the mutants proved to be quite problematic and involved many failed attempts to produce positive transformants. This might be due to various factors for instance, an older mutagenesis kit was initially being used so there is a possibility that the enzymes may have denatured or not had their full activity. In addition, other solutions such as the reaction buffer or dNTPs may also not be as active due to numerous freeze-thaw cycles. Mutagenesis was therefore repeated using a new kit. The transformation was originally performed into XL 10-Gold Ultracompetent cells and then changed to XL 1-Blue Supercompetent cells in an attempt to increase the transformation efficiency. Optimisations were also made to the volume of *Dpn*I-treated DNA transformed into cells and also the transformation solution that was plated to increase the probability of bacterial colony growth. The generation of other double, triple and quadruple mutants was undertaken, however due to time constraints were not produced. Nevertheless, it would be interesting if all the remaining double mutants and

the triple and quadruple mutants could be produced in order to analyse further the importance of the binding sites.

5.3   Y2H evaluation of predicted UBE2D2 binding site

Y2H screening was performed to assess the relative effects of mutations on protein interactions between mutant forms of UBE2D2 and OTUB1, $^{\Delta N}$OTUB1 or E3-RING proteins. Yeast clones of UBE2D2 mutants were prepared as described in Chapter 3. Wild type and mutant forms of UBE2D2 protein expressed as Y2H bait fusions were then mated with a OTUB1 prey clone to observe colony growth. Mutant UBE2D2 clones which no longer show interactions with OTUB1 indicate amino acid residues which in some way facilitate OTUB1:UBE2D2 interaction. However, it should be noted that mutations that disrupt interaction between UBE2D2s and OTUB1s may exert effects by changing the conformation of UBE2D2 hence making it a non-functional protein. To ensure that mutations which disrupt UBE2D2:OTUB1 interaction do not induce unwanted conformational changes, mutant forms of UBE2D2 were also screened against a series of known interaction partners. Therefore, a complementary Y2H assay was also performed to screen UBE2D2 WT and mutants against a selection of E3-RING ligases. In a study performed previously in this laboratory, RNF5 and RNF185 were found to interact with WT UBE2D2, whereas RNF144B did not, hence these clones acted as positive and negative controls for the Y2H screening between UBE2D2s and OTUB1s.

5.3.1   Y2H screen results

5.3.1.1 UBE2D2 WT and mutants versus OTUB1 full-length and truncated

Analysis of UBE2D2 mutants demonstrated a variety of effects on OTUB1 binding in which, some mutants do not influence the interactions while others abolished the interaction. The E9A mutation both as a single and double mutation with K8A and D12R, abolishes interaction with OTUB1 on both -WLA and -WLH(3-AT) plates (Figure 5.1A). Meanwhile, the K101A mutation appears to cause a reduction in the binding affinity between UBE2D2 and OTUB1 as the yeast growth only appears on the less stringent -WLH(3-AT) selection plate. In contrast, K8A and D12R mutants do not abolish protein interaction when present as

single mutants, however when combined as double mutants binding was clearly impaired as no growth was observed on -WLA selection. Interestingly, the K8A+K101A mutation actually induced interaction with OTUB1 even though the K101A single mutation showed less growth on the -WLH(3-AT) plate and no growth on the -WLA plate. Growth of the double mutant also appeared to be weaker on the -WLA plate. The K8A and D12R mutants alone still allowed protein interaction; however when E9A is combined with either K8A or D12R the interaction is completely abolished. This experiment shows that K8A mutant interacts with $^{\Delta N}$OTUB1 even though the wild type does not show any growth suggesting a decrease in the binding affinity as discussed in Chapter 3.

Overall, it can be seen that Glu9 of UBE2D2 is required for interaction with OTUB1. This is supported by the lack of yeast growth by the double mutants K8A+E9A and E9A+D12R even though K8A and D12R mutations do not abolish interaction by themselves, suggesting that the E9A mutation has a dominant negative effect over other mutations. Data from this study also suggests that Lys101 may also be involved in forming the interaction surface, but may not be absolutely vital for OTUB1 binding as some degree of interaction is still observed on the less stringent plate. The lack of growth caused by the K8A+D12R double mutant suggests that mutation of just one residue is not enough to abolish the interaction but with both substituted binding is no longer feasible. These amino acid residues may therefore still form part of the interaction surface. The growth seen in the presence of the double K8A+K101A mutation suggests that the K8A mutation exerts a dominant effect over the K101A mutation, so even though the K101A mutation causes a reduction in binding affinity, the presence of the K8A mutation is enough to maintain binding.

The importance of the N-terminus of OTUB1 was also demonstrated in this experiment by the lack of interaction observed between $^{\Delta N}$OTUB1 and WT UBE2D2 or all UBE2D2 mutants except K8A. Loss of yeast growth suggests that the ubiquitin-binding domain contained at the N-terminus of OTUB1 increases the binding affinity of OTUB1 for

UBE2D2. Previous studies have suggested that E2 enzymes charged with ubiquitin produced higher affinity binding for OTUB1 so it is possible that when OTUB1 cannot bind to ubiquitin attached to the E2, the proteins cannot interact or have lower affinity. Meanwhile, there is the possibility that the K8A mutation is causing a structural change to UBE2D2 that could open up, for example, a new binding site, or reveal a different residue or conformation which allows a weaker interaction with OTUB1 even though it is missing the N-terminus.

5.3.1.2   *Analysis of interactions between WT and mutants forms of UBE2D2 with known E3-RING interaction partners*

A complementary Y2H assay was performed to screen WT UBE2D2 and UBE2D2 mutants against three E3 RING ligases; RNF5, RNF185 and RNF144B. In previous Y2H screens performed in the Sanderson laboratory, RNF5 and RNF185 were shown to interact with WT UBE2D2, while RNF144B did not interact with WT UBE2D2. Results presented in Figure 5.1B show that the UBE2D2 mutant E9A abolished interaction with OTUB1, however there is still interaction with RNF185 shown on both the -WLA and -WLH(3-AT) plates. The K101A mutant also abolished interaction with OTUB1 but still interacts with RNF185, again on both plates. The double mutation K8A+D12R causes a decrease in the binding affinity to OTUB1, however with RNF185, high affinity binding is detected on both plates. The E9A+D12R double mutant also shows interaction with RNF185 however they completely abolished interaction with OTUB1. The K8A+E9A double mutation however abolishes interaction with both RNF185 and OTUB1. Interestingly, the K8A+D12R double mutant, but not WT UBE2D2, interacts with RNF144B. Only WT UBE2D2 and the D12R mutant maintains binding to RNF5. All other single and double mutations abolished this interaction.

The results from this Y2H study suggest that the mutations are not causing gross conformational changes in UBE2D2 that could inactivate the protein. Therefore, it would appear that these mutations specifically block interaction with OTUB1. The K8A+E9A double mutant shows no growth on both -WLA and -WLH(3-AT), suggesting that this

double mutation may cause a structural change in UBE2D2 that make it non-functional as this mutation also inhibits OTUB1 binding. However it is also possible that these two amino acids are required for the binding of both OTUB1 and E3-RINGs. Surprisingly, the K8A+D12R double mutation interacts with RNF144B, suggesting that these two mutations together may reveal or create a new binding site or residue that increases affinity of UBE2D2 binding to E3-RINGs. The binding of some but not all E3-RING proteins to UBE2D2 was inhibited by some of the same UBE2D2 mutations that inhibited OTUB1 binding, suggesting that in many cases OTUB1 and E3-RING binding to E2 proteins may be a competitive process.

(A)

(B)

(C)

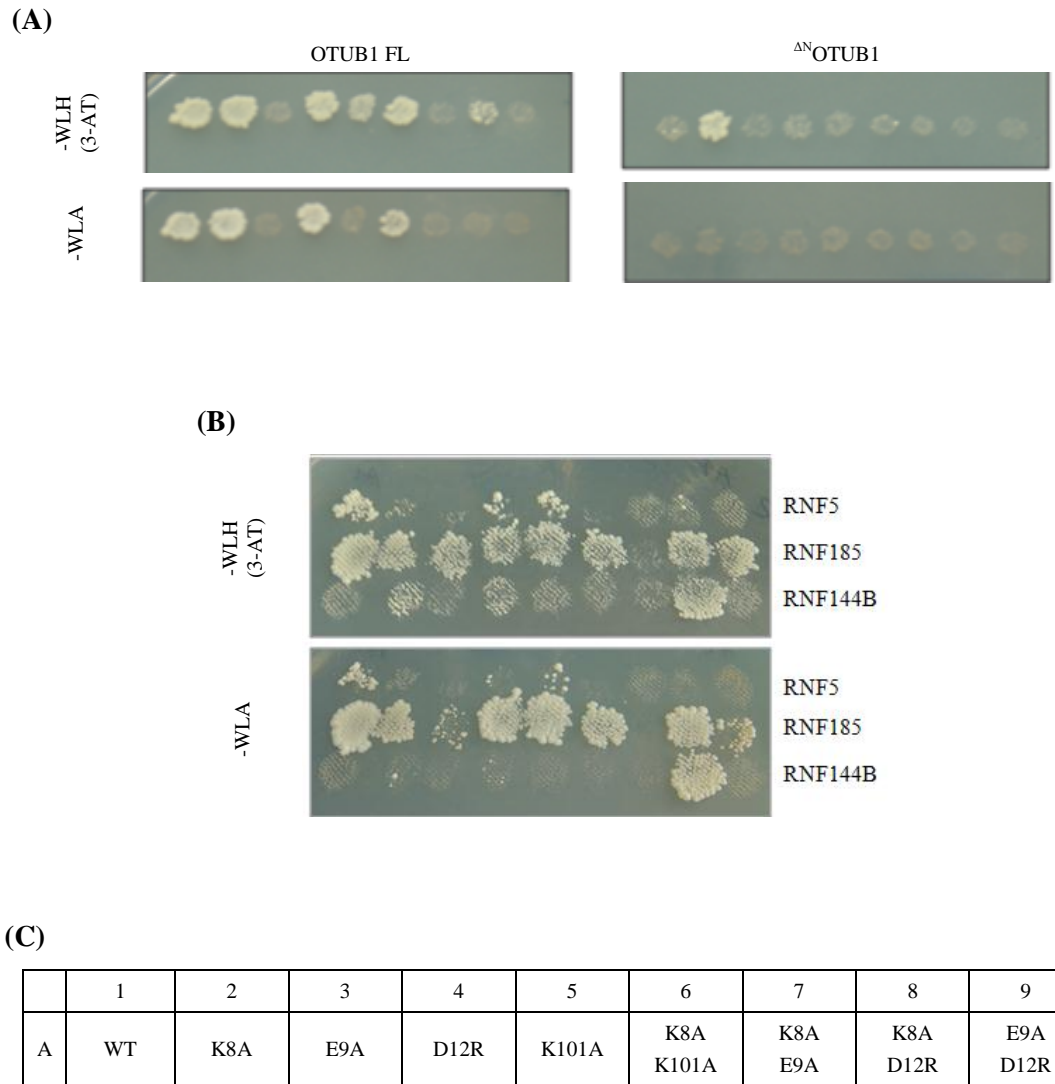|	1	2	3	4	5	6	7	8	9
A	WT	K8A	E9A	D12R	K101A	K8A K101A	K8A E9A	K8A D12R	E9A D12R

Figure 5.1 **Mutant UBE2D2 interactions in Y2H: (A)** OTUB1 FL and $^{\Delta N}$OTUB1 prey mated against mutant UBE2D2 baits on SD-WLA and SD-WLH(3-AT) at day seven. **(B)** E3 RING ligases prey mated against mutant UBE2D2 baits on SD-WLA and SD-WLH(3-AT) at day seven. **(C)** Grid positions of the UBE2D2 mutant baits assayed in Figure (A) and (B).

5.3.2   Competitive and non-competitive OTUB1:E3 RING:UBE2D2 binding models

Results in this experiment support both the competitive inhibition and non-competitive binding model presented in Figure 5.2. The data from Y2H assays show that some mutants do and others do not inhibit E3-RING binding. Figure 5.2A shows that when OTUB1 is bound to UBE2D2, this prevents E3-RING binding. Moreover, data from this experiment show that the interaction between RNF5 and UBE2D2 is inhibited by K8A, E9A and K101A mutations, and E9A+K101A also prevents OTUB1 binding. This suggests that RNF5 and OTUB1 use some of the same residues to bind to UBE2D2, and this supports a competitive mechanism for some E3-RING proteins. On the other hand, mutations that do not allow OTUB1 to bind but do not disrupt E3-RING RNF185 binding could suggest that RNF185 may bind in a slightly different way, which supports the non-competitive inhibition model as RNF185 interacts with all but one mutant analysed in this study, so could therefore be binding to UBE2D2 at a different site other than OTUB1.

However, it is also interesting to consider that mutations of UBE2D2 might change its specificity towards E3-RINGs. Based on their evolutionary comparison of the UBC fold in E2 proteins, it has been proposed to adopt 'three shells' of amino acids; the solvent accessible first shell which is directly involved in Ub/UBL conjugation, the second shell (or the inner core structures) partially covering the first shell and finally the third shell, the actual surface residues which mediate direct interactions with other partners like E1 and E3 proteins (Figure 5.3). If these properties are true then it should be possible to synthesise for example a UBE2N mutant that resembles UBE2D2 by mutating surface (third shell) residues. In addition, mutation of second shell residues also lead to changes in E3 interactions, as they support the correct organisation of third shell residues (van Wijk and Timmers, 2010).
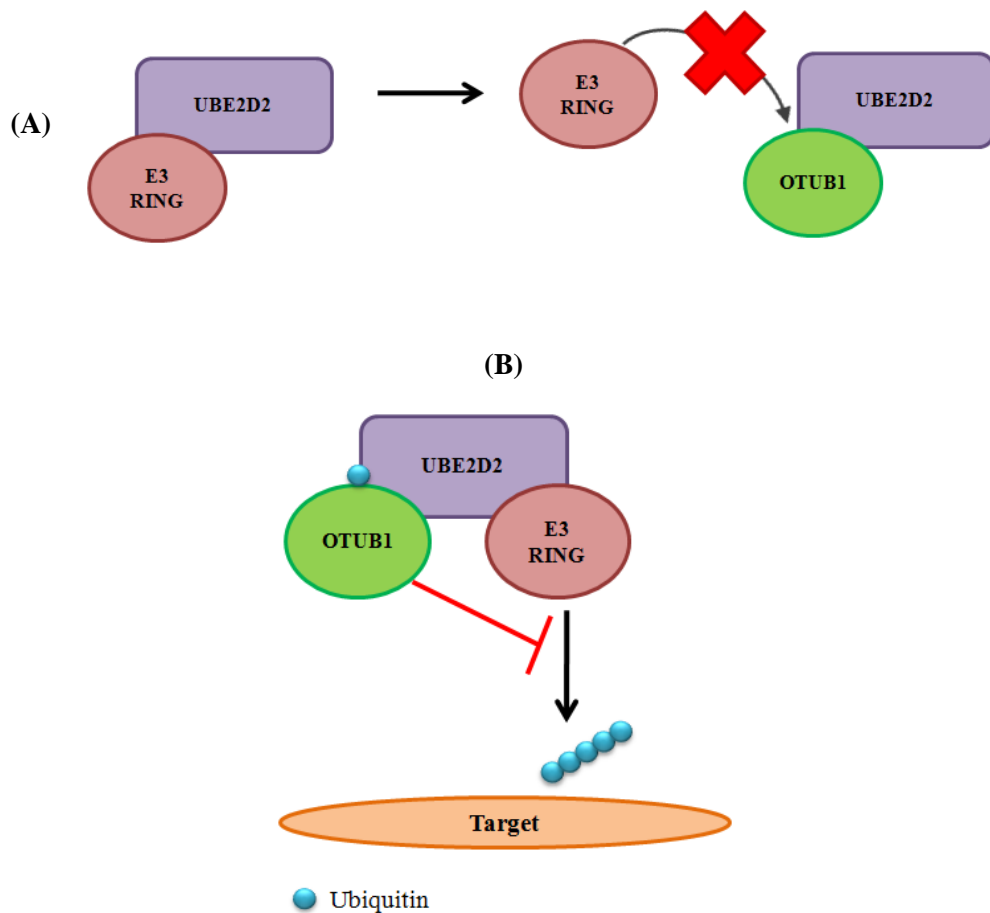
Figure 5.2 **Binding model of UBE2D2, OTUB1 and E3 RINGs: (A)** Competitive binding model. Y2H results support the competitive model that suggests OTUB1 and E3-RINGs compete for the same or similar binding site on UBE2D2. **(B)** Non-competitive binding model. This model suggests that E3 RINGs bind to residues that are not involved in OTUB1 binding. OTUB1 inhibits ubiquitination of targets through a non-competitive mechanism. Figure adapted from Nakada *et al.* (2010).
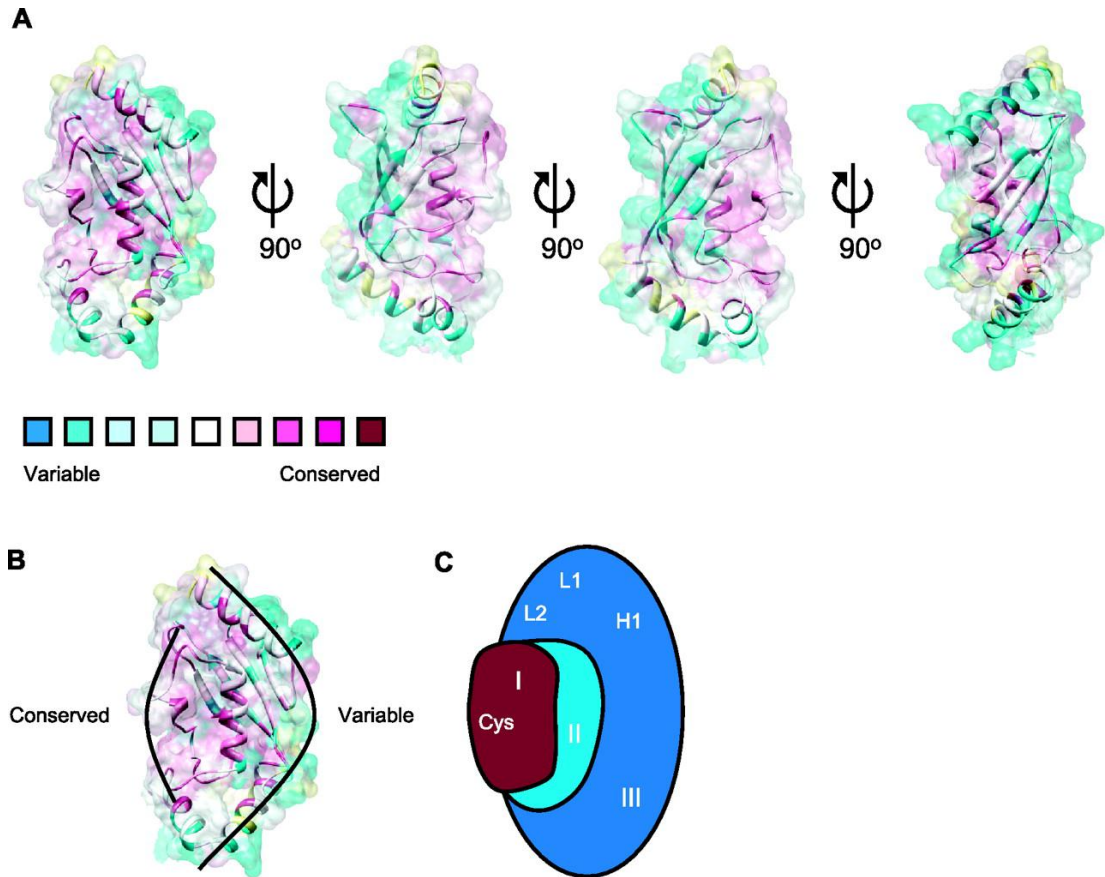
Figure 5.3   E2 shell-like model proposed to be important for selectivity of key enzymes (E1, E3, Ub/UBL) in directing Ub/UBL-conjugation pathways: (A) and (B) Evolutionary-diverged, mutational distances of amino acids in E2 enzymes concentrate at one side of protein according to the colour key. **(C)** Schematic representations of the shell-model for E2 enzymes. The first shell contains the catalytic cysteine, connected through the second shell with the third shell that contains structural elements that mediate selectivity towards Ub/UBL, E1s and E3s. Figure from van Wijk and Timmers (2010).

5.4   Expression and purification of UBE2D2 mutant proteins

The work in this part was done concurrently with experiments performed in Chapter 4 to sort out $^{\Delta N}$OTUB1 expression and purification problems. The UBE2D2 mutants which showed an interesting Y2H result were cloned into the pET151/D vector that fuses a His-tag to the protein and purified for NMR titration experiment with $^{\Delta N}$OTUB1, in case problems with production of $^{\Delta N}$OTUB1 could ever be resolved. Moreover, as an alternative to the pETM-11 vector, $^{\Delta N}$OTUB1 was also cloned into this vector to ensure that the problem of $^{\Delta N}$OTUB1 heterogeneity is not caused by pETM-11.

5.4.1   Directional TOPO® Cloning

The chosen method for cloning UBE2D2 mutants into suitable expression vector was to use the Champion TOPO® cloning kit from Invitrogen. The vector pET151/D contains the T7 bacteriophage promoter for a high inducible expression similar to pETM-11, and also has an N-terminal His-tag and TEV cleavage site (Figure 5.4A). TOPO® cloning does not require any DNA ligases to clone the targeted genes into specific vectors. This system merely relies on Taq polymerase to amplify the gene with primers containing TOPO® sites (CACC). The vital feature of this system is the pre-linearised vector containing topoisomerase at the 5' and 3' ends of the vector which functions as both restriction enzyme and ligase. Directional TOPO® cloning enables cloning of the blunt-ended PCR products in a 5´ to 3´ orientation directly into the vector in 5 minutes ligation reaction, thereby eliminating subcloning steps and saving a lot of time. The vectors contain a single-strand GTGG overhang on the 5´ end and a blunt end on the 3´ end. The four-nucleotide overhang invades the double-strand DNA of the PCR product and anneals to the CACC sequence at 5´ end and topoisomerase will then ligate the PCR product in the correct orientation, similar to the method described in Shuman (1994).
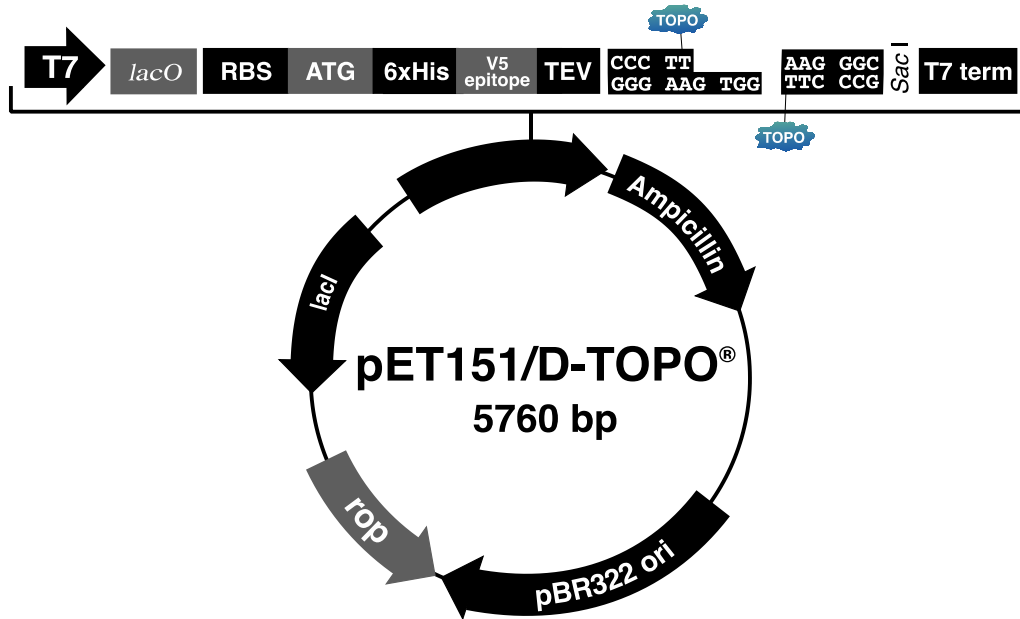
Figure 5.4   **Map of pET151/D-TOPO® vector:** Visible prior to the topoisomerase site are His-tag and TEV site, which function in producing the expressed protein with a cleavable N-terminal 6xHis-tag. Ampicillin resistance serves as a selectable marker for this vector.

5.4.1.1 Primer design and the TOPO® cloning principle

All forward primers have the CACC sequence at the 5' end of the primer; this was to ensure the topoisomerase reaction worked. The remainder of the primer was designed to provide a suitable annealing length complementary to the two strands of the target DNA, with a melting temperature (T_m) between 50°C-58°C and the T_m difference between the forward and reverse primers should be no more than 2°C to allow efficient PCR reaction.

Table 5.2   TOPO® forward and reverse primers: ORF Primer 5' to 3' (first base position indicated at 5' end) with CACC addition denoted by the green font.

	Forward primer	**Reverse primer**
OTUB1	5' **CAC C**AT GGC GGC GGA GGA AC 3'	5' CTA TTT GTA GAG GAT ATC GTA GTG TCC 3'
$^{\Delta N}$OTUB1	5' **CAC C**GA GAT TGC TGT GCA GAA CC 3'	5' CTA TTT GTA GAG GAT ATC GTA GTG TCC 3'
UBE2D2	5' **CAC C**AT GGC TCT GAA GAG AAT CCA C 3'	5' TTA CAT CGC ATA CTT CTG AGT CC 3'
UBE2E1	5' **CAC C**AT GTC GGA TGA CGA TTC G 3'	5' TTATGTAGCGTATCTCTTGGTCC 3'
UBE2N	5' **CAC C**AT GGC CGG GCT G 3'	5' TTA AAT ATT ATT CAT GGC ATA TAG CCT CGT CCA TG 3'

The PCR products were then incubated with the linear vector that already has the topoisomerase enzyme covalently attached to both of its strands' free 3' ends. When the free 5' ends of the PCR product strands attack the topoisomerase (or 3' end) of each vector strand, the strands are covalently linked by the already bound topoisomerase. This reaction proceeds efficiently when this solution is incubated at room temperature with required salt to prevent re-cutting by the topoisomerase. Reaction products were transformed into chemically competent cells and positive transformants were selected from ampicillin selection plate. Overall, TOPO® cloning proved to be a more convenient protocol compared to conventional cloning with the pETM-11 vector.

5.4.1.2 Purification of UBE2D2 mutant proteins

Following cloning of the UBE2D2 mutants into the His-tag pET151/D TOPO® vector, they were transformed into Rosetta expression cells for large-scale protein purification. Protein expression and purification were performed in the same manner as for wild type UBE2D2. Although some modifications in ion exchange chromatography were required due to the pI being either increased or decreased upon amino acids substitution, all mutants were purified in satisfying yield and were found to be correctly folded by analysis of 1D NMR spectra. However, no binding experiments could be done on them until the end of this project since the problem with $^{\Delta N}$OTUB1 could not be solved.

5.5   **Immunofluorescence study of OTUB1 and E2 conjugase proteins**

To determine the subcellular distribution of the enzymes, co-localisation experiments were performed by co-expressing the OTUB1 and the E2 conjugases in mammalian cells. For this study, OTUB1 and the selected E2 proteins were transferred from pDONR223 (or pDONR207) entry vectors into fluorescent-tagged mammalian expression constructs. Although in terms of validation, it has already been shown that OTUB1 interacts with these E2s, the location in which they interact has not yet been established. As individual enzyme, OTUB1 possesses a putative nuclear localisation signal and resides in the nucleus of porcine kidney cells (Balakirev *et al.*, 2003; Shan *et al.*, 2009). However, recent work revealed that OTUB1 regulates p53 in the cytoplasm (Sun *et al.*, 2012). On the other hand, analysis of the localisation of UBE2A/B, UBE2D1, UBE2D2, UBE2D3, UBE2E1, UBE2E2, UBE2N and UBE2R1 shows that the Ubc4/5 family proteins (UBE2D1, UBE2D2, and UBE2D3) specifically co-localised with c-Cbl in endosomes (Umebayashi *et al.*, 2008). UBE2D2 was also found to co-localised with Nedd4 in the cytoplasmic periphery (Anan *et al.*, 1998) and UBE2N has been shown to regulate p53 also in the cytoplasmic compartment (Laine *et al.*, 2006). While UBE2D2 and UBE2N seem to be cytoplasmic proteins, UBE2E1 co-immunoprecipitates and co-localises with Ataxin-1 protein in the nucleus (Hong *et al.*, 2008).

In this experiment, the main objective was to evaluate whether co-expression causes a change in the distribution that could suggest a particular localisation for interaction. OTUB1 and [ΔN]OTUB1 were transferred into pEGFP-N2 vector that fuses an N-terminal green fluorescent-tagged to the protein. It should be noted that OTUB1 in pDONR223 does not have any stop codon (Rual *et al.*, 2004), therefore only N-terminal tags will be suitable to OTUB1. Meanwhile, all the E2 enzymes were transferred into pG-cherry-mGR (kindly contributed by Sheila Ryan, University of Liverpool). ORF cloned in this vector will be expressed in mammalian cells as a tagged protein with C-terminal red fluorescent-tag. All clones were sequence-verified before transfecting into HeLa cells for immunofluorescence

studies. Vectors containing gene of interests were then singly transfected or co-transfected into HeLa cells to observe any overlap in their distribution, or changes when co-expressed to suggest the potential for interaction.

5.5.1 Localisation of OTUB1, $^{\Delta N}$OTUB1, UBE2D2, UBE2E1 and UBE2N

Immunofluorescence studies showed that OTUB1 localised to the same cellular compartments as the E2 enzymes, which shows that they have a potential for interaction *in vivo*. This experiment demonstrates that OTUB1, UBE2D2 and UBE2N are localised throughout the cell while UBE2E1 is predominantly nuclear.

In Figure 5.5, OTUB1 and $^{\Delta N}$OTUB1 prove to have a diffuse localisation throughout the cell. Meanwhile in Figure 5.6, UBE2E1 appears to have a predominantly nuclear localisation, consistent with literature reports. Also, it can be confirmed that similar to OTUB1 and $^{\Delta N}$OTUB1, UBE2D2 demonstrates a diffuse localisation throughout the cell, as does UBE2N, although it does appear to have a stronger localisation in the nucleus which would be consistent with its reported role in DNA repair (Nakada *et al.*, 2010).
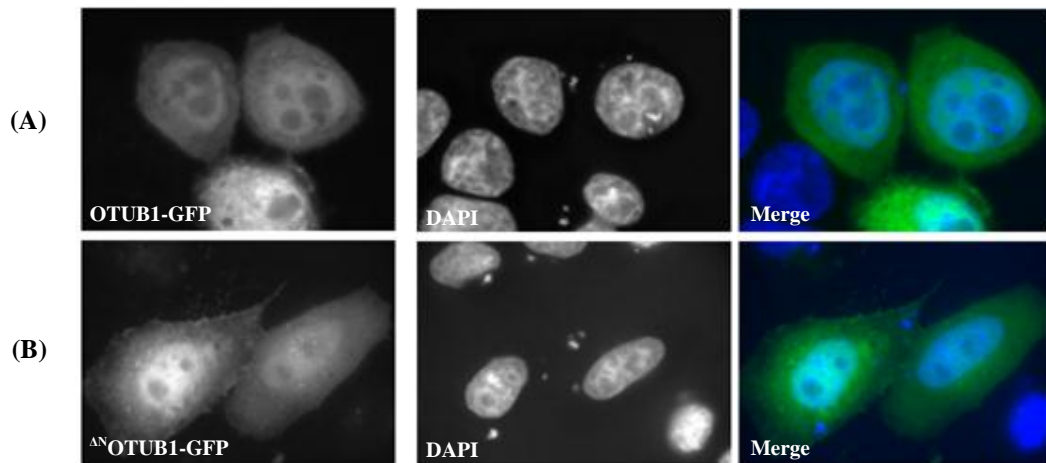
Figure 5.5 Localisation of (A) OTUB1 and (B) ^{ΔN}OTUB1 singly transfected into HeLa cells. Both transfections show a wide distribution throughout the cell. The DAPI blue stain shows cell nuclei while OTUB1 and ^{ΔN}OTUB1 are observed in green.
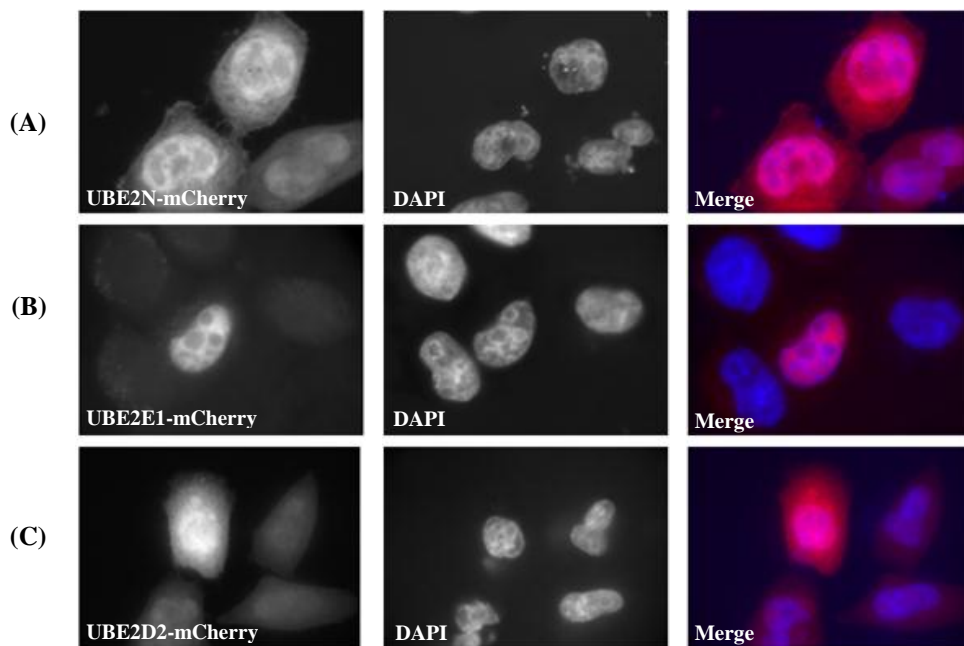

Figure 5.6   Localisation of (A) UBE2N, (B) UBE2E1 and (C) UBE2D2 in HeLa cells: UBE2N and UBE2D2 show a diffuse distribution throughout the cell while UBE2E1 shows a nuclear distribution. The DAPI blue stain shows cell nuclei while the E2s are observed in red.

5.5.2   Co-localisation study of OTUB1 and $^{\Delta N}$OTUB1, with E2 proteins

Co-localisation studies were performed to determine whether OTUB1 co-expressed with the

E2s could cause any recruitment of either protein to a particular subcellular localisation, thus

implying a physical interaction. As seen in Figure 5.7, when OTUB1 is co-expressed with

UBE2D2 and UBE2N, there does not appear to be any change in the subcellular distribution

of any of the proteins. However, when OTUB1 is co-expressed with UBE2E1, there does

appear to be some recruitment of OTUB1 to the nucleus compared to when OTUB1 is singly

transfected; however it is still localised throughout the cell. There is no change in the

subcellular distribution of UBE2E1 compared to its single transfections where it still shows

localisation predominantly to the nucleus and small punctuate structures in the cytoplasm.

In Figure 5.8, there do not appear to be any changes in the subcellular distribution of any of

the proteins when co-expressed in which, $^{\Delta N}$OTUB1 and UBE2D2 still show a wide

distribution as when singly transfected and UBE2N still appears to have a stronger

expression in the nucleus. In conclusion, co-expression of the E2 enzymes UBE2D2,

UBE2E1 and UB2N with both OTUB1 and $^{\Delta N}$OTUB1 did not alter the distribution of any of

the proteins; however they still have the potential to co-localise as they are expressed within

the same cellular compartments.
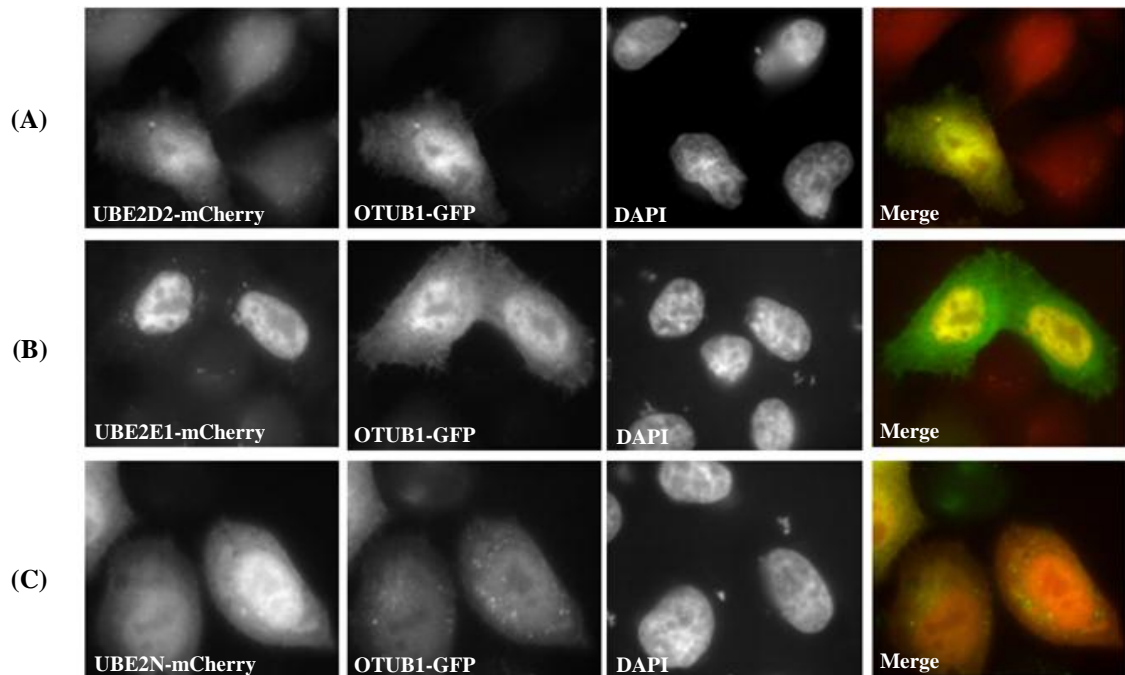
Figure 5.7   Co-localisation of (A) UBE2D2, (B) UBE2E1 and (C) UBE2N with OTUB1 in HeLa cells: When compared to single transfections the subcellular distribution of OTUB1 does not change when co-expressed with UBE2D2 and UBE2N, suggesting that one does not recruit the other to certain organelles when co-expressed. However OTUB1 seems to be slightly more nuclear when co-expressed with UBE2E1. The DAPI stain shows cell nuclei.

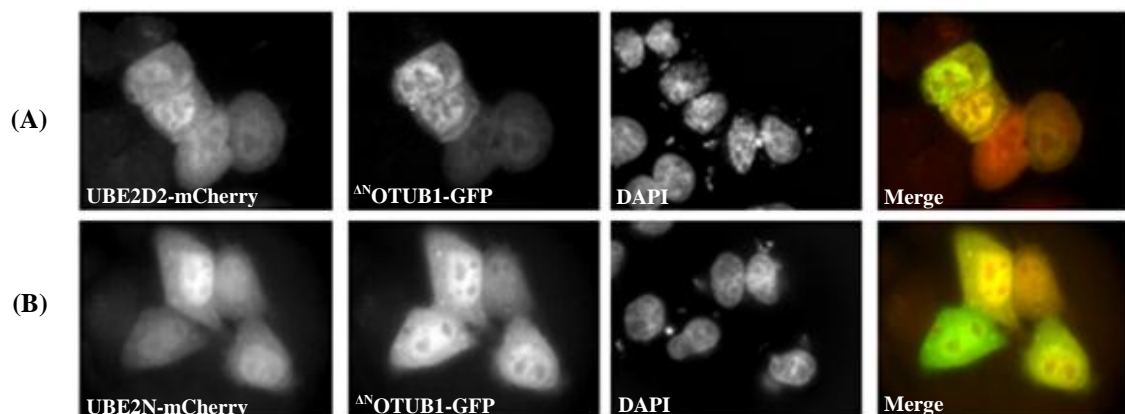**Figure 5.8   Co-localisation of (A) UBE2D2 and (B) UBE2N with ᐱNOTUB1 in HeLa cells:** There appears to be no change in subcellular distribution even though OTUB1 is missing its N-terminus. The DAPI stain shows cell nuclei.

Chapter Six:
DISCUSSION AND CONCLUSION

6.1 Solved structures of OTUB1:UBE2D2 and OTUB1:UBE2N complexes

Two comprehensive papers addressing the structural information of OTUB1:UBE2D2 and OTUB1:UBE2N complexes were published separately on February 10th and February 22nd 2012 (Juang *et al.*, 2012; Wiener *et al.*, 2012) while we were in the process of preparing this thesis. Juang *et al.* (2012) demonstrated that ubiquitin-charged UBE2D2 is recognised by OTUB1 through contacts with both donor ubiquitin and the E2 enzyme. By mimicking the Lys48-linked ubiquitin recognition, the free ubiquitin interacts with the N-terminal ubiquitin-binding site on OTUB1 to promote binding with the ubiquitin charged UBE2D2 protein (Juang *et al.*, 2012; Wiener *et al.*, 2012).

OTUB1 has been proven to have specificity towards K48-linked ubiquitin chains (Edelmann *et al.*, 2009). The interaction of the two ubiquitin molecules with OTUB1 puts the K48 of the donor ubiquitin in immediate proximity to the C-terminus of the distally bound ubiquitin, therefore this binding configuration might reflect how OTUB1 recognises its preferred substrate, K48-linked ubiquitin chains. In addition, the active site cysteine is well positioned to attack the K48-linked isopeptide bond. The interactions of OTUB1, UBE2D2 and two ubiquitin molecules (donor ubiquitin and free ubiquitin) are demonstrated in Figure 6.1. The specific surface residues involved in the interactions are defined in Table 6.1, which also revealed that the UBE2D2 surface obviously overlaps with E3 binding surface suggesting that OTUB1 and E3 would compete with each other to bind to E2 proteins (as summarised in Table 6.2).
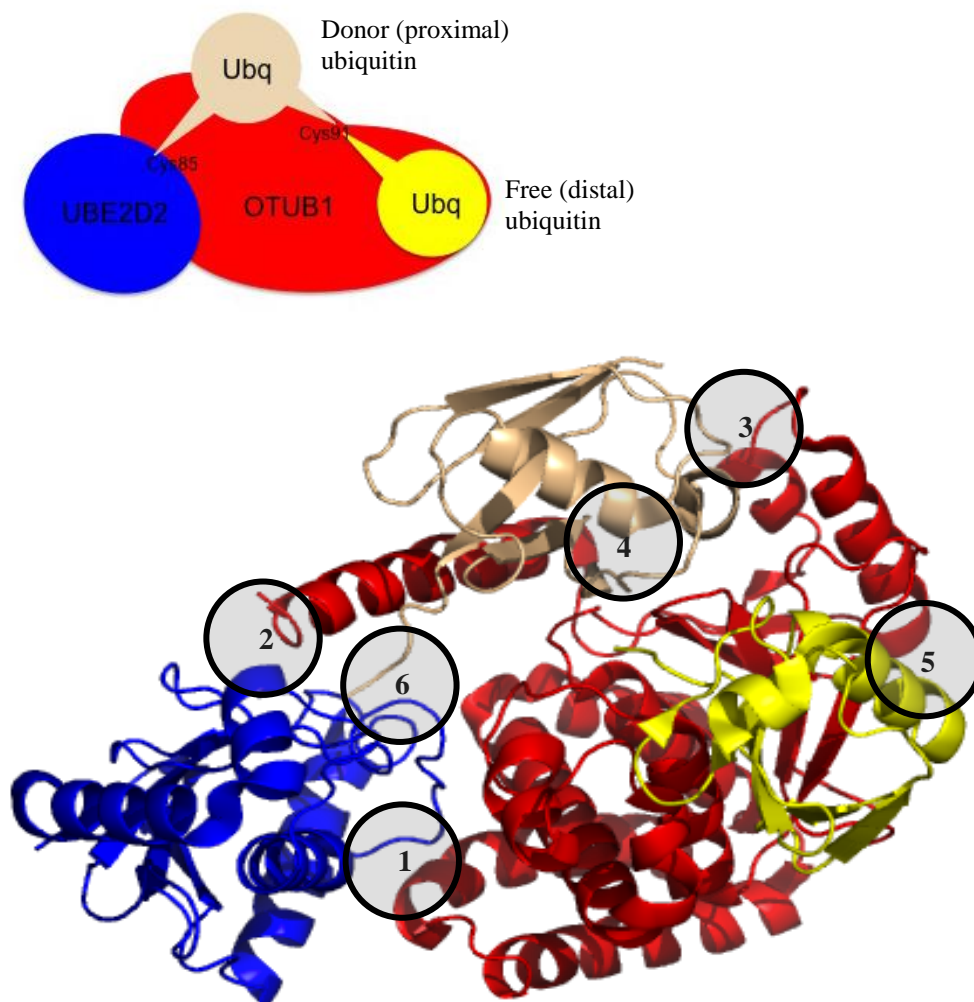
Figure 6.1   Schematic and ribbon representation of Ub~UBE2D2-OTUB1-Ub complex: (1) and **(2)** OTUB1:UBE2D2 binding surfaces consist of a larger surface in the center of OTUB1 (red) and UBE2D2 (blue) and a smaller contact at the catalytic cleft region of UBE2D2. **(3)** and **(4)** Two regions of OTUB1:donor Ub (brown) binding surfaces. **(5)** OTUB1:free Ub (yellow) binding interface. **(6)** UBE2D2:donor Ub binding surface involves little contact between C-terminus (Gly76) of the donor Ub and the catalytic cysteine of UBE2D2. Figure generated using MacPyMOL v1.3.

*Details on specific surface 1-5 residues are elaborated in Table 6.1. Meanwhile, surface 6 only involves Gly76 of ubiquitin and Cys85 of UBE2D2 hence no elaboration needed.

Table 6.1 **Specific residues that mediate interactions in Ub~UBE2D2-OTUB1-Ub complex: (A)** Binding interface between OTUB1 and UBE2D2. **(B)** OTUB1 and donor ubiquitin. **(C)** OTUB1 and distal (free) ubiquitin.

(A)

Surface 1		Surface 2	
OTUB1	UBE2D2	OTUB1	UBE2D2
helices α5, α7, and α9-α10 linker	helix α1, the α2-α3 linker, and the β3-β4 linker	helix αA	catalytic cleft region
Phe130	Met1	Asp27	Lys63
Thr131	Arg5	Glu28	Lys66
Phe133	Lys8	Met31	Ser83
Thr134	Asp12	Asp35	Arg90
Ile135	Pro61		Ser91
Asp137	Phe62		
Phe138	Lys63		
Asp169	Gln92		
Tyr170	Ser94		
Val173	Pro95		
Leu177	Ala96		
Gln206	Thr98		
Glu209	Lys101		
Pro210			
Met211			
Lys213			

(B)

Surface 3		Surface 4	
OTUB1	Donor Ubq	OTUB1	Donor Ubq
loop regions (α1-α2, b2-α3, β4-β5)	β3 to β5	N-terminal helix αA	β1, β3, β4, and intervening linkers
Tyr61	Lys48	Tyr26	Lys48
Ala62	Arg54	Ile30	Arg54
Asp65	Ser57	Gln33	Ser57
Tyr68	Asp58	Gln34	Asp58
Pro87	Tyr59	Arg36	Tyr59
Asp237	Asn60	Ile37	Asn60
Pro263	Gln62	Gln38	Gln62
His265		Ile41	

(C)

Surface 5			
OTUB1		Distal Ubq	
helix α10; strands β3, β3', and β4; and irregular loop region connecting helices α7 and α9		β1, β3, β4 and flanking loop regions	
Phe189	Gln225	Leu8	Gly47
His192	Tyr235	Thr9	His68
Phe193	Asp237	Ile36	Val70
Glu195	Asn245	Pro39	Leu71
Glu214	His247	Asp39	Leu73
Ser215	Phe249	Gln40	Gly75
Asp216	Glu251	Arg42	Gly76
His217	Tyr261	Ile44	
Ile218	Tyr266		
Ile221			

Table 6.2   Comparison of OTUB1 (DUB) and CNOT4 (E3) binding sites on UBE2D2 (E2): UBE2D2 amino acids involved in OTUB1 binding (Juang *et al.*, 20120) in comparison with CNOT4 binding site (Dominguez *et al.*, 2004).

1st α-helix		L1 loop of 3rd and 4th β-sheet		L2 loop of 4th β-strand and H2 α-helix	
OTUB1	CNOT4	OTUB1	CNOT4	OTUB1	CNOT4
Met1			Thr58	Ser83	
	Ala2		Asp59	Arg90	
	Leu3		Tyr60	Ser91	
Arg5	Arg5	Pro61		Gln92	
	Ile6	Phe62			Trp93
Lys8		Lys63	Lys63	Ser94	Ser94
	Glu9	Lys66		Pro95	
	Leu10			Ala96	Ala96
	Asn11			Thr98	Thr98
Asp12	Asp12				Ile99
				Lys101	Lys101
					Leu104

6.1.1   Comparison of methodologies

The structure of OTUB1:UBE2D2 complex was solved by crystallisation. As in our studies their initial crystal trials yielded numerous crystals corresponding to UBE2D2~Ub alone. However, they successfully crystallised the complex which was expressed from a single ORF containing a fusion of UBE2D2(ΔC85S):OTUB1(with a deletion of residues 1-24) (Juang *et al.*, 2012). The fused UBE2D2:OTUB1 protein was then charged by ubiquitin *in vitro* and the excess free ubiquitin was removed by gel filtration chromatography. Their approach in producing crystals was quite intriguing, it suggests that they also had problems purifying OTUB1 and solving the structure via NMR. Moreover, the 2D NMR experiment to verify the interaction observed in their crystal structure were performed with ^{15}N-labelled UBE2D3 (not UBE2D2) and UBE2N, while the OTUB1 was unlabelled and the reasons for this are unexplained.

An indirect approach was also observed in crystallisation of OTUB1:UBE2N complex in the work by Wiener *et al.* (2012) where the *C. elegans* (worm) OTUB1 was used instead of human. Worm OTUB1 only shares 34% sequence identity and 56% similarity with human OTUB1 and the N-terminal part is particularly poorly conserved. Because of this, crystal structures were also determined from hybrid OTUB1 containing the N-terminal 45 residues of human OTUB1 and the OTU domain of worm OTUB1 in quaternary complex with UbaI (ubiquitin aldehyde) and UBE2N^{DCA}~Ub. In contrast with the crystal production of the OTUB1:UBE2D2 as a protein fusion, the crystal of the worm OTUB1 and human UBE2N complex were grown from a 1:1 mix of the proteins.

6.1.2   The importance of OTUB1 N-terminal region

The N-terminal region of OTUB1 was truncated in our studies because it forms an unstructured region that proved to be problematic in protein structure determination through NMR and crystallisation (Edelmann *et al.*, 2009). Although removal of this domain advanced our studies, it would appear that this decision was unfortunate because the deleted

N-terminal region (aa1-39) contains a ubiquitin binding domain, which we now know increases binding affinity towards charged E2s (Nakada *et al.*, 2010). Indeed, certain DUBs require ubiquitin binding to obtain their active conformations and to prevent uncontrolled proteolytic activity and some DUBs have additional binding sites with affinity for the target protein that is ubiquitinated (Ventii and Wilkinson, 2008).

There is now growing evidence that intrinsically disordered proteins or protein regions may have an important role in many fundamental biological processes. In many cases, the disordered region undergoes from disordered to ordered transition upon binding with substrates or partners. For example, Ski interaction protein (SKIP) has a highly flexible region (residues 59-129) which is intrinsically unstructured but become ordered upon binding with peptidylprolyl isomerase-like protein 1 (PPIL1) (Wang *et al.*, 2010). In the experiment to test the ability of human OTUB1 to inhibit K63 polyubiquitin synthesis by UBE2N-UBE2V1, only deletion of aa residues 1-15 has no effect on inhibition of K63 ubiquitin synthesis whereas larger deletions exhibit defects (Wiener *et al.*, 2012), indicating that N-terminal residues from 16-39 that were truncated in our study are actually vital for activity.

6.1.3   Agreement with Y2H analysis of UBE2D2 binding site mutants

Lys8, Asp12 and Lys101 of UBE2D2 were mutated in our study (Chapter 5) to verify that they are indeed key residues in the formation of UBE2D2:OTUB1 complexes. These three residues were acknowledged to be involved in UBE2D2:OTUB1 interaction (Juang *et al.*, 2012) hence supporting the Y2H mutagenesis data. However, Glu9 which were also included in our mutagenesis studies appears not to be directly involved at the binding interface (Juang *et al.*, 2012).

Figure 6.2 clearly shows the side chain of Lys8, Asp12 and Lys101 facing towards the direction of OTUB1 binding surface and making direct contact with OTUB1 residues. Glu9 side chain is however not in a favorable position to interact with OTUB1 although this

mutation clearly inhibits complex formation in our Y2H studies. This discrepancy may be explained by subtle enforced changes in local conformation which reduce binding efficiency, even though this residue is not in direct physical contact with OTUB1.
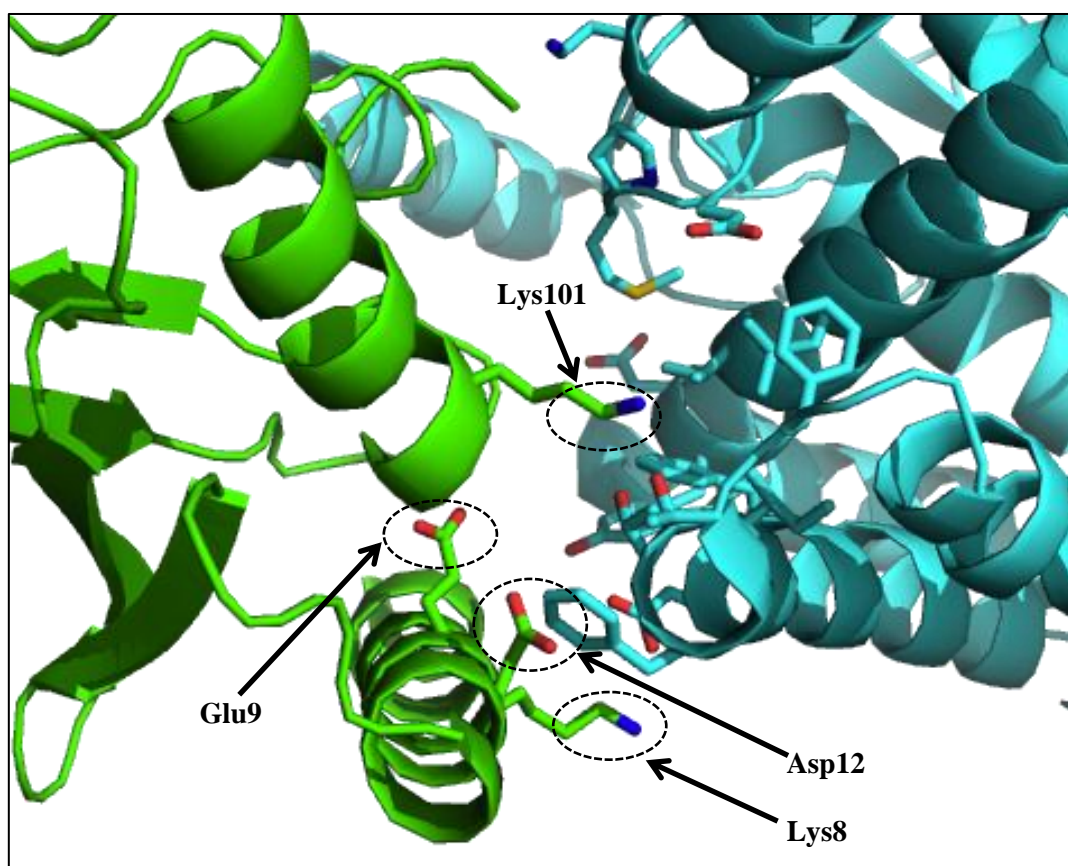

Figure 6.2   **Stereo view of the OTUB1:UBE2D2 binding interface:** Lys8, Asp12 and Lys101 side chains of UBE2D2 facing towards OTUB1 binding surface while Glu9 side chain extending away from OTUB1:UBE2D2 binding interface. OTUB1 is coloured cyan and UBE2D2 in green.

6.2   Insight into OTUB1 physiological function

Otubain-1 (OTUB1), a cysteine protease initially found in ovarian tumours is a DUB that contains a conserved OTU (ovarian tumour) domain which is present in all OTUB proteins (Edelmann *et al.*, 2009). Its N-terminus contains a ubiquitin-binding domain which is thought to interact with ubiquitin to increase its binding affinity to E2 enzymes. OTUB1 was initially proposed to provide an editing function for polyubiquitin chain growth by cleaving tetraubiquitin substrate (Balakirev *et al.*, 2003). Surprisingly, it was later revealed that OTUB1 promotes rather than inhibits the K48-linked self-ubiquitination and proteolysis of E3 RING ligase RNF128 (GRAIL) (Soares *et al.*, 2004). In the proteomic study by Edelmann *et al.* (2010), two more OTUB1 interactors were identified: FUS/TLS and Rack1, both of which are involved in RNA splicing. However, the significance of these interactions and whether FUS/TLS and Rack1 are deubiquitinated by OTUB1 remains unknown.

With a growing body of evidence that DUBs have non-canonical activity (Hanna *et al.*, 2006), OTUB1 proves this principle by its ability to regulate protein ubiquitination reactions. This is achieved by suppressing RNF168-dependent polyubiquitination by a mechanism that is independent of its catalytic activity, by simply binding to and blocking ubiquitin transfer and E3-RING docking to UBE2N. In other words they prevent ubiquitin attachment rather than detaching bound ubiquitin, thereby inhibiting DNA repair (Nakada *et al.*, 2010; Sun *et al.*, 2012) and presumably many other cellular processes as well. The fact that OTUB1 inhibits DNA repair could have therapeutic relevance. Nakada *et al.*, (2010) found that reducing the level of OTUB1 expression restores the process of homologous recombination in cells in which ATM (Ataxia telangiectasia mutated) kinase is inhibited. Thus, OTUB1 depletion can, in principle, mitigate DNA-repair effects. This observation makes the interaction between OTUB1 and UBE2N an attractive target for therapeutic intervention with particular relevance for disorders affecting DNA repair and possibly for use in combination with radiation therapy.

6.3   OTUB1 protein interaction network

The interaction of OTUB1 with UBE2D2 (and UBE2N) could be manipulated to inhibit the ubiquitination cascade in order to establish a new therapeutic drug targets. Being able to block OTUB1 would allow downstream cellular signalling pathways to occur. The involvement of UBE2D2 in the DNA damage response could be the target to allow the repair of DSBs and prevent chromosomal rearrangements that may lead to tumourigenesis and cancer. Although the downregulation of UBE2D2 by suppressing OTUB1 may seem promising, in our analysis OTUB1 interacts with a wider range of proteins. Therefore, we should consider whether the inhibition of OTUB1 will disrupt other physiologically important processes. Figure 6.3 shows the known OTUB1 interaction network which contains a broad range of binding partners, suggesting it has other functions other than DNA damage response. Interestingly, several binding partners are known to have roles in the cell cycle, which may represent an interesting area of future research.
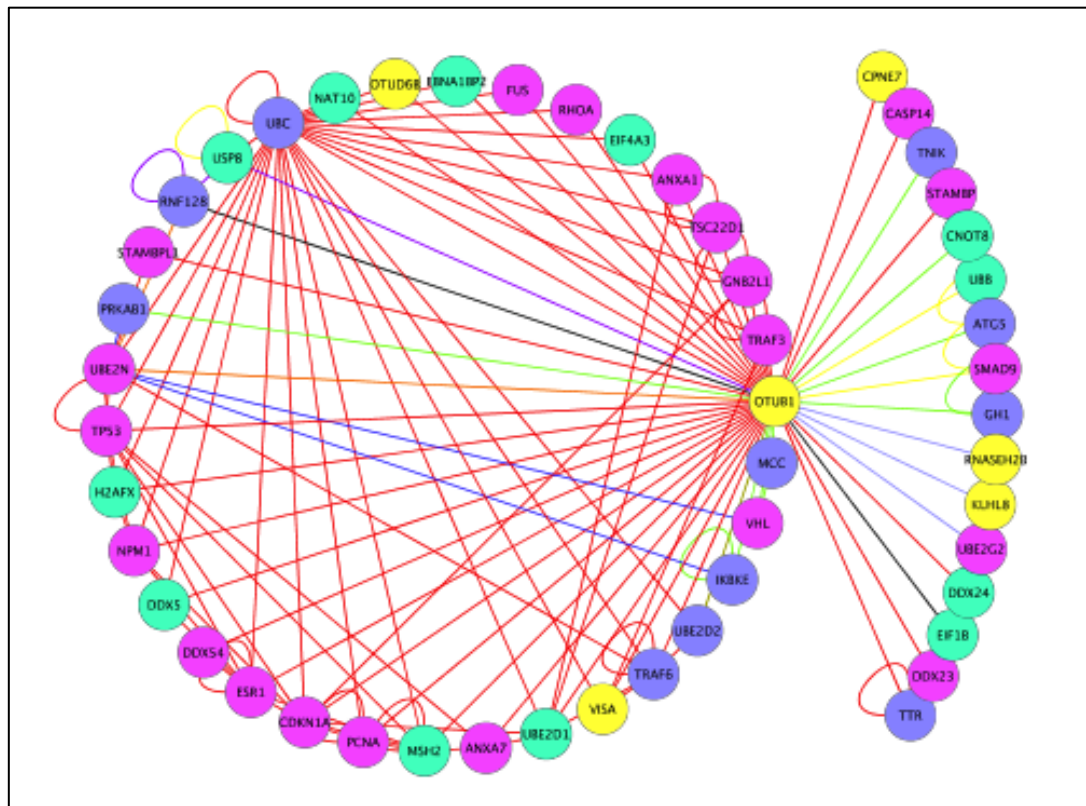
Edge legend:
— BioGRID
— BioGRID|Human
— BioGRID|Human|MS
— BioGRID|Ilog
— BioGRID|MS
— Human
— Ilog
— MS
— MS|Human

Node legend:
🔵 Cytoplasmic
🟢 Nuclear
🟣 Cytoplasmic|Nuclear
🟡 Unknown

Figure 6.3 **OTUB1 interaction network:** Figure shows OTUB1 and its interactors. The colour of the sphere containing the protein indicates the data source of type of experiment performed in order to get the interaction data. The colour of the node indicates the localisation of each protein. Figure generated using Cytoscape ver2.8.2.

6.4   Future directions

Work performed in this study has provided insight into the molecular mechanisms of the OTUB1:E2 protein interaction. In addition, we have generated a large collection of molecular reagents which can be used to further explore the nature and function of OTUB1:E2 and E2:E3-RING complexes in live cells. Preliminary work in this area is looking very promising. Using mutants and constructs generated in this project, other members of the Sanderson lab have successfully used fluorescence cross correlation spectroscopy (FCCS) methods to investigate the nature of OTUB1:E2 complexes in different cellular compartments in live cells. These studies confirmed many of predictions made in this thesis and allow the size, distribution and conditional nature of OTUB1:E2 protein interactions to be monitored in real time *in vivo*. Such methods may also provide a mechanism of screening for compounds which selectively disrupt different OTUB1:E2 or E2:E3-RING complexes, which in turn may offer new therapeutic opportunities.

References

Al-Hakim, A. K., Zagorska, A., Chapman, L., Deak, M., Peggie, M and Alessi, D. R. (2008). "Control of AMPK-related kinases by USP9X and atypical Lys29/Lys33-linked polyubiquitin chains." Biochem. J. **411**(2): 249-260.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002). Molecular Biology of the Cell. 4th edition. New York, Garland Science.

Alexandru, G., Graumann, J., Smith, G. T., Kolawa, N. J., Fang, R., Deshaies, R. J. (2008). "UBXD7 binds multiple ubiquitin ligases and implicates p97 in HIF1α turnover." Cell. **134**(5): 804-816.

Ambroggio, X. I., Rees, D. C. and Deshaies R. J. (2003). "JAMM: a metalloprotease-like zinc site in the proteasome and signalosome." PLoS Biol. **2**(1): e2.

Amerik, A. Y. and Hochstrasser, M. (2004). "Mechanism and function of deubiquitinating enzymes." Biochim. Biophys. Acta. **1695**(1-3): 189-207.

Anan, T., Nagata, Y., Koga, H., Honda, Y., Yabuki, N., Miyamoto, C., Kuwano, A., Matsuda, I., Endo, F., Saya, H. and Nakao, M. (1998). "Human ubiquitin-protein ligase Nedd4: expression, subcellular localization and selective interaction with ubiquitin-conjugating enzymes." Genes. Cells. **3**(11): 751-763.

Aranda, B., Achuthan, P., Alam-Faruque, Y., Armean, I., Bridge, A., Derow, C., Feuermann, M., Ghanbarian, A. T., Kerrien, S., Khadake, J., Kerssemakers, J., Leroy, C., Menden, M., Michaut, M., Montecchi-Palazzi, L., Neuhauser, S. N., Orchard, S., Perreau, V., Roechert, B., van Eijk, K. and Hermjakob, H. (2010). "The IntAct molecular interaction database in 2010." Nucleic Acids Res. **38**(suppl 1): D525-D531.

Arnason, T. and Ellison, M. J. (1994). "Stress resistance in *Saccharomyces cerevisiae* is strongly correlated with assembly of a novel type of multiubiquitin chain." Mol. Cell Biol. **14**(12): 7876-7883.

Baboshina, O. V. and Haas, A. L. (1996). "Novel multiubiquitin chain linkages catalyzed by the conjugating enzymes E2EPF and RAD6 are recognized by 26 S proteasome subunit 5." J. Biol. Chem. **271**(5): 2823-2831.

Balakirev, M. Y., Tcherniuk, S. O., Jaquinod, M. and Chroboczek, J. (2003). "Otubains: a new family of cysteine proteases in the ubiquitin pathway." EMBO Rep. **4**(5): 517-522.

Bartke, T., Pohl, C., Pyrowokalis, G. and Jentsch, S. (2004). "Dual role of BRUCE as an antiapoptotic IAP and a chimeric E2/E3 ubiquitin ligase." Mol. Cell. **14**(6): 801-811.

Bech-Otschir, D., Kraft, R., Huang, X., Heinklein, P., Kapelari, B., Pollmann, C. and Dubiel, F. (2001). "COP9 signalosome-specific phosphorylation targets p53 to degradation by the ubiquitin system." EMBO J. **20**(7): 1630-1639.

Bernard, P. and Couturier, M. (1992). "Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes." J. Mol. Biol. **226**(3): 735-745.

Berrow, N. S., Büssow, K., Coutard, B., Diprose, J., Ekberg, M., Folkers, G. E., Levy, N., Lieu, V., Owens, R. J., Peleg, Y., Pinaglia, C., Quevillon-Cheruel, S., Salim, L., Scheich, C., Vincentelli, R. and Busso, D. (2006). "Recombinant protein expression and solubility screening in *Escherichia coli*: a comparative study." Acta Crystallogr. D. Biol. Crystallogr. **62**(10): 1218-1226.

Bignell, G. R., Warren, W., Seal, S., Takahashi, M., Rapley, E., Barfoot, R., Green, H., Brown, C., Biggs, P. J., Lakhani, S. R., Jones, C., Hansen, J., Blair, E., Hofmann, B., Siebert, R., Turner, G., Evans, D. G., Schrander-Stumpel, C., Beemer, F. A., van Den Ouweland, A., Halley, D., Delpech, B., Cleveland, M. G., Leigh, I., Leisti, J. and Rasmussen, S. (2000). "Identification of the familial cylindromatosis tumour-suppressor gene." Nat. Genet. **25**(2): 160-165.

Borden, K. L. and Freemont, P. S. (1996). "The RING finger domain: a recent example of a sequence-structure family." Curr. Opin. Struct. Biol. **6**(3): 395-401.

Bram, R. J., Lue, N. F. and Kornberg, R. D. (1986). "A GAL family of upstream activating sequences in yeast: roles in both induction and repression of transcription." EMBO J. **5**(3): 603-608.

Breitkreutz, B. J., Stark, C., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., Oughtred, R., Lackner, D. H., Bähler, J., Wood, V., Dolinski, K. and Tyers, M. (2008). "The BioGRID interaction database: 2008 update." Nucleic Acids Res. **36**(suppl 1): D637-D640.

Brinkmann, U., Mattes, R. E. and Buckel, P. (1989). "High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the dnaY gene product." Gene. **85**(1): 109-114.

Brzovic, P. S. and Klevit, R. E. (2006). "Ubiquitin transfer from the E2 perspective: Why is UbcH5 so promiscuous?" Cell Cycle. **5**(24): 2867-2873.

Buechler, Y. J., Herberg, F. W. and Taylor, S. S. (1993). "Regulation-defective mutants of type I cAMP-dependent protein kinase. Consequences of replacing arginine 94 and arginine 95." J. Biol. Chem. **268**(22): 16495-16503.

Burnett, B., F. Li, F. and Pittman, R. N. (2003). "The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity." Hum. Mol. Genet. **12**(23): 3195-3205.

Caballero, O. L., Resto, V., Patturajan, M., Meerzaman, D., Guo, M. Z., Engles, J., Yochem, R., Ratovitski, E., Sidransky, D. and Jen, J. (2002). "Interaction and colocalization of PGP9.5 with JAB1 and p27(Kip1)." Oncogene. **21**(19): 3003-3010.

Chastagner, P., Israël, A. and Brou, C. (2006). "Itch/AIP4 mediates Deltex degradation through the formation of K29-linked polyubiquitin chains." EMBO Rep. **7**(11): 1147-1153.

Chatr-aryamontri, A., Ceol, A., Palazzi, L. M., Nardelli, G., Schneider, M. V., Castagnoli, L. and Cesareni, G (2007). "MINT: the Molecular INTeraction database." Nucleic Acids Res. **35**(suppl 1): D572-D574.

Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K. and Varshavsky, A. (1989). "A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein." Science. **243**(4898): 1576-1583.

Christensen, D. E., Brzovic, P. S. and Klevit, R. E. (2007). "E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages." Nat. Struct. Mol. Biol. **14**(10): 941-948.

Ciechanover, A., Orian, A. and Schwartz, A. L. (2000). "Ubiquitin-mediated proteolysis: biological regulation via destruction." BioEssays. **22**(5): 442-451.

Clague, M. J., Coulson, J. M. and Urbé, S. (2012). "Cellular functions of the DUBs." J. Cell Sci. **125**(2): 277-286.

Collins, F. S.: Mammalian Gene Collection (MGC) Program Team, contributed by Collins, F. S. (2002). "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. USA. **99**(26): 16899-16903.

Cooper, G. M. (2000). The Cell: A Molecular Approach. Sunderland (MA), Sinauer Associates.

David, Y., Ziv, T., Admon, A. and Navon, A. (2010). "The E2 ubiquitin-conjugating enzymes direct polyubiquitination to preferred lysines." J. Biol. Chem. **285**(12): 8595-8604.

Deshaies, R. J. and Joazeiro, C. A. P. (2009). "RING domain E3 ubiquitin ligases." Annu. Rev. Biochem. 78: 399-434.

Doherty, A. J., Ashford, S. R., Brannigan, J. A. and Wigley, D. B. (1995). "A superior host strain for the over-expression of cloned genes using the T7 promoter based vectors." Nucleic Acids Res. **23**(11): 2074-2075.

Dominguez, C., Bonvin, A. M., Winkler, G. S., van Schaik, F. M., Timmers, H. T. and Boelens, R. (2004). "Structural model of the UbcH5B/CNOT4 complex

revealed by combining NMR, mutagenesis, and docking approaches." <u>Structure.</u> **12**(4): 633-644.

Dye, B. T. and Schulman, B. A. (2007). "Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins." <u>Annu. Rev. Biophys. Biomol. Struct.</u>. **36**: 131-150.

Eddins, M. J., Carlile, C. M., Gomez, K. M., Pickart, C. M. and Wolberger, C. (2006). "Mms2-Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation." <u>Nat. Struct. Mol. Biol.</u> **13**(10): 915-920.

Edelmann, M. J., Iphöfer, A., Akutsu, M., Altun, M., di Gleria, K., Kramer, H. B., Fiebiger, E., Dhe-paganon, S. and Kessler, B. M. (2009). "Structural basis and specificity of human otubain 1-mediated deubiquitination." <u>Biochem. J.</u> **418**(2): 379-390.

Ennifar, E. (2012). "X-ray Crystallography as a tool for mechanism-of-action studies and drug discovery." <u>Curr. Pharm. Biotechnol.</u> **Epub**.

Fernández, C. and Wider, G. (2003). "TROSY in NMR studies of the structure and function of large biological macromolecules." <u>Curr. Opin. Struct. Biol.</u> **13**(5): 570-580.

Fields, S. and Song, O. (1989). "A novel genetic system to detect protein-protein interactions." <u>Nature.</u> **340**(6230): 245-246.

Finley, D., Bartel, B. and Varshavsky, A. (1989). "The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis." <u>Nature.</u> **338**(6214): 394-401.

Fisk, H. A. and Yaffe, M. P. (1999). "A role for ubiquitination in mitochondrial inheritance in *Saccharomyces cerevisiae*." <u>J. Cell Biol.</u> **145**(6): 1199-1208.

Flajolet, M., Rotondo, G., Daviet, L., Bergametti, F., Inchauspé, G., Tiollais, P., Transy, C. and Legrain, P. (2000). "A genomic approach of the hepatitis C virus generates a protein interaction map." <u>Gene.</u> **242**(1,Äì2): 369-379.

Flagg, J. L. and Wilson, T.H. (1976). "*lacY* mutant of *Escherichia coli* with altered physiology of lactose induction." <u>J. Bacteriol.</u> **128**(3): 701-707.

Formstecher, E., Aresta, S., Collura, V., Hamburger, A., Meil, A., Trehin, A., Reverdy, C., Betin, V., Maire, S., Brun, C., Jacq, B., Arpin, M., Bellaiche, Y., Bellusci, S., Benaroch, P., Bornens, M., Chanet, R., Chavrier, P., Delattre, O., Doye, V., Fehon, R., Faye, G., Galli, T., Girault, J. A., Goud, B., de Gunzburg, J., Johannes, L., Junier, M. P., Mirouse, V., Mukherjee, A., Papadopoulo, D., Perez, F., Plessis, A., Rossé, C., Saule, S., Stoppa-Lyonnet, D., Vincent, A., White, M., Legrain, P., Wojcik, J., Camonis, J. and Daviet L. (2005). "Protein interaction mapping: A *Drosophila* case study." <u>Genome Res.</u> **15**(3): 376-384.

Frias-Staheli, N., Giannakopoulos, N. V., Kikkert, M., Taylor, S. L., Bridgen, A., Paragas, J., Richt, J. A., Rowland, R. R., Schmaljohn, C. S., Lenschow, D. J., Snijder, E. J., García-Sastre, A., Virgin, H. W. 4th, (2007). "Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses." <u>Cell Host Microbe.</u> **2**(6): 404-416.

Gao, G., William, J. G. and Campbell, S. L. (2004). Protein-protein interaction analysis by Nuclear Magnetic Resonance spectroscopy. <u>Methods Mol. Biol.</u> H. Fu. Totowa NJ, Humana Press Inc. **261:** Protein-Protein Interaction: Methods and Protocol.

Garcia-Cuellar, M. P., Merderer, D. and Slany, R. K. (2009). "Identification of protein interaction partners by the yeast two-hybrid system." <u>Methods Mol. Biol.</u> **538**: 347-367.

Gill, G. (2004). "SUMO and ubiquitin in the nucleus: different functions, similar mechanisms?" <u>Genes and Dev.</u> **18**: 2046-2059.

Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla,

S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C. A., Finley, R. L. Jr, White, K. P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A.,McKenna, M. P., Chant, J. and Rothberg, J. M. (2003). "A protein interaction map of *Drosophila melanogaster*." Science. **302**(5651): 1727-1736.

Glickman, M. H. and Ciechanover, A. (2002). "The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction." Physiol. Rev. **82**(2): 373-428.

Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., dhe-Paganon, S., Park, H. W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S. H., Rao, Z., Shi, Y., Terwilliger, T. C., Kim, C. Y., Hung, L. W., Waldo, G. S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J. L., Stevens, R. C., Lesley, S. A., Wilson, I. A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M. I., Eschenfeldt, W. H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S. K., Emtage, J. S., Sauder, J. M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S. C., Bonanno, J. B., Fiser, A., Swaminathan, S., Studier, F. W., Chance, M. R., Sali, A., Acton, T. B., Xiao, R., Zhao, L., Ma, L. C., Hunt, J. F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C. K., Wang, D., Wang, H., Jiang, M., Montelione, G. T., Stuart, D. I., Owens, R. J,, Daenke, S., Schütz, A., Heinemann, U., Yokoyama, S., Büssow, K. and Gunsalus, K. C. (2008). "Protein production and purification." Nat. Methods. **5**(2): 135-146.

Grodberg, J. and Dunn, J. J. (1988). "ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification." J Bacteriol. **170**(3): 1245-1253.

Groettrup, M., Pelzer, C., Schmidtke, G. and Hoffman, K. (2008). "Activating the ubiquitin family: UBA6 challenges the field." Trends Biochem. Sci. **33**(5): 230-237.

Guo, D., Rajamäki, M. L, Saarma, M. and Valkonen, J. P. (2001). "Towards a protein interaction map of potyviruses: protein interaction matrixes of two potyviruses based on the yeast two-hybrid system." J. Gen. Virol. **82**(4): 935-939.

Haas, A. L., Warms, J. V., Hershko, A. and Rose, I. A. (1982). "Ubiquitin-activating enzyme. Machanism and role is protein-ubiquitin conjugation." J. Biol. Chem. **257**: 2543-2548.

Haglund, K., Di Fiore, P. P. and Dikic, I. (2003). "Distinct monoubiquitin signals in receptor endocytosis." Trends Biochem. Sci. **28**(11): 598-604.

Hanna, J., Hathaway, N. A., Tone, Y., Crosas, B., Elsasser, S., Kirkpatrick, D. S., Leggett, D. S., Gygi, S. P., King, R. W. and Finley, D. (2006). "Deubiquitinating enzyme Ubp6 functions noncatalytically to delay proteasomal degradation." Cell. **127**(1): 99-111.

Hartley, J. L., Temple, G. F. and Brasch, M. A. (2000). "DNA cloning using *in vitro* site-specific recombination." Genome Res. **10**(11): 1788-1795.

Hayes, S. D. (2009). "A systematic analysis of DUB and RING E3 ligase interactions. PhD, Universisty of Liverpool.

Hershko, A. and Ciechanover, A. (1998). "The ubiquitin system." Annu. Rev. Biochem. **67**(1): 425-479.

Hicke, L. and Dunn, R. (2003). "Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins." Annu. Rev. Cell Dev. Biol. **19**(1): 141-172.

Hochstrasser, M. (1996). "Ubiquitin-dependent protein degradation." Annu. Rev. Genet. **30**(1): 405-439.

Hochstrasser, M. (2009). "Origin and function of ubiquitin-like proteins." Nature. **458**(7237): 422-429.

Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G. and Jentsch, S. (2002). "RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO." Nature. **419**(6903): 135-141.

Hong, S., Lee, S., Cho, S. G. and Kang, S. (2008). "UbcH6 interacts with and ubiquitinates the SCA1 gene product ataxin-1." Biochem. Biophys. Res. Commun. **371**(2): 256-260.

Hosfield, D., Palan, J., Hilgers, M., Scheibe, D., McRee, D. E., Stevens, R. C. (2003). "A fully integrated protein crystallization platform for small-molecule drug discovery." J. Struct. Biol. **142**(1): 207-217.

Houben, K., Dominguez, C., van Schaik, F. M., Timmers, H. T., Bonvin, A. M. and Boelens, R. (2004). "Solution structure of the ubiquitin-conjugating enzyme UbcH5B." J. Mol. Biol. **344**(2): 513-526.

Ikeda, F. and Dikic, I. (2008). "Atypical ubiquitin chains: new molecular signals." EMBO Rep. **9**(6): 536-542.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. and Sakaki, Y. (2001). "A comprehensive two-hybrid analysis to explore the yeast protein interactome." Proc. Natl. Acad. Sci. USA. **98**(8): 4569-4574.

Iwai, K. and Tokunaga, F. (2009). "Linear polyubiquitination: a new regulator of NF-κB activation." EMBO Rep. **10**(7): 706-713.

James, P., Halladay, J. and Craig, E. A. (1996). "Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast." Genetics. **144**(4): 1425-1436.

Jancarik, J. and Kim, S. H. (1991). "Sparse matrix sampling: a screening method for crystallization of proteins." J. Appl. Cryst. **24**(4): 409-411.

Jin, L., Williamson, A., Banerjee, S., Philipp, I. and Rape, M. (2008). "Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex." Cell. **133**(4): 653-665.

Joazeiro, C. A., Wing, S. S., Huang, H. K., Leverson, J. D., Hunter, T., Liu, Y. C. (1999). "The tyrosine kinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin-protein ligase." Science. **286**(5438): 309-312.

Joazeiro, C. A. and Weissman, A. M. (2000). "RING finger proteins: mediators of ubiquitin ligase activity." Cell. **102**(5): 549-552.

Juang, Y. C., Landry, M. C., Sanches, M., Vittal, V., Leung, C. C., Ceccarelli, D. F., Mateo, A. R., Pruneda, J. N., Mao, D. Y., Szilard, R. K., Orlicky, S., Munro, M., Brzovic, P. S., Klevit, R. E., Sicheri, F. and Durocher, D. (2012). "OTUB1 co-opts Lys48-linked ubiquitin recognition to suppress E2 enzyme function." Mol. Cell. **45**(3): 384-397.

Keshava Prasad, T. S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A., Balakrishnan, L., Marimuthu, A., Banerjee, S., Somanathan, D. S., Sebastian, A., Rani, S., Ray, S., Harrys Kishore, C. J., Kanth, S., Ahmed, M., Kashyap, M. K., Mohmood, R., Ramachandra, Y. L., Krishna, V., Rahiman, B. A., Mohan, S., Ranganathan, P., Ramabadran, S., Chaerkady, R. and Pandey, A. (2009). "Human protein reference database--2009 update." Nucleic Acids Res. **37**(suppl 1): D767-D772.

Kim, H. T, Kim, K. P., Lledias, F., Kisselev, A. F., Scaglione, K. M., Skowyra, D., Gygi, S. P. and Goldberg, A. L. (2007). "Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages." J. Biol. Chem. **282**(24): 17375-17386.

Kim, H. T., Kim, K. P., Uchiki, T., Gygi, S. P. and Goldberg, A. L. (2009). "S5a promotes protein degradation by blocking synthesis of nondegradable forked ubiquitin chains." EMBO J. **28**(13): 1867-1877.

Kimura, Y. and Tanaka, K. (2010). "Regulatory mechanisms involved in the control of ubiquitin homeostasis." J. Biochem. **147**(6): 793-798.

Kirkpatrick, D. S., Hathaway, N. A., Hanna, J., Elsasser, S., Rush, J., Finley, D., King, R. W. and Gygi, S. P. (2006). "Quantitative analysis of *in vitro* ubiquitinated cyclin B1 reveals complex chain topology." Nat. Cell. Biol. **8**(7): 700-710.

Knopp, M., Schiffer, H. H., Rupp, S. and Wolf, D. H. (1993). "Vacuolar/lysosomal proteolysis: proteases, substrates, mechanisms." Curr. Opin. Cell. Biol. **5**(6): 990-996.

Komander, D., Clague, M. J. and Urbé, S. (2009). "Breaking the chains: structure and function of the deubiquitinases." Nat. Rev. Mol. Cell. Biol. **10**(8): 550-563.

Komander, D. (2009). "The emerging complexity of protein ubiquitination." Biochem. Soc. Trans. **37**(Pt 5): 937–953.

Koonin, E. V. and Abagyan, R. A. (1997). "TSG101 may be the prototype of a class of dominant negative ubiquitin regulators." Nature Genet. **16**(4): 330-331.

LaCount, D. J., Vignali, M., Chettier, R., Phansalkar A., Bell, R., Hesselberth, J. R., Schoenfeld, L. W., Ota, I., Sahasrabudhe, S., Kurschner, C., Fields, S. and Hughes, R. E. (2005). "A protein interaction network of the malaria parasite *Plasmodium falciparum*." Nature. **438**(7064): 103-107.

Laine, A., Topisirovic, I., Zhai, D., Reed, J. C., Borden, K. L. B. and Ronai, Z. (2006). "Regulation of p53 localization and activity by Ubc13." Mol. Cell Biol. **26**(23): 8901-8913.

Landy, A. (1989). "Dynamic, structural, and regulatory aspects of lambda site-specific recombination." Annu. Rev. Biochem. **58**(1): 913-941.

Leavitt, S. and Freire, E. (2001). "Direct measurement of protein binding energetics by isothermal titration calorimetry." Curr. Opin. Struct. Biol. **11**(5): 560-566.

Lederberg, J. (1948). "Gene control of beta-galactosidase in *Escherichia coli*." Genetics. **33**: 617.

Lehner, B., Semple, J. I., Brown, S. E., Counsell, D., Campbell, R.D. and Sanderson C. M. (2004a). "Analysis of a high-throughput yeast two-hybrid system and its use to predict the function of intracellular proteins encoded within the human MHC class III region." Genomics. **83**(1): 153-167.

Lehner, B. and Sanderson, C. M. (2004b). "A protein interaction framework for human mRNA degradation." Genome Res. **14**(7): 1315-1323.

Li, L., Liao, J., Ruland, J., Mak, T. W. and Cohen, N. (2001). "A TSG101/MDM2 regulatory loop modulates MDM2 degradation and MDM2/p53 feedback control." Proc. Natl. Acad. Sci. USA. **98**(4): 1619-1624.

Li, S., Armstrong, C., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D., Chesneau, A., Hao, T., Goldberg, D. S., Li, N., Martinez, M., Rual, J. F., Lamesch, P., Xu, L., Tewari, M., Wong, S. L., Zhang, L. V., Berriz, G. F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q., Gabel, H. W., Elewa, A., Baumgartner, B., Rose, D. J., Yu, H., Bosak, S., Sequerra, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., Van Den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K. C., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D. E. and Vidal, M. (2004). "A map of the interactome network of the metazoan *C. elegans*." Science. **303**(5657): 540-543.

Loda, M., Cukor, B., Tam, S. W., Lavin, P., Fiorentino, M., Draetta, G. F., Jessup, J. M. and Pagano, M. (1997). "Increased proteasome-dependent degradation of the cyclin-dependent kinase inhibitor p27 in aggressive colorectal carcinomas." Nat. Med. 3(2): 231-234.

Lorick, K. L., Jensen, J. P., Fang, S., Ong, A. M., Hatakeyama, S. and Weissman, A. M. (1999). "RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination." Proc. Natl. Acad. Sci. USA. **96**(20): 11364-11369.

Markson, G., Kiel, C., Hyde, R., Brown, S., Charalabous, P., Bremm, A., Semple, J., Woodsmith, J., Duley, S., Salehi-Ashtiani, K., Vidal, M., Komander, D., Serrano, L., Lehner, P. and Sanderson, C. M. (2009). "Analysis of the human E2 ubiquitin conjugating enzyme protein interaction network." Genome Res. **19**(10): 1905-1911.

Mastrandrea, L. D., You, J., Niles, E. G. and Pickart, C. M. (1999). "E2/E3-mediated assembly of lysine 29-linked polyubiquitin chains." J. Biol. Chem. **274**(38):27299-306.

Matic, I., van Hagen, M., Schimmel, J., Macek, B., Ogg, S. C., Tatham, M. H., Hay, R. T., Lamond, A. I., Mann, M. and Vertegaal, A. C. (2008). "*In vivo* identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an *in vitro* to *in vivo* strategy." Mol. Cell Proteomics. **7**(1): 132-144.

Maxwell, P. H., Wiesener, M. S., Chang, G. W., Clifford, S. C., Vaux, E. C., Cockman, M. E., Wykoff, C. C., Pugh, C. W., Maher, E. R. and Ratcliffe, P. J. (1999). "The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis." Nature. **399**(6733): 271.

McCraith, S., Holtzman, T., Moss, B. and Fields, S. (2000). "Genome-wide analysis of vaccinia virus protein-protein interactions." Proc. Natl. Acad. Sci. USA. **97**(9): 4879-4884.

McCullough, J., Clague, M. J. and Urbé, S. (2004). "AMSH *is* an endosome-associated ubiquitin isopeptidase." J. Cell Biol. **166**(4): 487-492.

McIntosh, L. P. and Dahlquist, F. W. (1990). "Biosynthetic incorporation of 15N and 13C for assignment and interpretation of nuclear magnetic resonance spectra of proteins." Q. Rev. Biophys. **23**(1): 1-38.

McPherson, A. (2001). "Crystallization of proteins and protein–ligand complexes." eLS. John Wiley & Sons, Ltd.

Merkley, N. and Shaw, G. S. (2004). "solution structure of the flexible class II ubiquitin-conjugating enzyme Ubc1 provides insights for polyubiquitin chain assembly." J. Biol. Chem. **279**(45): 47139-47147.

Michelle, C., Vourc'h, P., Mignon, L. and Andres, C. R. (2009). "What was the set of ubiquitin and ubiquitin-like conjugating enzymes in the eukaryote common ancestor?." J. Mol. Evol. **68**(6) 616-628.

MicroCal ITC Manual (2001). VP-ITC Instruction Manual Northampton, MA, Microcal Inc.

Nakada, S., Tai, I., Panier, S., Al-Hakim, A., Iemura, S., Juang, Y., O'Donnell, L., Kumakubo, A., Munro, M., Sicheri, F., Gingras, A., Natsume, T., Suda, T. and Durocher, D. (2010). "Non-canonical inhibition of DNA damage-dependent ubiquitination by OTUB1." Nature. **466**(7309): 941-946.

Nijman, S. M., Luna-Vargas, M. P., Velds, A., Brummelkamp, T. R., Dirac, A. M., Sixma, T. K. and Bernards, R. (2005). "A genomic and functional inventory of deubiquitinating enzymes." Cell. **123**(5): 773-786.

Nishikawa, H., Ooka, S., Sato, K., Arima, K., Okamoto, J., Klevit, R. E., Fukuda, M. and Ohta, T. (2004). "Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase." J. Biol. Chem. **279**(6): 3916-3924.

O'Brien, R., Ladbury, J. E. and Chowdry, B. Z. (2000). "Isothermal titration calorimetry of biomolecules." Protein-ligand interactions: hydrodynamics and calorimetry. Harding S. E. and Chowdry B.Z., eds, Oxford University Press.

Okuma, T., Honda, R., Ichikawa, G., Tsumagari, N. and Yasuda, H. (1999). "*In vitro* SUMO-1 modification requires two enzymatic steps, E1 and E2." Biochem. Biophys. Res. Commun. **254**(3): 693-698.

Oliva, A., Llabrés, M. and Fariña, J. B. (2004). "Applications of multi-angle laser light-scattering detection in the analysis of peptides and proteins." Curr. Drug. Discov. Technol. **1**(3): 229-242.

Olzmann, J. A., Li, L., Chudaev, M. V., Chen, J., Perez, F. A., Palmiter, R. D. and Chin, L. S. (2007). "Parkin-mediated K63-linked polyubiquitination targets misfolded DJ-1 to aggresomes via binding to HDAC6." J. Cell Biol. **178**(6): 1025-1038.

Ordureau, A., Smith, H., Windheim, M., Peggie, M., Carrick, E., Morrice, N. and Cohen, P. (2008). "The IRAK-catalysed activation of the E3 ligase function of

Pellino isoforms induces the Lys63-linked polyubiquitination of IRAK1." Biochem. J. **409**(1): 43-52.

Özkan, E., Yu, H. and Deisenhofer, J. (2005). "Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases." Proc. Natl. Acad. Sci. USA. **102**(52): 18890-18895.

Parrish, J. R., Gulyas, K. D. and Finley, R. L. Jr. (2006). "Yeast two-hybrid contributions to interactome mapping." Curr. Opin. Biotechnol. **17**(4): 387-393.

Pickart, C. M. (2001). "Mechanisms underlying ubiquitination." Annu. Rev. Biochem. **70**: 503-533.

Pickett, J. (2007). "UBE1, you're not alone." Nat. Rev. Mol. Cell. Biol. **8**(8): 599-599.

Ponting, C. P., Cai, Y. D. and Bork, P. (1997). "The breast cancer gene product TSG101: a regulator of ubiquitination?" J. Mol. Med. **75**(7): 467-469.

Ptak, C., Gwozd, C., Huzil, J. T., Gwozd, T. J., Garen, G., Ellison, M. J. (2001). "Creation of a pluripotent ubiquitin-conjugating enzyme." Mol. Cell. Biol. **21**(19): 6537–6548.

Radaev, S., Li, S. and Sun, P. D. (2006). "A survey of protein-protein complex crystallizations." Acta. Crystallogr. D. Biol. Crystallogr. **62**(6): 605-612.

Rahighi, S., Ikeda, F., Kawasaki, M., Akutsu, M., Suzuki, N., Kato, R., Kensche, T., Uejima, T., Bloor, S., Komander, D., Randow, F., Wakatsuki, S. and Dikic, I. (2009). "Specific recognition of linear ubiquitin chains by NEMO is important for NF-κB activation." Cell. **136**(6): 1098-1109.

Raiborg, C. and Stenmark, H. (2009). "The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins." Nature. **458**(7237): 445-452.

Redman, K. L. and Rechsteiner, M. (1989). "Identification of the long ubiquitin extension as ribosomal protein S27a." Nature. **338**(6214): 438-440.

Reyes-Turcu, F. E., Ventii, K. H. and Wilkinson, K. D. (2009). "Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes." Annu. Rev. Biochem. **78**: 363-397

Rhodes, G. (1993). "Chapter 3 - Protein Crystals." Crystallography Made Crystal Clear (Third Edition). Burlington, Academic Press**:** 31-47.

Roberts, B. E. and Paterson, B. M. (1973). "Efficient translation of Tobacco Mosaic Virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ." Proc. Natl. Acad. Sci. USA. **70**(8): 2330-2334.

Rual, J. F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P., Clingingsmith, T. R., Hartley, J. L., Esposito, D., Cheo, D., Moore, T., Simmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C., Vandenhaute, J., Cusick, M. E., Albala, J. S., Hill, D. E. and Vidal, M. (2004). "Human ORFeome version 1.1: a platform for reverse proteomics." Genome Res. **14**(10b): 2128-2135.

Rual, J., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D. S., Zhang, L. V., Wong, S. L., Franklin, G., Li, S., Albala, J. S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R. S., Vandenhaute, J., Zoghbi, H. Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M. E., Hill, D. E., Roth, F. P. and Vidal, M. (2005). "Towards a proteome-scale map of the human protein-protein interaction network." Nature. **437**(7062): 1173-1178.

Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. and Howley, P. M (1990). "The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53." Cell. **63**(6): 1129-1136.

Schuberth, C. and Buchberger, A. (2008). "UBX domain proteins: major regulators of the AAA ATPase Cdc48/p97." Cell Mol. Life Sci. **65**(15): 2360-2371.

Semple, J. I., Prime, G., Wallis, L. J., Sanderson, C. M. and Marki, D. (2005). "Two-hybrid reporter vectors for gap repair cloning." BioTechniques. **38**(6): 927-934.

Shan, T., Tang, Z. L., Guo, D. Z., Yang, S. L., Mu, Y. L., Ma, Y. H., Guan, W. J. and Li, K. (2008). "Partial molecular cloning, characterization, and analysis of the subcellular localization and expression patterns of the porcine OTUB1 gene." Mol. Biol. Rep. **36**(6): 1573-1577.

Shembade, N., Ma, A. and Harhaj, E. W. (2010). "Inhibition of NF-κB signaling by A20 through disruption of ubiquitin enzyme complexes." Science. **327**(5969): 1135-1139.

Shuman, S. (1994). "Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase." J. Biol. Chem. **269**(51): 32678-32684.

Soares, L., Seroogy, C., Skrenta, H., Anandasabapathy, N., Lovelace, P., Chung, C. D., Engleman, E. and Fathman, C. G. (2004). "Two isoforms of otubain 1 regulate T cell anergy via GRAIL." Nat. Immunol. **5**(1): 45-54.

Sowa, M. E., Bennett, E. J., Gygi, S. P. and Harper, J. W. (2009). "Defining the human deubiquitinating enzyme interaction landscape." Cell. **138**(2): 389-403.

Spence, J., Gali, R. R., Dittmar, G., Sherman, F., Karin, M. and Finley, D. (2000). "Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain." Cell. **102**(1): 67-76.

Stanyon, C. A., Liu, G., Mangiola, B. A., Patel, N., Giot, L., Kuang, B., Zhang, H., Zhong, J. and Finley, R. L. Jr. (2004). "A *Drosophila* protein-interaction map centered on cell-cycle regulators." Genome Biol. **5**(12): R96.

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenher, A., Koeppen, S., Timm, J., Mintzlaff, S., Abraham, C., Bock, N., Kietzmann, S., Goedde, A., Toksöz, E., Droege, A., Krobitsch, S., Korn, B., Birchmeier, W., Lehrach, H. and Wanker, E. E. (2005). "A human protein-protein interaction network: a resource for annotating the proteome." Cell. **122**(6): 957-968.

Sun, S. C. (2008). "Deubiquitylation and regulation of the immune response." Nat. Rev. Immunol. **8**(7): 501-511.

Sun, T., Chance, R. R., Graessley, W. W. and Lohse, D. J. (2004). "A study of the separation principle in size exclusion chromatography." Macromolecules. **37**(11): 4304-4312.

Sun, X. X., Challagundla, K. B. and Dai, M. S. (2012). "Positive regulation of p53 stability and activity by the deubiquitinating enzyme Otubain 1." EMBO J. **31**(3): 576-592

Swaminathan, S., Amerik, A. Y. and Hochstrasser, M. (1999). "The Doa4 deubiquitinating enzyme is required for ubiquitin homeostasis in yeast." Mol. Biol. Cell. **10**(8): 2583-2594.

Tan, J. M. M., Wong, E. S. P., Kirkpatrick, D. S., Pletnikova, O., Ko, H. S., Tay, S. P., Ho, M. W. L., Troncoso, J., Gygi, S. P., Lee, M. K., Dawson, V. L., Dawson, T. M. and Lim, K. L. (2008). "Lysine 63-linked ubiquitination promotes the formation and autophagic clearance of protein inclusions associated with neurodegenerative diseases." Hum. Mol. Genet. **17**(3): 431-439.

Tashiro, K., Pando, M. P., Kanegae, Y., Wamsley, P. M., Inoue, S. and Verma, I. M. (1997). "Direct involvement of the ubiquitin-conjugating enzyme Ubc9/Hus5 in the degradation of IκBα." Proc. Natl. Acad. Sci. USA. **94**(15): 7862-7867.

Tatham, M. H., Jaffray, E., Vaughan, O. A., Desterro, J. M., Botting, C. H., Naismith, J. H. and Hay, R. T. (2001). "Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9." J. Biol. Chem. **276**(38): 35368-35374.

Thrower, J. S., Hoffman, L., Rechsteiner, M. and Pickart, C. M. (2000). "Recognition of the polyubiquitin proteolytic signal." EMBO J. **19**(1): 94-102.

Uetz, P., Dong, Y. A., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., Rajagopala, S. V., Roupelieva, M., Rose, D., Fossum, E. and Haas, J. (2006). "Herpesviral protein networks and their interaction with the human proteome." Science. **311**(5758): 239-242.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin,

B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S. and Rothberg, J. M. (2000). "A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*." Nature. **403**(6770): 623-627.

Umebayashi, K., Stenmark, H. and Yoshimori, T. (2008). "Ubc4/5 and c-Cbl continue to ubiquitinate EGF receptor after internalization to facilitate polyubiquitination and degradation." Mol. Biol. Cell. **19**(8): 3454-3462.

van Wijk, S. J., de Vries, S. J., Kemmeren, P., Huang, A., Boelens, R., Bonvin, A. M. and Timmers, H. T. (2009). "A comprehensive framework of E2-RING E3 interactions of the human ubiquitin-proteasome system." Mol. Syst. Biol. **5**:295.

van Wijk, S. J. and Timmers, H. T. (2010). "The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins." FASEB J. **24**(4): 981-993.

VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M. and Wolberger, C. (2001). "Molecular insights into polyubiquitin chain assembly: crystal structure of the Mms2/Ubc13 heterodimer." Cell. **105**(6): 711-720.

Velazquez-Campoy, A., Leavitt, S. A., Freire, E. (2004). Characterization of protein-protein interactions by Isothermal Titration Calorimetry. Methods Mol. Biol. H. Fu. Totowa, NJ, Humana Press Inc. **vol. 261:** Protein-Protein Interactions: Methods and Protocols**:** 35-54.

Ventii, K. H. and Wilkinson, K. D. (2008). "Protein partners of deubiquitinating enzymes." Biochem. J. **414**(2): 161-175.

Wang, G., Yang, J. and Huibregtse, J. M. (1999). "Functional domains of the Rsp5 ubiquitin-protein ligase." Mol. Cell Biol. **19**(1): 342-352.

Wang, X., Zhang, S., Zhang, J., Huang, X., Xu, C., Wang, W., Liu, Z., Wu, J. and Shi, Y. (2010). "A large intrinsically disordered region in SKIP and its disorder-order transition induced by PPIL1 binding revealed by NMR." J. Biol. Chem. **285**(7): 4951-4963.

Waterman, H, Levkowitz, G., Alroy, I. and Yarden, Y. (1999). "The RING finger of c-Cbl mediates desensitization of the epidermal growth factor receptor." J. Biol. Chem. **274**(32): 22151-22154.

Wei, N., G. Serino, G. and Deng, X. W. (2008). "The COP9 signalosome: more than a protease." Trends Biochem. Sci. **33**(12): 592-600.

Weissman, A. M. (2001). "Themes and variations on ubiquitylation." Nat. Rev. Mol. Cell Biol. **2**(3): 169–178.

Welchman, R. L., Gordon, C. and Mayer, R. J. (2005). "Ubiquitin and ubiquitin-like proteins as multifunctional signals." Nat. Rev. Mol. Cell Biol. **6**(8): 599-609.

Wertz, I. E., O'Rourke, K. M., Zhou, H., Eby, M., Aravind, L., Seshagiri, S., Wu, P., Wiesmann, C., Baker, R., Boone, D. L., Ma, A., Koonin, E. V. and Dixit, V. M. (2004). "De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-κB signalling." Nature. **430**(7000): 694-699.

Wiener, R., Zhang, X., Wang, T. and Wolberger, C. (2012). "The mechanism of OTUB1-mediated inhibition of ubiquitination." Nature. **483**(7391): 618-622.

Wilkinson, K. (1997). "Regulation of ubiquitin-dependent processes by deubiquitinating enzymes." FASEB J. **11**(14): 1245-1256.

Winkler, G. S., Albert, T. K., Dominguez, C., Legtenberg, Y. I., Boelens, R. and Timmers, H. T. (2004). "An altered-specificity ubiquitin-conjugating enzyme/ubiquitin-protein ligase air." J. Mol. Biol. **337**(1): 157-165.

Wooten, M. W., Geetha, T., Babu, J. R., Seibenhener, M. L., Peng, J., Cox, N., Diaz-Meco, M. T. and Moscat, J. (2008). "Essential role of sequestosome 1/p62 in regulating accumulation of Lys63-ubiquitinated proteins." J. Biol. Chem. **283**(11): 6783-6789.

Xirodimas, D. P., Sundqvist, A., Nakamura, A., Shen, L., Botting, C. and Hay, R. T. (2008). "Ribosomal proteins are targets for the NEDD8 pathway." EMBO Rep. **9**(3): 280–286.

Xu, P., Duong, D. M., Seyfried, N. T., Cheng, D., Xie, Y., Robert, J., Rush, J., Hochstrasser, M., Finley, D. and Peng, J. (2009). "Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation." Cell. **137**(1): 133-145.

Yang, Z. R., Thomson, R., McNeil, P. and Esnouf, R. M. (2005). "RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins." Bioinformatics. **21**(16): 3369-337.

Ye, Y., Meyer, H. H. and Rapoport, T. A. (2001). "The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol." Nature. **414**(6864): 652-656.

Ye, Y. and Rape, M. (2009). "Building ubiquitin chains: E2 enzymes at work." Nat. Rev. Mol. Cell Biol. **10**(11): 755-764.

Zacksenhaus E. and Sheinin R. (1990). "Molecular cloning, primary structure and expression of the human X linked A1S9 gene cDNA which complements the ts A1S9 mouse L cell defect in DNA replication." EMBO J. **9**(9): 2923-2929.

Zacksenhaus E., Sheinin R. and Wang, H. S. (1990). "Localization of the human A1S9 gene complementing the ts A1S9 mouse L-cell defect in DNA replication and cell cycle progression to Xp11.2----p11.4." Cytogenet. Cell Genet. **53**(1):20-2.

Zhang, X. C., Chen, J., Su, C. H., Yang, H. Y. and Lee, M. H. (2008). "Roles for CSN5 in control of p53/MDM2 activities." J. Cell Biochem. **103**(4): 1219-1230.

Zheng, N., Wang, P., Jeffrey, P. D. and Pavletich, N. P. (2000). "Structure of a c-Cbl-UbcH7 Complex: RING domain function in ubiquitin-protein ligases." Cell. **102**(4): 533-539.

Zhao, C., Beaudenon, S. L., Kelley, M. L., Waddell, M. B., Yuan, W., Schulman, B. A., Huibregtse, J. M. and Krug, R. M. (2004). "The UbcH8 ubiquitin E2 enzyme is also the E2 enzyme for ISG15, an IFN-α/β-induced ubiquitin-like protein." Proc. Natl. Acad. Sci. USA. **101**(20): 7578-7582.