University of Nottingham

UK | CHINA | MALAYSIA

Analysis of substrate selectivity determinants in Ubiquitin Specific Proteases 11 and 15

Matthias P. Mayer

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

March 2023

Word Count: 24,469

Abstract

Ubiquitin Specific Proteases (USPs) have long been studied for various physiological effects, being involved in many cellular processes as well as many different cancers. The variety of the USP family as well as the large variation in substrates has made them the subject of drug trials in the past. However, as one USP can target many different substrate in cells, more work should be done to analyse what directs this substrate specificity.

This investigation aims to elucidate which amino acids direct the substrate specificity of USPs.

USP11, a poor cleaver of linear diubiquitin, was analysed using sequence analysis and a previous structure to identify amino acids which affect substrate specific activity. This aspartic acid residue proximal to the catalytic histidine (USP11D886) was shown to affect cleavage of linear diubiquitin, with a glycine in this position drastically increasing cleavage rates. The reverse mutation in USP15 (USP15G860D) demonstrated complementary results. However, two binding forms leads to a complex analysis of the interaction through isothermal titration calorimetry (ITC) and microscale thermophoresis (MST). An X-ray crystallography derived structure with novel protein tags to aid crystallographic study shows evidence of multiple binding sites between USP11 and diubiquitin. This showcases a promising target for substrate specific activity, with the proximal aspartic acid/glycine residue being a druggable target for specific interactions.

COVID-19 Impact Statement

During March 2020, the coronavirus pandemic affected non-essential lab work throughout the United Kingdom. This included the closure of the labs used for the work in this thesis for a time of six months. This led to severe disruption of the first and second year research. When the labs reopened, due to safety requirements there was a maximum lab occupancy below half the number of users. This led to a part time/shift pattern work schedule for the following six months. This in effect led to a total 9 months of lost time for completing this work. The time after the global pandemic also affected the supply lines of plastic ware and reagents, leading to difficulties in performing lab work. Additional problems incurred includes accessing facilities and expertise both within and without Nottingham, i.e. the nanomicroscale research centre at the University of Nottingham for previously planned electron microscopy (EM/cryoEM) experiments, and MST/SPR experiments at Manchester Universities' Biomolecular analysis platform. Before the pandemic, significant time was spent on training for electron microscopy analysis, but after the shutdown, lack of access to facilities even once the labs had reopened meant this was no longer feasible.

Mitigations for these issues included a three month funded extension through the University of Nottingham which recovered time taken by a mandatory three month industrial student placement scheme. Consumables with a long lead time were bulk ordered to prevent this from delaying work. However, the main mitigation of coronavirus factors involved reframing the scope of the project. Limiting experiments to equipment and expertise accessible techniques,

focussing more time and effort on research which was more likely to yield results such as only utilising protein sequences which could be readily produced with large yields. Additionally assessing the purity needs of proteins and occasionally allowing for less stringent purification to reduce time requirements. Additionally, structure determination was focussed on USPs with pre-existing structures to showcase the feasibility of otherwise highly time intensive structural techniques.

Acknowledgments

First and foremost my supervisor Dr Ingrid Dreveny and her right hand Dr Sigrun Maurer, if not for their ceaseless effort, constant supervision, expert tutelage, and inspired contribution this work would not be possible.

Additionally, Dr Sigrun Maurer and Dr Simon Caulton performed much previous work, including identification of novel protein tags. The entire structural biology group at Nottingham contributed immensely, including Hyojung Kim, Thomas Warwick, William Richardson, Alexander Brown, Phillip Smith, Monika Pathak, Louise O'Hara, Catherine Harrison, Zoe Markham-lee, Hannah Mbiwan, Jiatong Zhang, Jed Allen, Yewei Ma, Reyes Ruiz Campoy, and Craig Charles for his funk and soul show, the soundtrack to my thesis.

I would like to acknowledge the funding received through the BBSRC, the University of Nottingham as my host institution, and the administrators of my course who assisted this work as much as governor Inalchuq contributed to the future of the Khwarazmian Empire.

The entire team of Diamond Light Source, their funders, and especially the operators of beamline I24 without whom the structural data would not be possible. Other contributions include the nanomicroscale research centre in Nottingham, the Midlands Regional Cryo-Electron Microscope Facility, the Manchester University Biomolecular analysis platform, Dr Lodewijk Dekker.

Personally, to my mother who brought me into Biology as a field, my family and friends as a whole whom I have shamefully ignored and neglected when attempting this work. Most importantly, my partner Dr Lauren Smith, without her constant support and wise guidance this work would not have been possible.

CONTENTS

List of Abbreviations

USP – Ubiquitin Specific Protease

PDB- Protein Data Bank

Ub - Ubiquitin

M – Methionine

G – Glycine

ITC – Isothermal Titration Calorimetry

MST – Microscale Thermophoresis

WT – Wildtype

FL – Full Length

cat – Catalytic domain construct

LB – Luria Broth

CCP4 - Collaborative Computational Project Number 4

D1D2 – Catalytic Domain 1 and Catalytic Domain 2

BL1 – Blocking Loop 1

BL2 – Blocking Loop 2

SL – Switching Loop

CCL – Catalytic Cleft Loop

LMB - Laboratory of Molecular Biology

DUB – Deubiquitinase

Ub-AMC – Ubiquitin 7-Amino-4-methylcoumarin

MJD – Machado-Joseph disease protease

UCH – Ubiquitin Carboxy-terminal hydrolases

JAMM - JAB1 MPN Mov34 metalloenzymes

MINDY - motif-interacting novel DUB family proteases

ZUFSP - Zinc finger with UFM1-specific peptidase domain protein

List of Tables

List of Figures

List of Appendices

1. *Chapter 1: Introduction*

1.1 Ubiquitin

Ubiquitin is a 76 amino acid protein (Swatek & Komander, 2016) which is ubiquitously present across eukaryotes (Cappadocia & Lima, 2018), with proteins exhibiting ubiquitin like folds being present in some bacteria and archaea (Burroughs, Balaji, Iyer, & Aravind, 2007). The importance of ubiquitin was uncovered when ubiquitination of substrates led to protein degradation in the proteasome (Coux, Tanaka, & Goldberg, 1996). Since then ubiquitin and the ubiquitin system has been shown to affect cell cycle progression, transcriptional regulation, signal transduction, the immune response, receptor regulation, endocytosis and others (Hershko & Ciechanover, 1998).

The action of ubiquitin binding is a multi-step ATP dependent process where (most commonly) the C-terminal glycine 76 residue is bonded to a target lysine substrate, although cysteine, threonine, and serine residue bonds have also been reported (Scaglione, et al., 2013). This occurs through three classes of enzymes, the ubiquitin activating E1 enzymes form a thioester linkage with ubiquitin and carries out ATP-AMP exchange reactions (Ciechanover, Elias, Heller, & Hershko, 1982). The activated ubiquitin is then transferred to the E2 ubiquitin conjugating enzyme trough a similar thioester linkage (Zhang & Sidhu, 2014). E3 enzymes recruit the E2-ubiquitin complex, recognise target substrates and catalyse the ubiquitin ligation to the target residue (most commonly a lysine) (Bernassola, Karin, Ciechanover, & Melino, 2008). (shown in figure 1.1).

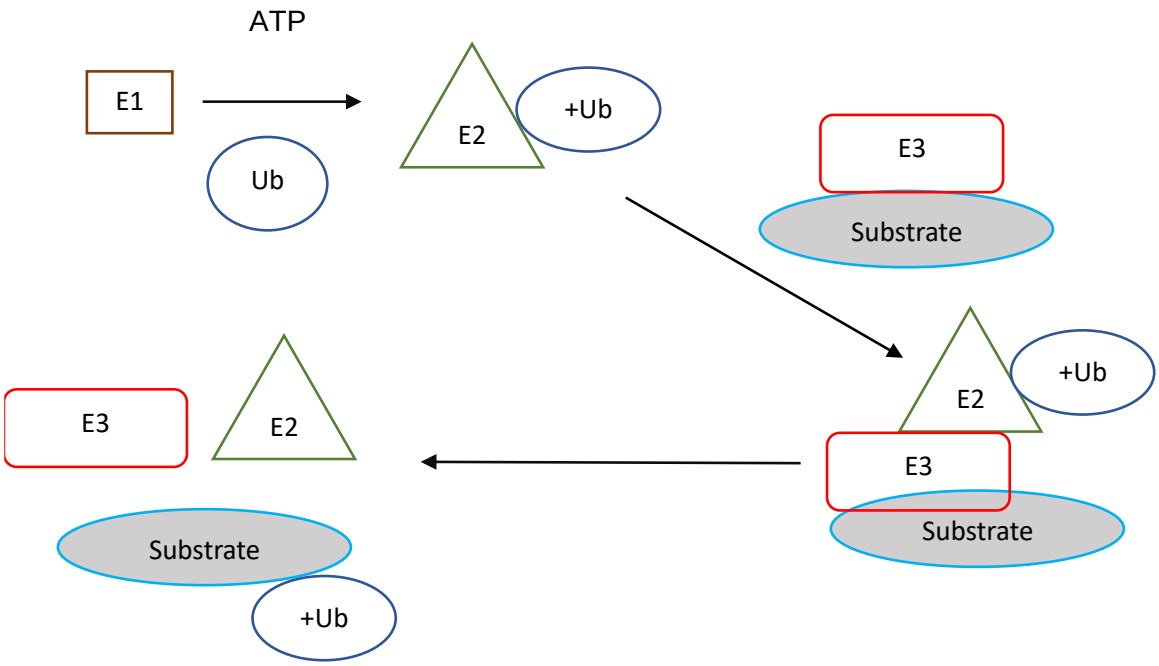

Figure 1.1, Representation of ubiquitin binding through E1, E2, and E3 enzymes. This indicates how E1 ubiquitin activation enzymes are conjugated to ubiquitin through an ATP dependent process. E1 enzymes then confer the ubiquitin to the E2 conjugation enzymes. The E2 enzymes then associate with the E3 enzymes which conjugate with the substrate target. The E3 ubiquitin ligases then transfer the activated ubiquitin to the substrate. The E2 and E3 proteins disassociate from the complex (Bernassola, Karin, Ciechanover, & Melino, 2008).

There are different members of the E1, E2, and E3 cascade families. With two confirmed E1 enzymes, approximately 40 E2 members, and somewhere between 600-1000 E3 ubiquitin conjugating enzymes. These different enzymes coordinate ubiquitination and importantly substrate specificity for this ubiquitination (Bui, Hong, Kwak, Lee, & Lee, 2021). Ubiquitin is itself susceptible to ubiquitination, containing seven documented ubiquitination sites (K6, K11, K27, K29, K33, K48, and K63) (Xu & Peng, 2008) shown in figure 1.2. This can lead to chains of ubiquitin which can alter the effects on target proteins dependent on the ubiquitin chains attached to them (Castañeda, et al., 1993). Mono-ubiquitinated proteins can have affects such as down regulation of transcription, such as with histone modifications (Oss-Ronen, Sarusi, & Cohen, 2022). Poly-ubiquitinated K48 chains are closely associated with degradation of the target substrate through the proteasome degradation pathway, whereas K63 chains have been linked to signalling pathways such as endocytosis, inflammation signalling, and DNA repair (Bello, et al., 2022). This lends itself to the importance of substrate specificity of particularly E3 ubiquitin ligases when directing ubiquitin ubiquitination, which accounts for the large number of members in the E3 ligase family (Smith, et al., 2019).

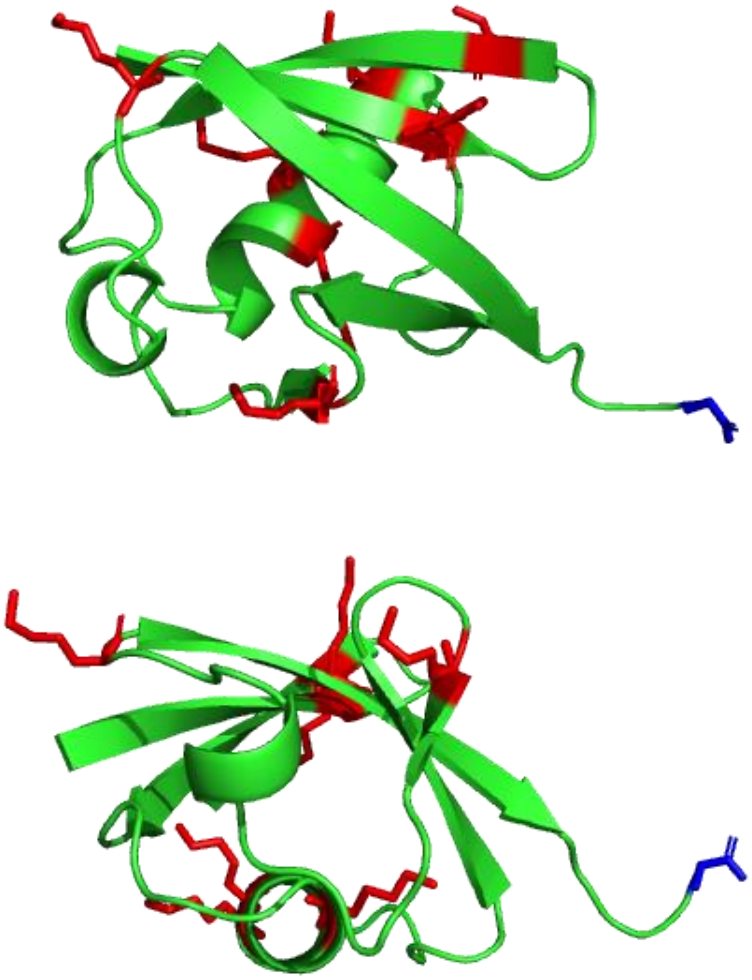

Figure 1.2 Structure of Ubiquitin (PDB-ID: 1UBQ. The structure of ubiquitin shows a five-stranded β-sheet opposite one α-helix with a short 3_{-10} helix in between. Ubiqiutin shows a hydrophobic core with lysine residues (highlighted in red) which can themselves be ubiquitinated but the C-terminal glycine residue (blue). This shows the possible variety of ubiquitin chains through the potential for six lysines to be ubiquitinated (K6, K11, K27, K29, K33, K48, and K63) (Reyes-Turcu & Wilkinson, 2009) (Vijay-Kumar, Bugg, & Cook, 1987).

1.2 Deubiquitinases

With ubiquitination requiring a large suite of proteins, there exists a superfamily of proteins which reverses this action, called deubiquitinases (Lange, Armstrong, & Kulathu, 2022). This balance between deubiquitination and ubiquitination has been deemed essential for eukaryotic homeostasis, especially with regards to protein cycling, DNA transcription and cell cycle progression (March & Farrona, 2018). Deubiquitinases exhibit high substrate specificity, both with regards to ubiquitinated protein, and the type of ubiquitination (i.e. different ubiquitin chains) (Clague, et al., 2013). There are at present, roughly 100 deubiquitinating enzymes, which have been separated into seven distinct families. These are called ubiquitin specific proteases (USPs), ovarian tumour proteases (OTUs), Machado-Joseph disease proteases (MJDs), ubiquitin carboxy-terminal hydrolases (UCHs), JAB1 MPN Mov34 metalloenzymes (JAMM), motif-interacting novel DUB family proteases (MINDY) (Maurer & Wertz, 2016), and the most recently defined Zinc finger with UFM1-specific peptidase domain protein (ZUFSP) (Kwasna, et al., 2018). JAMM enzymes are zinc metalloproteases, with the remainder of DUBs being cysteine proteases (Harrigan, Jacq, Martin, & Jackson, 2018). These DUBs have distinct characteristics and are diverse in their structure and substrate targets. ZUFSP localises to DNA lesions where it selectively cleaves K63 polyubiqiutin (Kwasna, et al., 2018). MINDY proteins are highly selective for K48 linked polyubiqiutin chains, cleaving from the distal end (Abdul Rehman, et al., 2016). JAMM metalloenzymes exhibit target promiscuity (relative to other DUBs) but many require the association of multisubunit complexes for complete activity (Cao, et al., 2017). UCH enzymes were investigated as a family which

cleaves small molecules at the c-terminus of ubiquitin (Wilkinson, 1995) (Qiu, et al., 2006). MJD deubiquitinases contain the highly conserved Josephin domain, and the dysfunction of MJDs is linked to the titular Mochado-Joseph disease, most commonly through MJD ataxin-3 (Zeng, et al., 2020). Most OTU proteases are involved in important signalling cascades and exhibit strong ubiquitin chain specificity with dysregulation implicated in tumour progression (Mevissen, et al., 2013). USPs are the largest of the DUB families with at least 58 members, with a wide variety of substrate and ubiquitin chain specificities (Sun, Shi, Mamun, & Gao, 2020).

1.3 Ubiquitin Specific Proteases

USPs share a catalytic core USP domain, which has been described as containing three regions known as the thumb, fingers, and palm regions (figure 1.2) (Yuan, et al., 2018). This catalytic core domain itself has some degree of variation within it, and ranges in size between USPs to be between 295 and 850 residues in length (Ye, Scheel, Hofmann, & Komander, 2009). Whilst the full length USP sequences range in size from 330 to 3500 amino acids in length with a variety of different domains potentially present such as zinc finger domains, ubiquitin like domains (UBL), domain present in USP (DUSP) domains, and others (Nijman, et al., 2005). There are also different target substrates for USPs between members of the family. For example, USP5 and USP13 have been shown to target the linkages in unattached ubiquitin chains for ubiquitin cycling within cells (Reyes, Shanks, Komander, & Wilkinson, 2008), whereas USP9X releases ubiquitin chains from the substrates via internal cleavage (Al-Hakim, et al., 2008).

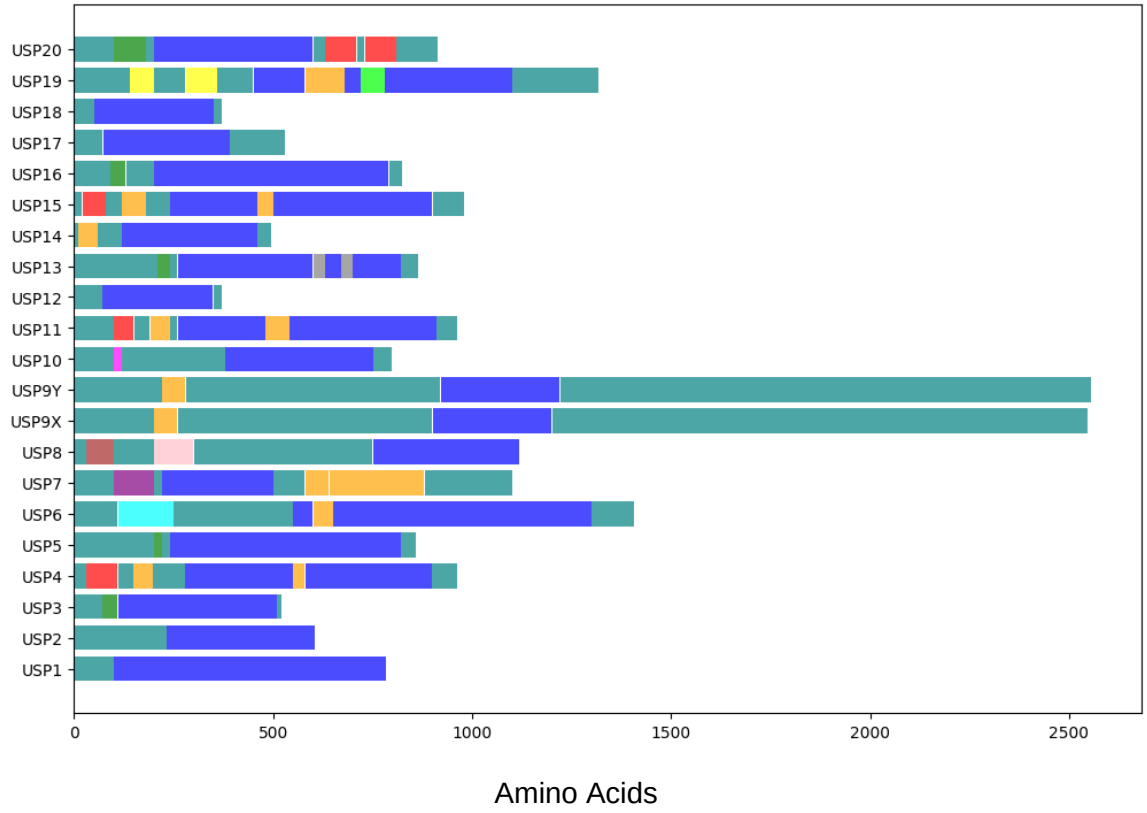

Figure 1.3, Representation of USP domain boundaries. This shows twenty one USP proteins with their domains represented in a colour coded format. The domains are as follows, USP catalytic domain = blue, Zinc finger ubiquitinated-specific protease domain(ZnF-UBP) = green, domain present in ubiquitin specific proteases (DUSP) = red, ubiquitin-like domain (UBL)= orange, MATH domain = purple, microtubule interaction and trafficking domain (MIT) = brown, Rhodanese domain = pink, Rab GTPase activating protein TBC domain (TBC-RABGAP) = cyan, Ataxin-2C = magenta, CS-domain= teal, ZnF-MYND= lime, and non domain sequences = light blue. Amino acid number is on the x-axis (Komander, Clague, & Urbe, 2009).

This substrate variation and interaction with near ubiquitious post translational protein modification has led USPs to be involved in a wide variety of cellular processes such as USP44s involvement in the cell cycle (Stegmeier, et al., 2007), USP1s involvement in DNA repair (Oestergaard, et al., 2007), or USP9Xs involvement in TGFβ signalling and embryonic development (Dupont, et al., 2009). The variety of USPs is shown by a representative 21 USPs in figure 1.3. With a wide variety of substrates, the USP family also is involved with many disease states. This includes USP22s effect on *F.tularensis* proliferation (Akimana, Al-Khodor, & & Abu Kwaik, 2010), or the effect of USP16 on insulin and diabetes (Balaji, Pokrzywa, & Hoppe, 2018). However, USPs have mostly been investigated with their varied and complicated relationship with cancers. This includes a wide range of lung, breast, carcinoma, liver, colorectal, oesophageal, pancreatic and more cancers (Harrigan, Jacq, Martin, & Jackson, 2018). Aside from gene based knockdown or analysis studies (Koulich, Li, & DeMartino, 2008), USPs and cancer has also been investigated through structural biology and the designing of inhibitory drug molecules (Bonacci & Emanuele, 2020).

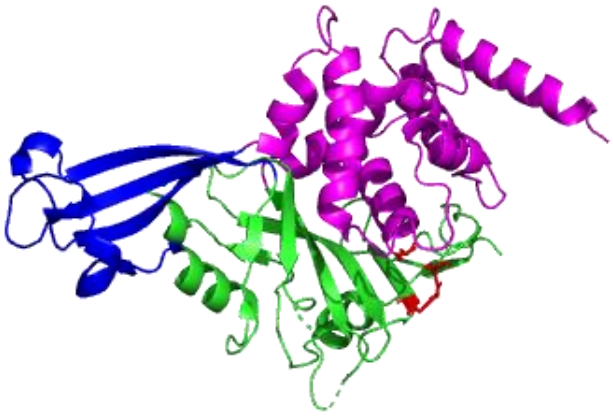

Figure 1.4, Structure of USP7 (PDB-ID:1NB8) showing the typical USP fold structure. This structure of USP7 highlights the three catalytic triad residues C223, H464, and D481 (red). Along with the “finger”, “palm”, and “thumb” domain typical of USP catalytic domains (blue, green, and magenta respectively) (Kim, van Dijk, & Sixma, 2016) (Hu, et al., 2002).

1.3.1 USP7 as a historically drugable target

Whilst USP7 mutations were found in patients showing neurodevelopmental disorders (Fountain, et al., 2019). USP7 was identified as a drug-able target when it was linked to cellular levels of p53, a tumour suppressor which is affected in a majority of tumour types (Hu, 2006). This was confirmed when elevated USP7 levels negatively affected prognosis in leukaemia (Carrà, et al., 2017), prostate cancer (Song, et al., 2008), glioblastoma, and others (Cheng, Niu, Yang, Wang, & Lu, 2013). USP7 is a 135 kDa seven domain protein consisting of a TRAF-like (Tumor necrosis factor Receptor–Associated Factor) domain, followed by the USP catalytic core domain (figure 1.4), and then five ubiquitin like domains (Pozhidaeva & Bezsonova, 2019). Analysis of the USP7 structure, using structure based hit discovery through molecular docking and hit to lead analysis, led to several inhibitors being developed (Schauer, Magin, Liu, Doherty, & Buhrlage, 2020). These include some of the earliest compounds “P5091” (Chauhan, et al., 2012) and “HBX-41108” (Colland, et al., 2009). Whilst this was followed by high throughput screening studies, with similar USPs as controls for enhanced selectivity, led to additional inhibitors being developed such as GNE-6776, which showed high selectivity if low activity (Kategaya, et al., 2017). Investigation is currently ongoing for these USP7 inhibitors including upcoming clinical trials, this case of USP7 demonstrates the viability of structure based drug design for USP selective small molecule inhibitors, and the potential this represents for cancer therapeutics (Oliveira, Guedes, & Salvador, 2022).

1.3.2 USP15

USP15 is a protease linked to several cellular signalling pathways including the COP9-singalosome (Hetfeld, et al., 2005), TGF-β (Eichhorn, et al., 2012), NF-κB (Schweitzer, Bozko, Dubiel, & Naumann, 2007), β-catenin (Huang, Langelotz, Hetfeld-Pechoc, Schwenk, & Dubiel, 2009), and p53 (Zou, et al., 2014). All of which are pathways heavily associated with oncogenic activity (Chou, et al., 2017). USP15 has also been investigated with its USP specific antagonistic activity with parkin mediated mitophagy in extreme cases of Parkinson's disease (Cornelissen, et al., 2014), and for its role in the innate immune response through regulating TRIM-25-RIG-1 antiviral signalling (Pauli, et al., 2014).

USP15 exhibits a multi domain structure containing an N-terminal DUSP-UBL domain (Faesen, Luna-Vargas, & Sixma, The role of UBL domains in ubiquitin-specific proteases, 2012), with the catalytic core separated by an additional UBL insertion, this identifies the catalytic core domains 1 and 2 (D1D2) as amino acids 294 – 484 (D1) and 775 – 963 (D2). The catalytic triad residues have been identified as C269, H862, and D879 (Ward, et al., 2018) (figure 1.5).

With the catalytic USP core domain being split in USP15, it has previously been investigated with relation to two other USPs which exhibit a similar catalytic domain insertion UBL domain structure. These being USP4 and USP11 with a 56.9% and 43.2% sequence homology respectively (Vlasschaert, Xia,

Coulombe, & Gray, 2015). With regards to further analysis, USP11 and USP15 share some targets but not others, with binding targets in p53 and RNA splicing pathways being selective between them (Sowa, Bennett, Gygi, & Harper, 2009). This represents an opportunity for potentially identifying key differences in otherwise similar structures that direct this substrate specificity.

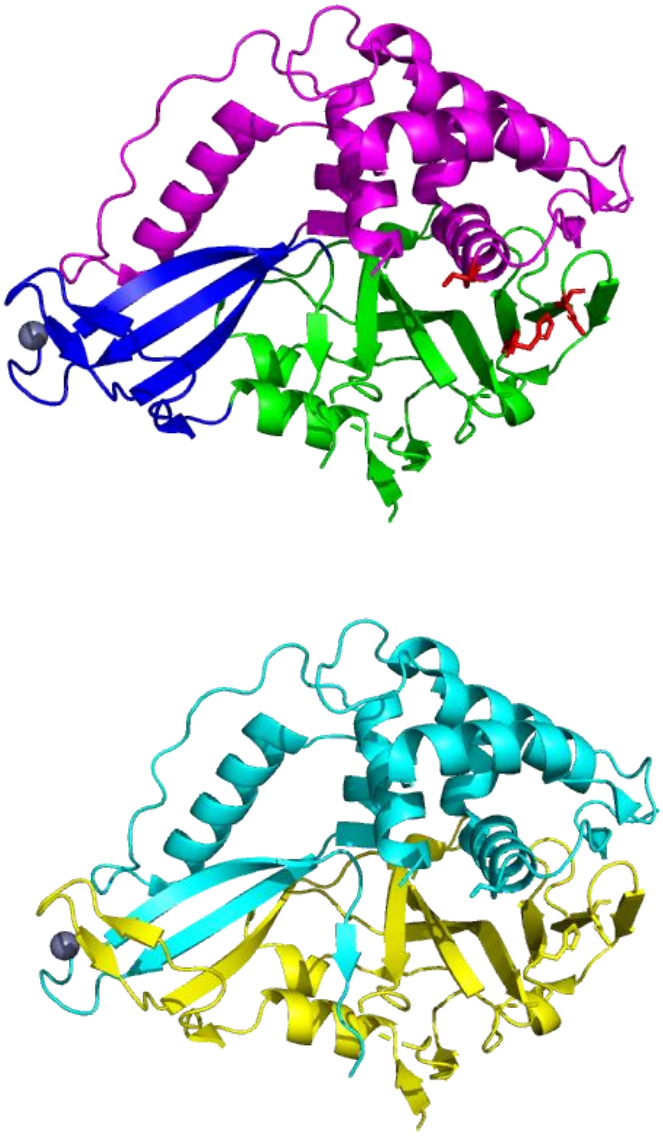

Figure 1.5, Structure of USP15 catalytic core (PDB-ID: 6GHA). The top structure of USP15 shows the catalytic triad C269, H862, D879 (red) with the finger (blue), palm (green), and thumb (magenta) domains. The structure below shows the USP15 catalytic core split into the domains which are usually separated by the ubiquitin like (UBL) insertion domain. This was removed for crystallographic study, which shows the first half of the catalytic domain (D1) in cyan with the second half of the catalytic domain (D2) shown in yellow (Ward, et al., 2018).

1.3.3 USP11

First investigated for its role in cell cycle regulation through interaction with RanBPM and the formation of microtubules (Ideguchi, et al., 2002). Other binding partners were discovered including BRCA2 impacting DNA repair and breast cancer (Schoenfeld, Apgar, Dolios, Wang, & Aaronson, 2004), IAP2 which inhibits apoptosis and affects melanoma (Lee, et al., 2015), H2AX which is essential for double-strand break DNA repair (Yu, et al., 2016), and others impacting the TGFβ and TNFα pathways (Guo, Tang, Yuan, Zhang, & Wang, 2022).

USP11, along with USP4 and USP15, contains an N-terminal DUSP-UBL domain, which has been investigated for the potential to direct substrate specificity (Elliott, et al., 2011). The structures of this DUSP domain has been elucidated for these three proteins (figure 1.6). Other similarities include the USP catalytic domain being separated in USP11 into two catalytic domains (D1/D2) with an inserted ubiquitin like domain (UBL) domain between them. This leads the catalytic domain of USP11 to be amino acids 242 – 445 (D1) and 736 – 931 (D2) with the catalytic triad residues being C318, H888, and D906 (Harper, et al., 2014).

Some substrates of USP11 are still targeted with a USP11 N-terminal deletion, such as the VGLL4 transcriptional regulator (Zhang, et al., 2016). This indicates that there are other components directing USP11 substrate specificity for targets.

Common substrates can be used to determine the effects of different factors on USP function, such as ubiquitin-AMC. This has been used to assess inhibitors and mutations for a variety of USPs (Gjonaj, et al., 2019). Other common substrates have also been used which are cleaved by USPs with drastically different affinities. One such substrate is linear diubiquitin, which has been shown to be a poor substrate for cleavage by USP11 (Faesen, et al., 2011). USP15, despite having a similar sequence structure, has a relatively high affinity for linear diubiquitin cleavage. This identifies an easily assessable substrate which highlights the differences between these two similar USPs (Ward, et al., 2018). USP11 having been shown to cleave K63 linked ubiquitin chains confirms the ability for USP11 to interact with multimers of ubiquitin, as opposed to USP11 being specialised for monoubiquitin interaction (Liao, Zhou, Wang, Yang, & Jiang, 2022).

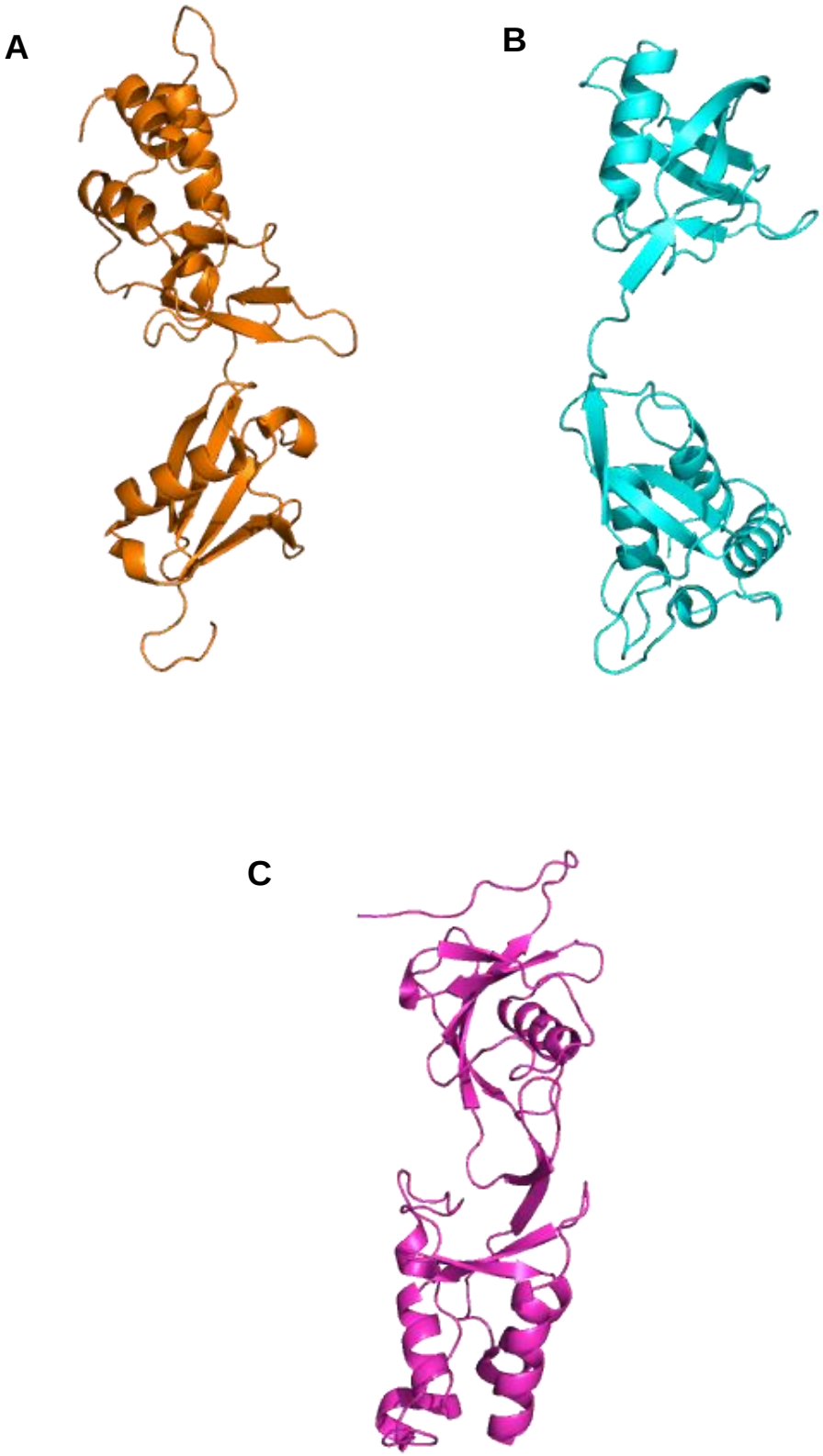

Figure 1.6, Structure of DUSP domains, USP4, 11, and 15. A), shows the USP4 DUSP structure (orange, PDB ID: 3JYU), B) shows the USP11 DUSP structure (cyan, PDB ID: 4MEL), C) shows USP15 DUSP structure (purple, PDB ID: 3T9L). (Bacik, et al., 2009) (Harper, et al., 2014) (Harper, 2011)

1.3.4 USP17

Originally researched in mice as deubiquitinating enzymes (DUB family) (Zhu, Pless, Inhorn, B., & D'Andrea, 1996), the human analogue called USP17 was identified (Burrows, McGrattan, & Johnston, 2005). USP17 actually consists of a small family of USP17 like proteins which can contain as much as an 8 % difference in sequence similarity (Burrows, Scott, & Johnston, 2010). However, differences in expression patterns and specificities have not been identified (Ducker & Shaw, 2021). This paper, and others like it, do not make a distinction between the USP17 like proteins.

A distinction between USP17 and the previous USPs mentioned is that USP17 does not have the DUSP domain, and the catalytic core of USP17 is not separated into two domains with an insertion UBL domains, the catalytic core containing amino acids 72-390, with the catalytic triad C89, H334, and D350 (Ducker & Shaw, 2021) (Hjortland & Mesecar, 2016). The only other identified domain in USP17 (aside from the USP catalytic domain) are two hyaluronan binding motifs, which also have been implicated in binding RNA (Shin, Yoo, Kim, Kim, & Baek, 2006).

USP17 was found to be expressed transiently in response to cytokines, with constitutive expression shown to block cell proliferation, to the point of inducing apoptosis (Burrows J. , et al., 2004). USP17 transcription occurs mostly during G2/M phase transition and late G1 stage of the cell cycle, with high transcript levels detected in the pancreas, heart, and liver (Shin, Yoo, Kim, Kim, & Baek,

2006) (McFarlane, et al., 2010). However one study of the human protein atlas indicated low tissue specificity in USP17 production (Uhlén, et al., 2015).

Complementary to expression pattern of USP17, it has been investigated in relation to the cell cycle. Such as the investigation showing USP17 deubiquitinates and stabilises CDC25a, which is vital for G1/S cell cycle progression (Pereg, et al., 2010). USP17 has been shown to interact with other cell cycle targets such as cyclinA (Hu, et al., 2019), and SETD8 (Li, et al., 2022). USP17 has other targets influencing innate immunity/inflammation, such as RIG-1 and MDA-5 (Chen, et al., 2010), and cell migration such as SLUG and TWIST (Lin, et al., 2017).

With the relation between USP17 and the cell cycle, cell migration, inflammation, and apoptosis, USP17 has been investigated for its role in tumorigenesis. This shows USP17 differentially effects different cancer phenotypes (Ducker & Shaw, 2021). USP17 is present in higher levels (compared to non-carcinogenic expression) for breast cancer (Shaw, et al., 2012) and lung cancer tissue (Lu, et al., 2018). However USP17 was shown to be downregulated in glioma (Hu, et al., 2016). USP17 has also been shown to affect pathogenesis, whereby USP17 levels correlated with non-small-cell lung carcinoma disease recurrence at distant sites (McFarlane, et al., 2013). Breast cancer metastasis was similarly effects by USP17 knockdown/rescue experiments (Liu, et al., 2017). This highlights the importance of substrate specificity, where complete knockdown of USP17 activity could have off target effects.

Despite the difference in sequence and domain architecture between USP17 and the USPs with the catalytic domain separated by a UBL domain, there are some similarities between them. For example, USP17 shares the USP11 poor ability to cleave linear diubiquitin (Hjortland & Mesecar, 2016). In this way, the comparison of USP11 with USP15 and USP17 could identify key amino acids directing the substrate specificity differences between these proteins. With USP11 and USP15 having a similar sequence but different substrate activity, USP11 and USP17 having a dissimilar sequence but similar substrate activity.

1.4 Biophysical Techniques

1.4.1 Isothermal Titration Calorimetry (ITC)

ITC is a technique which analyses the thermodynamics of the interaction between subjects, commonly soluble protein samples with ligands (Baranauskiene, Kuo, Chen, & Matulis, 2019). This requires protein samples with a high purity and quality, referring to homogeneity and structural conformity (Lamanna, et al., 2018). Advantages of ITC include the absence of labelling, immobilisation, chemical modification, optical considerations, or molecular weight limits for the samples used. But this must be balanced with the large protein requirement and a high water solubility (Reimund, Kovrov, Olivecrona, & Lookene, 2017).

ITC can be used to determine a range of different constants and parameters including enthalpy change (ΔH), binding constant (K_b), dissociation constant (K_d), or entropy change (ΔS), and by extension the Gibbs energy of a reaction (ΔG). ITC additionally identifies the stoichiometry of reactions through binding molarities. This highlights ITC as an information rich but high resource use technique (Falconer, 2016).

The technique studies the biomolecular interaction through addition of the biomolecule of interest into a metallic calorimetric sample cell, a companion cell (filled commonly with H_2O) is used to compare differences in temperature between these two cells (Freyer & Lewis, 2008). Both cells use heaters with a constant reference power, and sensors to detect changes in temperature, the sensors fuel a feedback circuit engaging the heaters. The graph is generated through the power required to maintain an equal temperature between the cells. The stepwise injection of a second molecule interacting with the first causes changes in the in the thermodynamics of the sample cell, which eventually returns to equilibration with the baseline (Buurma & Haq, 2007). The deviation from and resumption of this baseline input power, combined with the concentrations of the reactants, is used to generate the reaction constants (Velazquez-Campoy, Leavitt, & Freire, 2004). This requires high accuracy for reactant concentrations, optimisation for concentrations and injections which yield a full curve, careful loading of the sample cell, and control experiments to identify errant heat effects (Pierce, Raman, & Nall, 1999). ITC also gives information about binding stoichiometry including number of binding sites (Duff, Grubbs, & Howell, 2011). But more detailed information about binding stoichiometry either requires complementary techniques such as SPR

(Krishnamoorthy, Alluvada, Hameed Mohammed Sherieff, Kwa, & Krishnamoorthy, 2019) or currently developing analysis techniques (Erlekam, Zumbansen, & Weber, 2022).

1.4.2 Microscale Thermophoresis (MST)

MST is a technique relying on the movement molecules undergo through microscopic temperature gradients, called thermophoresis (or thermodiffusion) (Duhr & Braun, 2006). As different titrations of a substrate are added to a target molecule, MST can detect the changes in the molecules' thermophoretic shift (Wienken, Baaske, Rothbauer, Braun, & Duhr, 2010). This is attributed to the Soret effect, which is dependent on the size, charge, and surface properties of a molecule. These characteristics change enough when undergoing a binding event that the outcome can be measured through MST (Wienken, Baaske, Duhr, & Braun, 2011).

MST is measured through the UV absorbance which sometimes utilises the fluorescence of aromatic amino acids, but most often a fluorophore is tagged to the molecule of interest eliminating interference common with untagged MST (Sparks, Lawless, Arango, Tajkhorshid, & Fratti, 2022). As an example using the Monolith NT.115 nanotemper system, 16 samples are made with an increasing titre of substrate. The base absorbance is measured, then an infrared laser creates a localised temperature gradient, decreasing the absorbance read as the particles migrate from the heat source (Sparks & Fratti, 2019). The chance in fluorescence signal varies with increasing substrate,

allowing MST to quantify the interaction and determine constants such as the K_d (Bartoschik, et al., 2018).

The measurement of fluorescence can also be varied. The initial sudden change of the fluorophore properties with heating, less than one second, is regarded as the temperature jump (T-jump). This is influenced by local surroundings of the fluorophore, whereas the thermophoresis signal occurs over the seconds to minute timescale and is influenced by global properties of the entire complex (Jerabek-Willemsen, et al., 2014). These can be used to infer mechanistic information in certain instances, such as in the case of EcoSSB exhibiting 2:1 binding stoichiometry with a 35 oligonucleotide sequence. This was the case where the T-jump signal exhibited a typical sigmoidal curve but the thermophoresis signal contained a central peak with a 1:10 ratio of EcoSSB to DNA before levelling off at the 1:1 molar ratio (Bujalowski & Lohman, 1989). However, most MST analysis is based on the assumption of a 1:1 binding model, with some specialist analysis being performed on different binding parameters, including the Hill equation. But these methods only account for one inflection point, leaving a lot of work still to come in determining a widely accepted standard for other binding stoichiometry and their analysis with MST (Tso, et al., 2018).

1.4.3 X-Ray Crystallography

X-ray crystallography is a well-known structural biology technique with a long history, starting with the discovery of X-rays in 1895, the diffraction of X-rays

through a crystal lattice in 1912, to the first biological structure of myoglobin in 1957 (Shi, 2014). The repository of X-ray structures had 171,000 depositions as of 2022 (Protein Data Bank, 2023).

The technique requires pure protein samples to be concentrated into a crystalline lattice, which remains the main bottleneck of generating X-ray crystallographic data (Pardon, et al., 2014). After this has been achieved often requiring numerous optimisations of a myriad of conditions (precipitant, concentration, buffer, pH, drop diffusion method, temperature, additives) (Thakur, et al., 2007), the crystal is commonly frozen in liquid nitrogen and exposed to a high powered single wavelength beam of X-rays. This commonly requires a specialist facility called a synchrotron to generate these X-ray beams (Grabowski, et al., 2021). This outputs a diffraction image, commonly this is then combined with many diffraction images with the diffraction spots being indexed. Collecting this data requires making sure all reflections of the asymmetric unit are collected. The asymmetric unit represents the base repeating unit of the crystal lattice. Collecting this data requires rotating the crystal to collect a range of diffraction images during a single exposure to the high energy beam (Dauter, 2017). The intensity of these spots are measured (Otwinowski & Minor, 1997). The intensity of these spots is a result from the amplitude of waves used to construct them and the phase difference between them when they are diffracted into the detector. The amplitude can be calculated with relative ease (Collaborative Computational Project, Number 4, 1994), counter to what has often been termed the phase problem which must be calculated indirectly (Barbarin-Bocahu & Graille, 2022).

A common method for solving the phase problem is molecular replacement. This requires a high homology structure to be available. This works by using the structure factors and phase information from a similar structure and using them to calculate the new factors and phase information. Molecular replacement involves fitting the similar structure to the experimental unit cell to calculate test phases (Ilari & Sevano, 2008). This requires iterative refinement in an attempt to reduce previous model bias and create accurate data (Rossmann, 1990).

This combination of spot amplitude and phase information can be utilised through Bragg's Law (Bragg, 1913)and fast Fourier transformation (FFT) (Ten Eyck, 1973)to calculate the electron density map. The electron density is a three dimensional map in which a model of the target protein can be fitted (Jones, Zou, Cowan, & Kjeldgaard, 1991). This usually requires additional rounds of refinement and the quality of the model can be assessed through the R factor which compares the calculated model structure factors with the observed structure factors. This eventually yields the structures deposited on the protein data bank (Arnold & Rossmann, 1986). Most of the complicated mathematics can be accomplished through a suite of software programs packaged into CCP4 or Phenix (Krojer, et al., 2017).

1.5 Previous Work

This investigation built upon work previously undertaken. This includes work performed identifying novel protein tags to aid in expression, stability, and crystallisation of proteins of interest. Work performed by Caulton used literature

searches as well as analysis of the protein data base (PDB) and identified several novel protein tags (Caulton, 2016).

These tags included dimerisation tags, such as the PDB ID 3WVA. Identified as HcgF from the bacterium *Methanocaldococcus jannaschii*. This protein is used in the biosynthesis of an iron-guanylylpyridinol cofactor (Fujishiro, Kahnt, Ermler, & Shima, 2015). This tag was investigated for its ability to strongly dimerise, when attached to the N or C-terminus of a target protein, this would lead to the dimerisation of the entire complex. This was performed with a view to aid in crystal packing, electron microscopy, stabilisation, and protein production (figure 1.5) (Maurer S. , 2021).

Additional tags investigated include PDB IDs 1W41 and 2GKG. 1W41 is the structure of ribosomal protein L30e from the bacterium *Thermococcus celer*. This ribosomal protein was investigated for its high thermal stability, where mutations of this sequence led to a reduction in thermal degradation temperature (Lee, Allen, Bycroft, & Wong, 2005). The 2GKG structure is the receiver domain of FrsZ, a motility protein from *Myxococcus xanthus* .The FrsZ protein locates within the cell at poles to direct social motility. Mutations at points in the 2GKG structure prevents this pole localisation and inhibits cell motility (Fraser, et al., 2007). These two proteins were used as tags with a view to stabilise flexible loop regions (figure 1.5). This was investigated with a view to aid crystallisation for reduced flexibility, increase protein production, and increase protein stability (Caulton, 2016). Another proposed tag was the 1BKR structure, a calponin homology domain, which acts as the actin binding site of cytoskeleton proteins (Bañuelos, Saraste, & Djinović Carugo, 1998). This was used to increase protein stability and aid in protein production (Caulton, 2016).

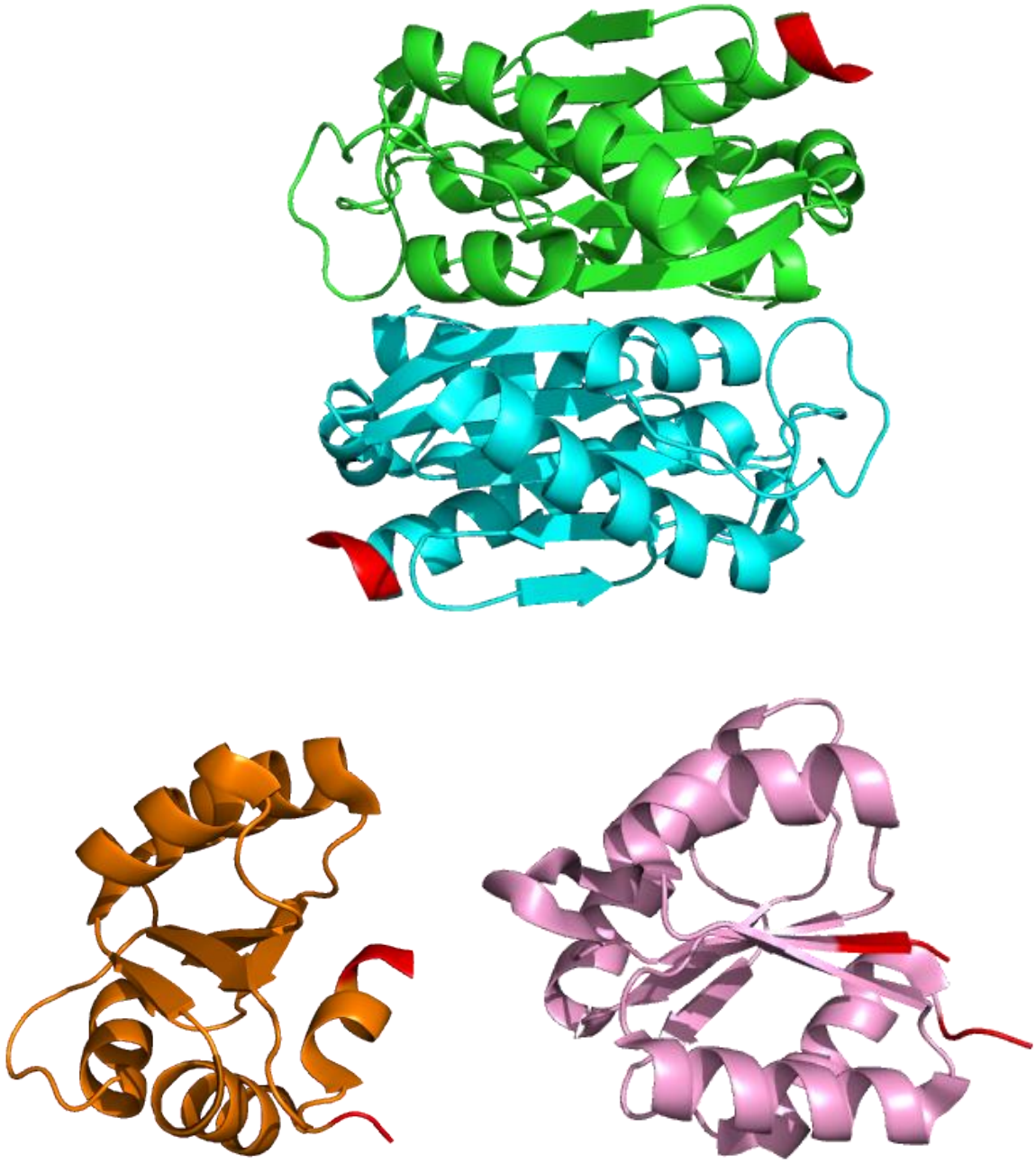

Figure 1.5, Structures of proteins used as peptide tags. The top structure shows a 3WVA dimer (green and blue), with red highlighting the amino acids used to tag the protein of interest. The two structures on the bottom show 1W41 (left, orange) and 2GKG (right, pink) which use both the C and N terminus residues (red) which are mutated into a protein of interest in an attempt to stabilise flexible loop regions (Fraser, et al., 2007) (Fujishiro, Kahnt, Ermler, & Shima, 2015) (Lee, Allen, Bycroft, & Wong, 2005) (Caulton, 2016) (Maurer S. , 2021).

Additional work previously completed heavily impacted the direction of this investigation. A structure was elucidated of a USP11-2GKG catalytic domain construct bound to monoubiquitin with 3 additional glycine residues (ubiquitinGGG) (figure 1.6). The proximity of amino acids to the ubiquitin structure led to an investigation into which of these amino acids may be directing the complex substrate specificity known to USPs. This led to an analysis of these loop regions and their sequence. This was performed to identify links between these loops and substrate specific activity. USP11 was compared to the similar USP structures USP4, and USP15 which had much higher diubiquitin cleavage activity (Faesen, et al., 2011). Additionally, these loop regions were compared to a lower sequence identity candidate USP17, this was performed due to USP17 being previously identified as a similarly poor cleaver of linear diubiquitin compared to USP11 (Ducker & Shaw, 2021). These sequence comparisons in figure 1.7 identify areas in which USP11 differs from the similar D1D2 USPs. However, USP11 D886 was highlighted as a prime candidate for further investigation as the aspartic acid residue was only present in the poor cleavers of linear diubiqiutin (USP11 and USP17). Conversely the higher affinity cleavers of linear diubiquitin contained a glycine residue in this location (two amino acids before the catalytic Histidine residue in these sequences).

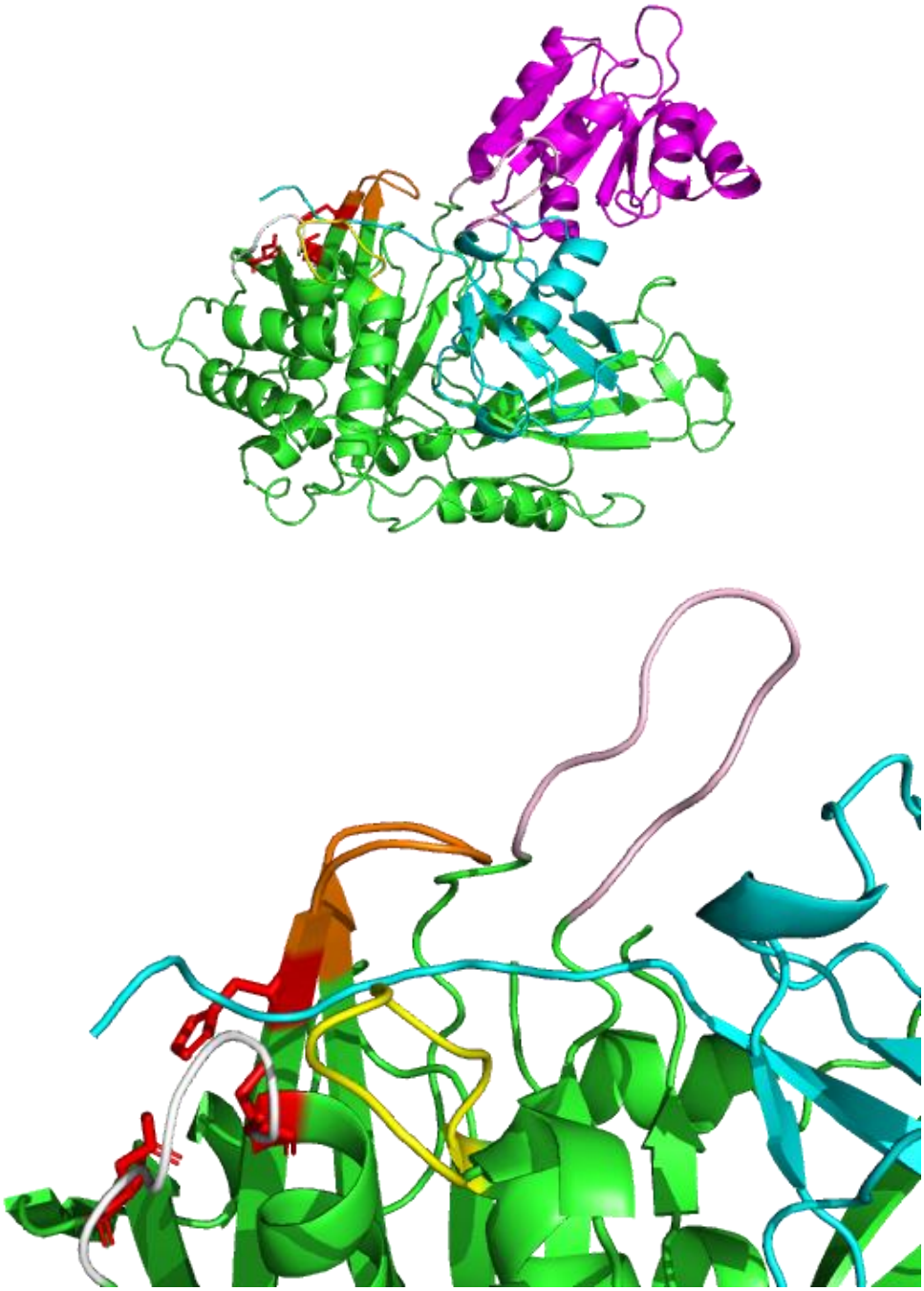

Figure 1.6, USP11 bound to monoubiquitin. The top structure shows a catalytically inactive USP11 D1D2 mutant (green) bound to ubiquitin (blue). This includes the 2GKG protein tag (purple) cloned into the USP11 construct to stabilise a flexible loop region. The bottom structure shows a focussed view of the catalytic triad residues (red) around the c-terminal glycine residue of the bound ubiquitin (blue). Also focussed are four different loop regions, identified as the catalytic cleft loop (CCL, white), the switching loop (SL, yellow), the blocking loop 1 (BL1, pink), and the second blocking loop (BL2, orange). These regions were considered for the potential to effect substrate binding.

		CCL							
USP11	**312**	**T**	**N**	**L**	**G**	**N**	**T**	**C**	**F**
USP4	**305**	G	N	L	G	N	T	C	F
USP15	**263**	S	N	L	G	N	T	C	F
USP17	**83**	Q	N	M	G	N	T	C	Y

		SL									
USP11	**391**	**S**	**Q**	**F**	**L**	**G**	**Y**	**Q**	**Q**	**H**	**D**
USP4	**384**	P	Q	F	S	G	Y	Q	Q	Q	D
USP15	**342**	P	Q	F	S	G	Y	Q	Q	Q	D
USP17	**151**	A	G	F	H	R	G	K	Q	E	D

		BL1							
USP11	**831**	**F**	**S**	**Y**	**T**	**K**	**F**	**S**	**R**
USP4	**828**	F	S	Y	N	R	Y	W	R
USP15	**809**	F	S	Y	S	R	Y	M	R
USP17	**283**	F	S	D	V	T	G	N	K

		BL2							
USP11	**882**	**G**	**G**	**M**	**R**	**D**	**G**	**H**	**Y**
USP4	**875**	G	A	M	G	V	G	H	Y
USP15	**856**	G	G	M	G	G	G	H	Y
USP17	**328**	W	S	C	H	D	G	H	Y

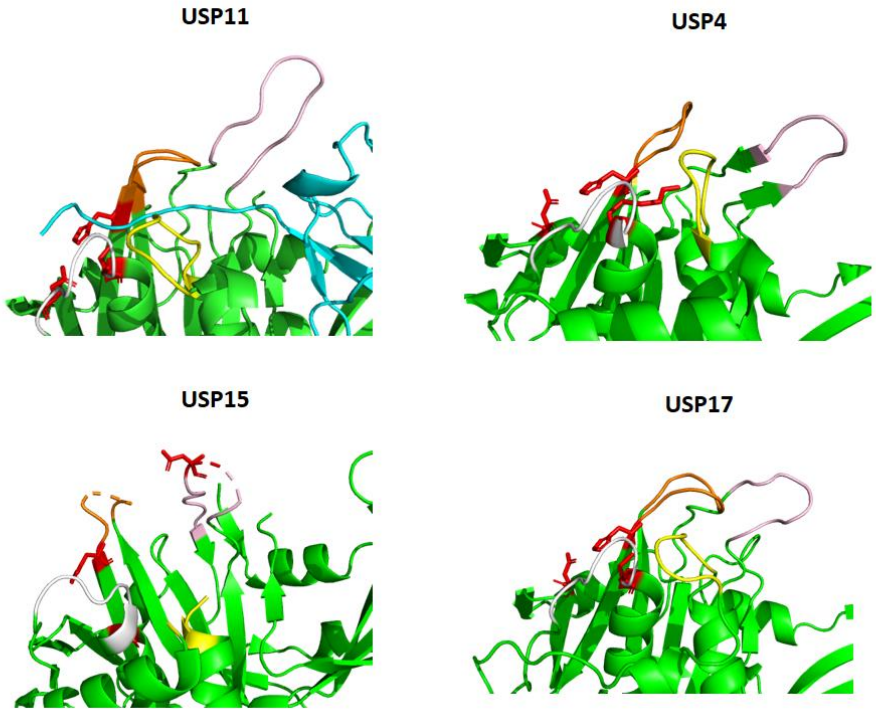

Figure 1.7, Sequence alignment, USPs of loop regions. Previously identified loop regions, thought to effect substrate binding, were compared between USPs of interest. The loops colour coded from the sequence alignment (top) to the structures of USPs (bottom), CCL (white), SL (yellow), BL1 (pink), BL2 (orange). This shows the USP11 (structure to be deposited, UniProt P51784), USP4 (PDB ID: 2Y6E, UniProt Q13107), USP15 (PDB ID: 6GHA, UniProt Q9Y4E8-2), and USP17 (AlphaFold predicted structure, UniProt Q6R6M4) sequences and structures. This comparison is used to identify potential mutation targets.

1.6 Project Aims

The general aims of this project are to identify which amino acids close to the catalytic triad are capable of affecting substrate specificity for USPs. This involved

- Using sequence analysis to identify amino acids with the potential to direct substrate specificity concerning differences in diubiquitin cleavage ability with USP11 and USP15, investigating this with fluorescence based monoubiquitin and gel based diubiquitin assays.
- Further characterise these interactions using quantitative analytical techniques ITC and MST to generate binding constants
- Utilise X-ray crystallography to generate a co-crystallisation structure of USP11 mutants bound to linear diubiquitin to demonstrate this mutation and the importance of this amino acid for substrate specificity.

2. Chapter 2: Materials and Methods

2.1 Sequences used

To keep data comparable and consistent to previous literature, USP constructs used were the catalytic domains. In the case of USP11 (UniProt P51784) and USP15 (Q9Y4E8-2) this involved combining the D1 and D2 domains. The USP11 construct without mutations was bounded by the sequence 252-455, 736-839, this construct is later referred to as USP11 wildtype (USP11wt), or throughout this paper simply as USP11 with a molecular weight of 42.46 kDa. The USP15 construct without mutations was bounded by the D1 and D2 domains, sequence 255-440, 756-919, this construct was referred to throughout as USP15 D1D2 or USP15 wildtype (USP15wt), or simply USP15 with a molecular weight of 41 kDa. The USP17 (UniProt Q6R6M4) sequence was used both full length (1 – 530) at 59.9 kDa, but where stated the catalytic domain was used (72 – 390) at 35.5 kDa (USP17cat). Ubiquitin constructs (UniProt P0CG48) for monoubiquitin used the full ubiquitin chain length (1 – 76) at 8.5 kDa. DiUbiquitin constructs were expressed with two sequences (i.e, 1-152) at 17.1 kDa. Where the diUbiquitin deletion was used, this labelled as the L149X mutant refers to the second ubiquitin in the chain, i.e. in a monoubiquitin construct, this would refer to L73X, removing the terminal RGG sequence from the ubiquitin chain. Monoubiquitin substrates with additional amino acids, i.e. Ub-M had the additional amino acids added to the C-terminal residue, making the sequence X77M. Additional tags used where stated also impacted the

molecular weight observed in gels, notably for USP11_2GKG and USP17cat1BKR constructs with 53.0 kDa and 48.9 kDa respectively.

2.1.1 Primers and cloning

The full length USP17 construct was ordered from GenScript in the GST tagged plasmid pGEX1 for USP17fl-GST tagged expression. Restriction enzymes Nde1 and Xba1 were used to clone the sequence into the his-tag plasmid pCOLD1 for his-tag purification. This used the forward primer GGAATTCCATATGGAGGACGACTCACTCTAC, and the reverse primer CTAGTCTAGATTACTGGCACACAAGCAGAGCCC. The USP17 catalytic domain sequence was ordered from GenScript and restriction enzyme cloning was used to input the sequence into a pRSF-13 plasmid which was pre-cloned with the PDB-ID 1BKR. The restriction enzymes used were Not1 and Xho1. The primers forward primer used was ATAAGAATGCGGCCGCAAGCAGGAGACCTGCTGCGG, and the reverse primer was CCGCTCGAGTTATTCCCTGCCTCTTGACACACTC. The USP17 constructs USP17(fl)1W41, and USP17cat3WVA were created using restriction free cloning. The primers used are as follows, 1W41 forward: ACCTGGACATCGCCCTGGATATCCAGTCTATGGTTGATTTTGCTTTCGAA CTCCG, 1W41 reverse: TGTTCCAAAGCTTGCTGGACACTCTGTGCGCCCAACGCCAG, 3WVA forward: CGCTTCCTGGAAATCCTGAAAAAAGAATACTTCGCCGCAAGTAGCAGGA

GACCTGCTG, 3WVA reverse: GCTTTGTTAGCAGCCGGATCTCATTCCTCTTCTTCCTCTTCTTCCCTGCC TCTTGACACAC. The USP17_2GKG construct was ordered from GenScript in the plasmid pET21-d. USP11 constructs were as previously stated all the D1D2 catalytic domain constructs. Originally ordered from GenScript in the plasmid pET26-b. These constructs were subject to numerous mutations utilising a quick-change mutagenesis protocol (below). The mutants and primers used are as follows. C318S forward: CAATCTGGGCAACACGTCCTTCATGAACTCGGCCCTG, C318S reverse: CAGGGCCGAGTTCATGAAGGACGTGTTGCCCAGATTG, H399Q forward: GGCTACCAGCAGCAAGACTCTCAGGAGCT, H399Q reverse: AGCTCCTGAGAGTCTTGCTGCTGGTAGCC, R885G forward: CCATTATGGGGGCATGGGTGATGGACACTACAC, R885G reverse: GTGTAGTGTCCATCACCCATGCCCCCATAATGG, D886G forward: CATTATGGGGGCATGCGTGGTGGACACTACACAACATTTGC, D886G reverse: GCAAATGTTGTGTAGTGTCCACCACGCATGCCCCCATAATG. In the case of multiple mutants, quick change mutagenesis was first run for one set of primers, then the protocol was repeated for another set of primers (after SourceBioscience sequencing confirmed the first mutant was present). USP11 was cloned with PDB-ID 2GKG with restriction free cloning utilising the primers forward tag sequence using the following primers forward: CCTTCTGCTACCTCAGTGTTCCACTGCCGGGTGCGAAAAAAATTCTGATT GTGGAAAG and reverse: CTCAATGCACTCCTGCAGCCGCACCGGCGGAAAGCCAATCAGCGCG.

The USP15D1D2 construct, originally ordered from GenScript in plasmid

pET21-d underwent quick change mutagenesis to achieve the G860D mutant. This used the primers forward: CTATGGAGGGATGGGAGACGGACACTATACTGCTTTTG, and reverse: CAAAAGCAGTATAGTGTCCGTCTCCCATCCCTCCATAG.

Ubiquitin constructs, ordered from GenScript (both mono 1-79, and diubiquitin 1-149) were cloned into pCDF-DUET-1 using Nde1 and Kpn1 restriction digest protocols. This used the primers forward: GGGAATTCCATATGCAGATCTTCGTGAAGACTC, and reverse: CGGGGTACCTCAACCACCACCCCCACCTCTGAGACGGAG. This also extended to ubiquitin which had the sequence extended by 1 glycine (ubi-G), 3 glycines (ub-GGG), and a methionine (Ubi-Met). This also is applied to the diubiquitin L149X mutant.

2.1.2 PCR mutation and amplification

For quick change mutagenesis, a mix was created using 0.5 µL of 100 ng/µL original plasmid, 5 µL of 10 × PFU Ultra reaction buffer, 0.5 µL of forward primer (125 ng/µL), 0.5 µL reverse primer (125 ng/µL), 1 µL DNTP mix, 1 µL 50% (v/v) DMSO, 40.5 µL dH_2O, and 1 µL of PFU Ultra II HF DNA Polymerase. This mix was added to a PCR machine with the following cycle, intial denaturation 95 °C for 5 minutes, denaturation 95 °C 30 seconds, annealing 55 °C 1 minute, extension 72 °C 8 minutes, the denaturation, annealing, and extension steps were repeated til a total of 38 cycles had been completed, then the final 72 °C

extension step was performed for 15 minutes. The PCR machine then cooled the samples to 4 °C til they were retrieved. The mix was then digested of previous original plasmid with the addition of 1 µL Dpn1 enzyme before incubation at 37 °C for three hours. After this time heat shock transformation and sequencing performed by Source Bioscience was performed to confirm presence of mutation in the plasmid. Restriction enzyme digest was performed in several steps, first by the creation of two separate mixes. Both contained 1 µL of each of the two chosen restriction enzymes (as above), 2 µL of 10 × cutsmart buffer (New England Biolabs), 8 µL of DNA (one mix containing the insert to be cloned in, the second mix containing the host plasmid), and 8 µL dH_2O. This was left overnight at room temperature. Self ligation was prevented with the addition of Antarctic phosphatase for 1 hour 30 minutes room temperature before heat inactivation at 37 °C for 30 minutes. After this, 1 µL of T4 Ligase buffer (New England Biolabs) was added to 6 µL of the insert DNA, with 2 µL of the plasmid DNA and 1 µL of T4 ligase. This was left to incubate overnight at 16 °C before being heat shock transformed into nova blue cells. Presence of correct insert size was confirmed by agarose gel electrophoresis and sequencing (Source Bioscience). Restriction free cloning was performed in a two-step process. The first step is a mix using 0.5 ng/µL of the original plasmid sequence, 5 µL of 10 × PFU Ultra reaction buffer, 0.5 µL of forward primer (125 ng/µL), 0.5 µL reverse primer (125 ng/µL), 1 µL DNTP mix, 1 µL 50% (v/v) DMSO, 40.5 µL dH_2O, and 1 µL of PFU Ultra II HF DNA Polymerase. This was added to a PCR machine with 95 °C for 15 minutes as an initial denaturation, then denaturation (95 °C for 30 seconds), annealing (55 °C for 30 seconds), and extension (72 °C for 2 minutes) was repeated 40 × before a final extension

at 72 °C for 15 minutes. The PCR machine then cooled the sample to 4 °C. The mix was run through a PCR prep kit (QIAGEN, Cat. No. / ID: 28104) to purify the “megaprimers” used in restriction free cloning. These megaprimers were added to a mix, consisting of 0.5 µL of 100 ng/µL megaprimer, 0.5 ng/µL of 100ng/µL target plasmid sequence, 5 µL of 10 × PFU Ultra reaction buffer, 1 µL DNTP mix, 1 µL 50% (v/v) DMSO, 40.5 µL dH_2O, and 1 µL of PFU Ultra II HF DNA Polymerase. This was then run in a PCR machine under the following conditions: 95 °C for 15 minutes as an initial denaturation, then denaturation (95 °C for 30 seconds), annealing (55 °C for 30 seconds), and extension (72 °C for 8 minutes) was repeated 40 × before a final extension at 72 °C for 15 minutes. The PCR machine then cooled the sample to 4 °C. At the finish of the protocols, all samples were cleaned using the Qiagen PCR cleanup kit. DNA concentration was assessed through addition of 1 µL of sample to a blanked Thermofisher Nanodrop Spectrophotometer

2.2 Recombinant Protein Production

2.2.1 Heat Shock Transformation

2 µL of 100ng/µL plasmid DNA was added to 100µL of competent cell mix under Bunsen aseptic conditions (BL21 DE3 Codon Plus cells for expression, Nova Blue cells for plasmid replication). This mix was kept on ice for 15 minutes. The mix was then heated to 42 °C for 45 seconds. The mix was cooled on ice for 10 minutes before 500 µL of LB media was added to the mix under Bunsen aseptic conditions. This was then incubated at 37 °C with 180 rpm shaking in an

incubator for 1 hour. Next to a Bunsen burner, 200 µL of this mix was spread on an LB agar plate with a working concentration of the antibiotic corresponding to the resistance in the plasmid. This was left in a static incubator at 37 °C overnight and assessed for colony growth the next day (~18 hours). Exceptions to this are the co-expressions between the USP11 construct and linear diubiquitin, whereby the protocol was kept the same except 1 µL of 100ng/µL of each plasmid were combined and mixed first, before 2 µL of this plasmid mix was added to competent cells. The rest of the protocol was kept the same except for 300 µL of the cell mix was spread on the agar plate after incubation. Plasmid preparation from nova blue competent cells was prepared according to the Qiagen plasmid mini prep kit (Cat. No. / ID: 27104).

2.2.2 Bacterial Recombinant Protein Expression

All work performed next to a Bunsen burner. 100 mL of LB media in a 250 mL flask was inoculated with a working concentration of antibiotics dictated by plasmid resistance. The flask was then inoculated with cells using a pipette tip to scrape either a colony from an agar plate, or cells from a glycerol stock. This was left overnight shaking at 180 rpm in a 37 °C incubator. Cells were retrieved and 8 mL of overnight culture was added to 800 mL of 2×YT media (for USP constructs) or 800 mL of LB media (for ubiquitin constructs) in a 2.5L flask with appropriate antibiotics at working concentrations, this was repeated twice for 3 × 800mL flasks (2.4L total). The flasks were then incubated at 37°C with 180

rpm shaking (for smooth flasks) or 120 rpm shaking (for baffled flasks). The OD was measured after the first 2 hours, then again until the cells reached an OD_{600} of 0.6 (OD_{600} 1.2 for p.Cold1 constructs). The flasks were then inoculated with 400 µL of 1 M IPTG each (final concentration 0.5 mM IPTG). The temperature was reduced to 18 °C and the cells were left incubated overnight (12 °C and left for 2 days for pCold1 constructs).

The next day (~18 hours) the flasks were retrieved from the incubator and centrifuged at 4250 r.c.f. (g), for 30 minutes at 4 °C, in 600 mL containers (4 × 600 mL = 2.4L total). These cells once pelleted were resuspended in ~30 mL of purification buffer (histrap buffer A for USP constructs, Cation Exchange buffer A for ubiquitin constructs). The resuspended cell pellet was centrifuged again at 4250 r.c.f. for 30 minutes to collect the cell pellet from each 2.4L into one falcon tube which was stored at -20 °C until used for purification.

Modifications to this protocol include media changes to autoinduction media (Studier, 2005) where stated, which contains 6 g Na_2HPO_4, 3 g KH_2PO_4, 20 g tyrptone, 5 g yeast extract, 5 g NaCl, 10 mL of 60 % v/v glycerol, 5 mL of 10 % w/v glucose, 25 mL of 8 % w/v lactose into 1 L of H_2O. This was incubated with overnight cultures and grown at 18 °C for 26 hours before harvesting.

2.2.3 Affinity/Cation exchange/size exclusion chromatography purification

Histrap purification buffer A: 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM Imidazole, 5 % v/v Glycerol. Histrap purification buffer B: 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM Imidazole.Cation exchange buffer A: 20 mM

ammonium acetate pH 5.1. Cation Exchange buffer B: 20 mM ammonium acetate pH 5.1, 1 M NaCl. Gel filtration buffer: 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% v/v Glycerol.

All buffers were degassed via attaching to a pump and were filtered through a 0.22 µm filter. This was then left with a magnet stirrer agitating the buffer and the vacuum pump degassing the buffer for at least 1 hour.

Cell pellets, retrieved from -20 °C freezer, were resuspended in 30mL purification buffer A (histrap purification buffer A for USP constructs, cation exchange buffer A for ubiquitin constructs). Cells were then sonicated for a total of 8 minutes (20 seconds on, 20 seconds off) on ice. Lysed cells were then centrifuged at 28,000 r.c.f for 60 minutes at 4 °C. The ubiquitin constructs had the cell lysate then had the pH reduced with 1M acetic acid to 4.2 for precipitation of contaminant proteins, the pH was then raised with 1M NH4OH to 4.8 to prevent further precipitation. Ubiquitin constructs were then centrifuged again at 28,000 rcf for 60 minutes at 4 °C. The lysate was then syringe filtered through a 0.45 micron filter before being loaded onto a prewashed column with buffer A. The column was then transferred to an Akta system (pure or start Cytiva) where it was washed with 5 column volumes of buffer A. A gradient was setup and 2.5 mL fractions were collected as this was run over the column, 1% buffer B increase per mL. These fractions were analysed for purity on SDS-PAGE gels before the appropriate fractions were pooled and concentrated for subsequent purification or experimentation.

Purity assessed by SDS-PAGE gels led to many constructs needing size exclusion gel filtration purification. This was performed by concentrating the

desired fractions to between 2-3 mL in volume and being loaded into a 5 mL loop. This was then loaded and eluted from a pre-equilibrated 16/600 column (superdex200 for USP constructs, superdex 75 for ubiquitin constructs) with gel filtration buffer (50 mM Tris-HCl pH7.5, 150 mM NaCl, 1 % v/v glycerol). 2 mL fractions were collected after 40mL (void volume) had been run through the column. These fractions were assessed for purity using SDS-PAGE analysis and were then concentrated and used for future experimentation, sometimes requiring snap freezing in liquid nitrogen for longer term storage at -20 °C (no longer than 6 months).

2.2.4 Protein Concentration

Protein concentration was performed using Sartorius protein concentrators (Vivaspin centrifugal concentrator). 30 kDa cutoff concentrators were used for USP constructs (minimum MW of USP constructs = 42160). 10 kDa cutoff concentrators were used for diubiquitin constructs, with 5 kDa MW cutoff concentrators used for monoubiquitin constructs. The concentrators were first rinsed with distilled H_2O, then were spun for 1 minute at 3,000 r.c.f. with 10 mL Gel filtration buffer. The concentrators were then loaded with the fractions (varying volume dependent on purification analysis) and spun at 4,000 r.c.f. for 10 minutes at 4 °C. After this the concentrator was assessed for volume and then the concentrator was respun at the same conditions until the volume decreased to a desired point (~ 2.5 mL for gel filtration, < 1 mL dependent on downstream experimental requirements). After every spin, the supernatant was resuspended via a pipette to prevent collection of material precipitating at the

filter of the concentrator. Final concentration was assessed through addition of 1 µL of sample to a blanked thermofisher nanodrop spectrophotometer.

2.3 SDS-PAGE Analysis

Gels were prepared in sets of 4 plates simultaneously using the Biorad Mini-PROTEAN short gel plates and gel casette. Plates were collected in the clamps with the shorter plates facing forwards, the large plates behind with the raised edges between the two. These clamps and plates were put in the housing on top of a sponge material to prevent leakage. The gel was then prepared, all gels had 18 % v/v acrylamide final concentration and were prepared by adding liquid in the following order. First the resolving gel was prepared by adding 9.6 mL of 30% (v/v) acrylamide (Sigma Aldrich), 4 mL 1.5 M Tris-HCl pH 8.8 (pre prepared stock), 2.06 mL H_2O, 160 µL 10% (w/v) SDS (pre prepared stock), 160 µL 1% w/v APS (pre prepared stock kept at – 20 °C for storage), 16 µL TEMED (tertramethylethyldiamine**,** Sigma Aldrich). The TEMED and APS solutions were added immediately before mixing and applying to the gel plates, these plates were filled to roughly between 3/4 and 4/5 of the total capacity with resolving gel solution. The gels then had 100 % Isopropanol poured gently on top of them to draw out excess liquid, prevent air bubble formation, and ensure a flat interface between the resolving gel and the future stacking gel. This was left to set for a minimum of 30 minutes. After which the ispropanol was gently rinsed off the gel surface by applying H_2O and pouring off repeatedly, (5-7 times) before drying the edges of the plate. The stacking gel (6 % v/v acrylamide

final concentration) was then prepared and applied to the gels. This was made by combining 2 mL 30 % (v/v) acrylamide, 2.5 mL 0.5 M Tris-HCl pH 6.8, 5.3 mL H_2O, 100 µL 10 % (w/v) SDS (pre prepared), 100 µL 1% w/v APS (pre prepared), 10 µL TEMED (tertramethylethyldiamine**,** Sigma aldrich). The TEMED and APS solutions were added immediately before mixing and applying to the gel plates. The plates were filled to a maximum capacity with stacking gel before 15 well combs (SIGMA) were added to the tops of the gels, taking care to not introduce bubbles between the combs and the stacking gel. This was left to set for at least 30 minutes before the combs were removed, the gel plates were removed and kept at 4 °C wrapped in a damp condition for future use (no longer than 2 weeks storage).

All gels were run at 180V constant voltage for 60 minutes using the biorad PowerPac™ basic power supply. The gels were then removed from the power packs and gel plates before being washed with 20 mL distilled H_2O and heated in a microwave to remove excess SDS. The gels then had the H_2O removed and 20 mL of Coomassie stain (made up to 1L with 4 mL 37% (v/v) HCl (~12M) and 2 g Coomassie brilliant blue R powder (thermo-scientific 20278). This was left staining for 30 minutes before the coomassie stain was removed and the gel was destained by heating H_2O for 30 seconds in the microwave and left agitated on the gel for a further 30 minutes.

2.3.1 Agarose Gel Electrophoresis

1 g of Agarose (Merck) was heated and diluted with 100mL 1 × TAE buffer (50 mM Tris, 20 mM acetic acid, 1 mM EDTA). This was left to cool to below 50 °C

and poured into a gel casing with a comb in place (sigma Aldrich). Before this was solidified 1 µL of SYBR safe stain (Thermo Fisher) was added and mixed. The DNA sample is mixed in a 5:1 ratio with a 6 × gel loading buffer (30 % v/v glycerol, 0.25 % w/v Bromophenol blue, 0.25 % w/v Xylene cyanol FF dye). When the gel has fully set, the DNA is added 5 µL to each well before being run at 80 V for 40 minutes. Imaging is performed on a VWR geldoc using blue light for illumination. Samples were compared to a ladder run simultaneously in the gel, using the 1Kb ladder from ThermoFisher Scientific.

2.4 Analytical Methods

2.4.1 Ubiquitin AMC assays

Conjugated Ubiquitin-AMC was purchased from Ubiquigent and 2 µL aliquots of 50 mg/mL were prepared and stored at -80 °C. For assays, one aliquot would be prepared with a final in well concentration of 750 nM (unless otherwise stated) with target 75 nM USP (unless otherwise stated) in a reaction buffer (AMC reaction buffer**:** 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM DTT). The ubiquitin-AMC was only added to wells immediately before plate reading was performed using a Perkin Elmer Envision plate reader. Arbitrary fluorescence units were read with an excitation wavelength of 380 nM (Fura2 BFP) and a read wavelength of 460 nM (Umbelliferone). Negative and positive controls used where necessary (previously identified functional cleaving USP proteins as positive controls, heat deanatured protein/absent protein as negative

controls). Reads were performed every minute for a total of 90 reads. This was kept at a constant temperature of 25 °C. Experiments were performed in triplicate and an average of reads was taken.

2.4.2 Isothermal Titration Calorimetry (ITC)

ITC was performed using the Malvern Panalytical MicroCal PEAQ-ITC. The ITC reaction buffer was kept constant as 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 % (v/v) glycerol with both the syringe (ubiquitin constructs) and cell (USP constructs) peptides being dialysed overnight using gentle stirring at 4 °C in a cold room into the same buffer. Dialysis membrane used was the 3.5 kDa limit mini dialysers (mini Slide-A-Lyzer ThermoFisher Scientific).

Unless otherwise stated (as determined by different parameters needed) the experimental conditions were as follows: Cell protein concentration 30 µM, syringe protein concentration 300 µM, 23 total injections, 0.2 µL initial injection followed by 1.7 µL injections, 160 seconds spacing between injections, constant temperature 25 °C, reference power 5 µcal/s. Unless otherwise stated, dissociation constants were determined from 3 repeats of the same conditions for each experiment. Spent reactions were discarded. Results were performed and analysed using the Malvern MicroCal PEAQ-ITC analysis software. Experiments were performed in triplicate with results shown being an average of these.

2.4.3 Microscale Thermophoresis (MST)

Microscale Thermophoresis was performed on the Monolith NT.119 ™ from NanoTemper. The labelling kit used was the nanotemper RED-tris-NTA kit, a hexahistidine tag to specifically bind 1:1 to histidine tagged proteins. This involved creating 2 µL aliquots of concentrating labelling reagent (Red-Tris-NTA) and mixing with 98 µL of aliquoted buffer reagent (PBS-T Buffer). This was mixed with protein diluted to 200 nM (in PBS buffer) in a 1:1 ratio to give a final concentration of 100 nM protein. This was left at room temperature for 30 minutes to allow binding of labelling molecule before the mix was centrifuged at 13,000 r.c.f. for 10 minutes at 4 °C. A dilution series (in PBS buffer) was created of target ubiquitin construct, with a 1:1 dilution starting at a final concentration of (unless otherwise stated) 250,000 nM (250 µM). The labelled protein was also added to each of these dilution tubes to a final concentration of 10 nM (unless otherwise stated). These were thoroughly mixed to create sample homogeneity before being centrifuged at 13,000 r.c.f. for 10 minutes at 4 °C to remove precipitates. This series was then loaded into standard capillaries (Nanotemper monolith capillaries) (except for premium capillaries used where stated). These capillaries were scanned at first 20% MST power, then again at 40 % MST power and unless otherwise stated at 40% LED power. The scanning of a capillary first used 5 seconds MST power off, 35 seconds MST power on, then finally 5 seconds MST power off. The results compare a difference between T-jump, thermophoresis, and T-jump plus thermophoresis where shown. Shown in some experiments, the PBS was supplemented with a final concentration of 0.05 % v/v PEG 400 (Sigma Aldrich/Merck) to assess the

effect on capillary sticking. Experiments were completed in triplicate with results containing an average.

2.4.4 Gel Cleavage Assays

Gel cleavage assays were performed using a reaction buffer of 1 mM DTT, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl. Diubiquitin constructs were diluted to a final concentration of 5 µM and USP constructs were diluted to a final concentration of 400 nM. Reactions were left for varying timings listed in results at 22 °C without agitation. Reactions timings were taken by removing some of the reaction mixture (usually 15 µL) and adding this to 0.5 × the volume of 3 × SDS-PAGE loading buffer (usually 7.5 µL). After this, 8 µL of the timings samples were added to an SDS-PAGE gel and run as per the SDS-PAGE gel section above.

Gel images were analysed using the Fiji ad-on for the image J software. Whereby the intensity of the diubiquitin bands were compared to the mono-ubiquitin bands. This was used to generate the percentage of diubiquitin and was plotted using the graphpad prism software ™ (GraphPad Software, San Diego, 2023). Experiments were completed in triplicate with results being an average of these.

2.5 X-ray Crystallography

Purified protein was used for crystallography trials. This began by using pre-made screens purchased from molecular dimensions, these mixes had 70 µL pipetting into the corresponding wells of a 96 well plate (Molecular Dimensions, Swissci PS MRC crystallisation plate). Where indicated, zinc chloride was added to the mother liquor of selected screens to a final concentration of 10 µM in well solution (5 µM in drop). Then a mosquito robot (SPT-LabTech) was used to mix 200 nL of ubiquitin specific protease (usually ~ 8 mg/mL for USP11 constructs, 1 mg/mL for USP17 constructs) with 200 nL of associated mother liquor per well onto the sitting drops. These plates were sealed with clear-tape (ClearVue sealing sheets, Molecular Dimensions). The plates were then kept in an incubator at 20 °C, and were assessed for crystalline growth roughly following the pattern: 1 day, 3 days, 7 days, 10 days, 2 weeks, 3 weeks, 1 month, 1.5 months, 2 months, 2.5 months etc. Assessment was performed optically using a Leica microscope and promising hits were further assessed using the UV microscope (JANSI UVEXp UV). Optimised screens for select conditions were setup by the addition of 1 mL total mother liquor into selected wells of a 24 well plate. The mother liquor was then added to protein samples at a 1:1 ratio of 3 µL each for the sitting drop stand. Plates were sealed using clear tape (ClearVue sealing sheets, Molecular Dimensions) and kept at 20 °C to be checked periodically (as above). Select promising hits were cryo cooled by picking from the wells with cryo-loops (Hampton Research, 18 mm CryoLoops (different loop sizes dependent on crystal size)) and flash frozen in liquid nitrogen before being sent to Diamond Light Source synchrotron (Oxford, UK) for analysis. Where indicated, wells had additional cryo-protectants added,

commonly a final concentration of 30 % glycerol, before cryo cooling occurred. Data analysis was performed following the CCP4 (Winn, et al., 2011) and Phenix (Adams, et al., 2010) pipelines, with diamond Light Source post processing performed by Xia2 Dii3, and analysis performed using CCP4.truncate, CCP4.freeRflag, phenix.PhaserMR, phenix.Autobuild, phenix.refine. With visualisation performed using Pymol (Schrödinger, LLC, 2023) and Coot (Emsley, Lohkamp, Scott, & Cowtan, 2010) programs.

3. Chapter 3: Analysis of USP substrate selectivity

3.1 Introduction

In order to investigate USP interactions, expression trials of USP11 were performed to ensure sufficient amounts of high purity protein could be produced. This is one of the first bottlenecks for protein investigation. Analysis of previous USP11 structures led to an additional USP protein being assessed, USP15 was picked deliberately to analyse the “reverse” mutations and the effects of amino acid sequence changing on substrate specificity. Initial analysis of USPs and their mutants was carried out, first SDS-PAGE gels to certify purity, then Ub-AMC and diubiquitin cleavage gels to elucidate preliminary binding effects. This was used to inform further experimentation which would build on the data herein.

3.2 Expression and purification of USP11 proteins

USP11 selected first as the primary investigated protein, was expressed and purified as standard according to the methods, using 2.4L of bacterial culture grown before a two step FPLC purification.

Figure 3.1, shows the purification of USP11D1D2 (referred to as USP11) from cell lysate. The construct of USP11 produced a large quantity of protein from 2.4L of bacterial culture, able to reach approximately 2 mg of purified product. Following the isolation of the catalytic core of USP11, further purification was performed on different USP11D1D2 mutants, identified from the USP11C318S

monoubiquitin structure. These mutants were expressed and purified to ensure downstream analysis would be feasible. Figure 3.2 shows the expression of two of these mutants and their expression profiles. Whilst the expression was lower than for the wildtype USP11 catalytic domain, the yield was sufficient without further expression optimisation (e.g. ~ 1.5 mg/2.4L, ~0.7 mg/2.4L expression media). Despite the presence of some bands indicating contamination in both the USP11 WT and select mutants (such as USP11H399Q), this was determined to be sufficient purity as these bands are only viewable with severely overloaded wells. This can be compared to lower concentrations of protein in the bands of Figure 3.4, which does not show contamination bands.

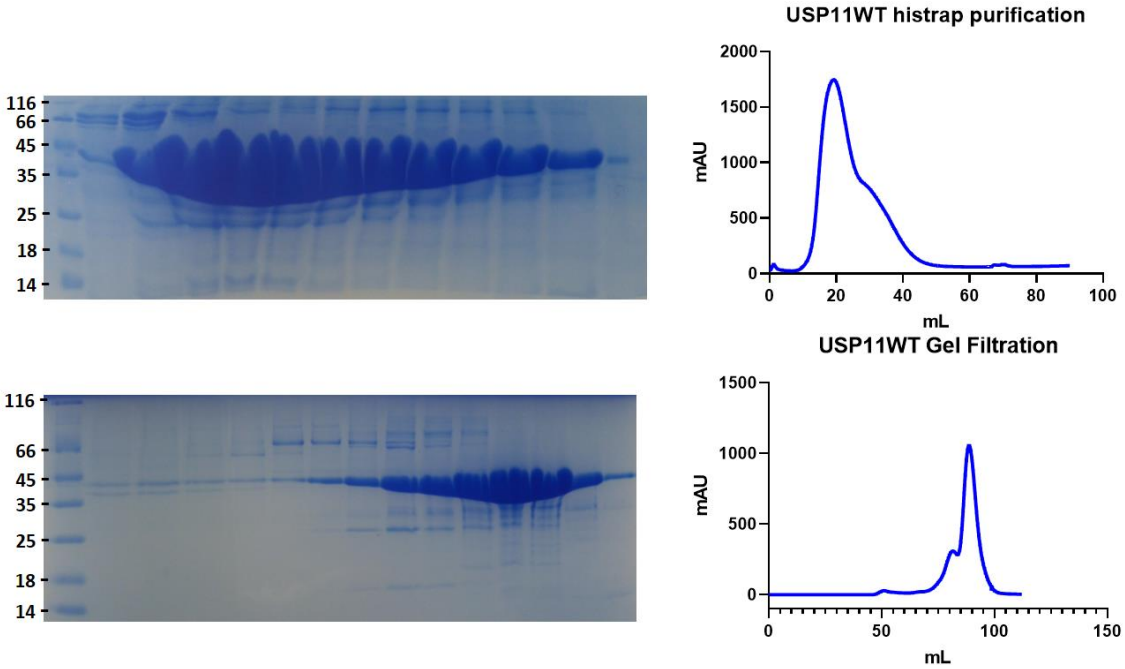

Figure 3.1, SDS-PAGE and chromatogram analysis of USP11 purification from cell lysate. Cell lysate was sonicated and centrifuged to clear insoluble cell debris. Fractions from nickle ion affinity chromatography (top) and size exclusion chromatography (bottom) were run on 18% acrylamide SDS-PAGE gels to separate and identify contamination bands. Protein production was assessed using absorbance units at 280 nM. Although large quantities of USP11 was produced from just a histrap purification, the presence of excess protein bands caused a second purification to be performed. Sufficiently isolated protein fractions were combined, concentrated, and either used for more purification/experimentation or flash frozen with liquid nitrogen and stored at -20 °C.

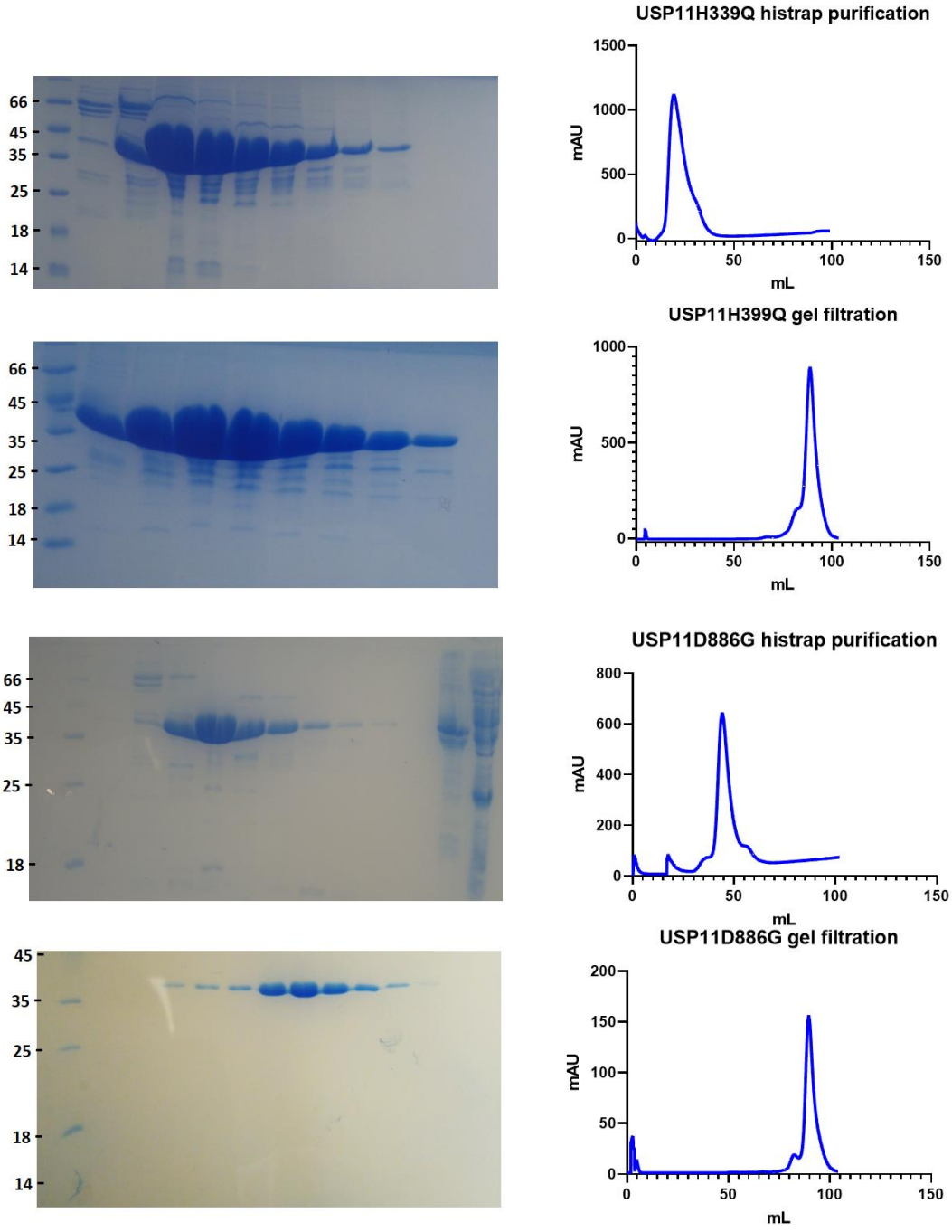

Figure 3.2, SDS-PAGE and chromatogram analysis of USP11 mutants. Cell lysate was sonicated and centrifuged to clear insoluble cell debris. Fractions from nickle ion affinity chromatography (top) and size exclusion chromatography (bottom) were run on 18% acrylamide SDS-PAGE gels to separate and identify contamination bands. Protein production was assessed using absorbance units at 280 nM. Although the yield was less than for the WT USP11 catalytic domain, it was deemed that the expression of these mutants (and other similar mutants) was sufficient for further experimentation without the need for optimisation at this stage. Sufficiently isolated protein fractions were combined, concentrated, and either used for more purification/experimentation or flash frozen with liquid nitrogen and stored at -20 °C.

Purified USP11 mutants were assayed for activity first with Ub-AMC, a fluorophore bound to ubiquitin used in numerous studies (Rut, et al., 2020). This was to confirm activity of the mutants to show they retained deubiquitinating activity (as opposed to the catalytic cysteine residue mutation USP11C318S which binds but does not cleave the Ub-AMC molecule). Once completed, a mutant with the largest deviation from USP11WT was selected for further analysis. The difference was further quantified with varying concentrations of Ub-AMC to generate the initial velocity of the reaction. USP11D886G was then compared to USP11WT with another substrate, linear diubiqiutin through cleavage gels, where incubation of the two proteins was stopped at different timepoints to chart the reaction. Figures 3.3 and 3.4 show the results of these experiments, outlining that whilst USP11D886G had a lower affinity for the Ub-AMC substrate, this mutation led to an increase reaction rate for the cleavage of linear diubiquitin into monomers.

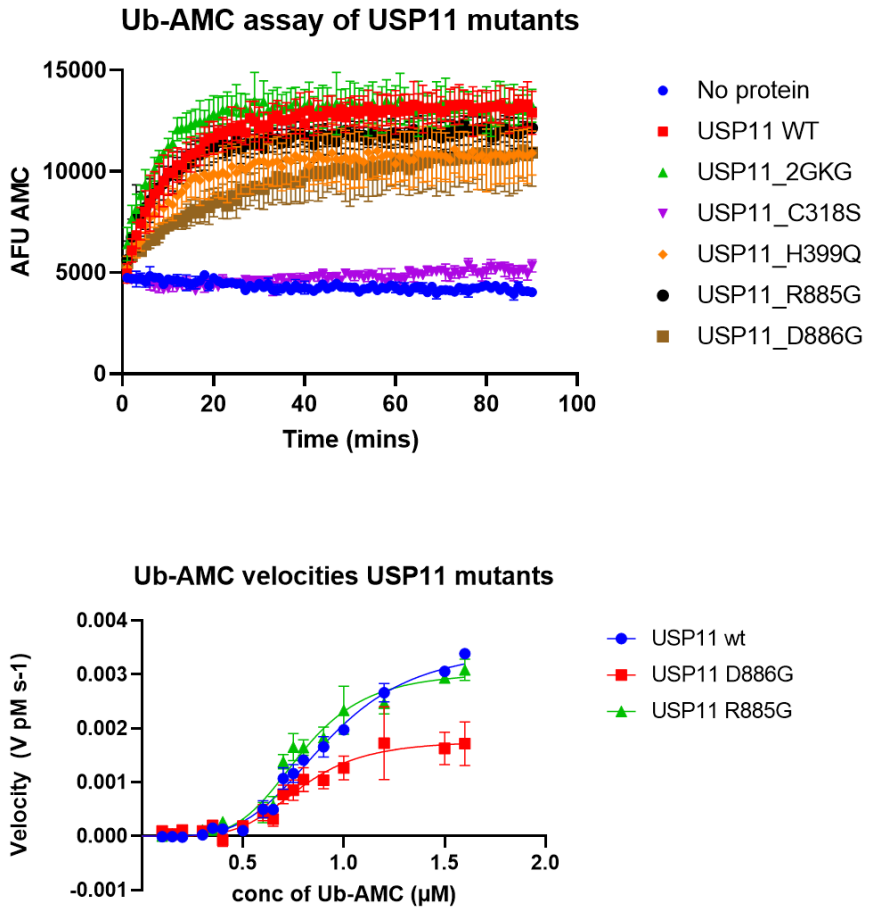

Figure 3.3, USP11 mutants Ub-AMC cleavage. USP11 mutants were expressed and compared for their Ub-AMC cleavage activity. The top graph shows the increase in arbitrary fluorescence units against time of the reaction in minutes. Despite the standard deviation range, trends in the data show the USP11D886G mutant as having the lowest activity (excluding the complete activity knockout of USP11C318S). This led to USP11D886G being compared to USP11WT and USP11R885G for the increasing Ub-AMC concentration assay, allowing an increasing velocity curve (pM per second against concentration of Ub-AMC substrate). This generates more reliable and precise evidence for the difference between the mutants and their catalytic activity.

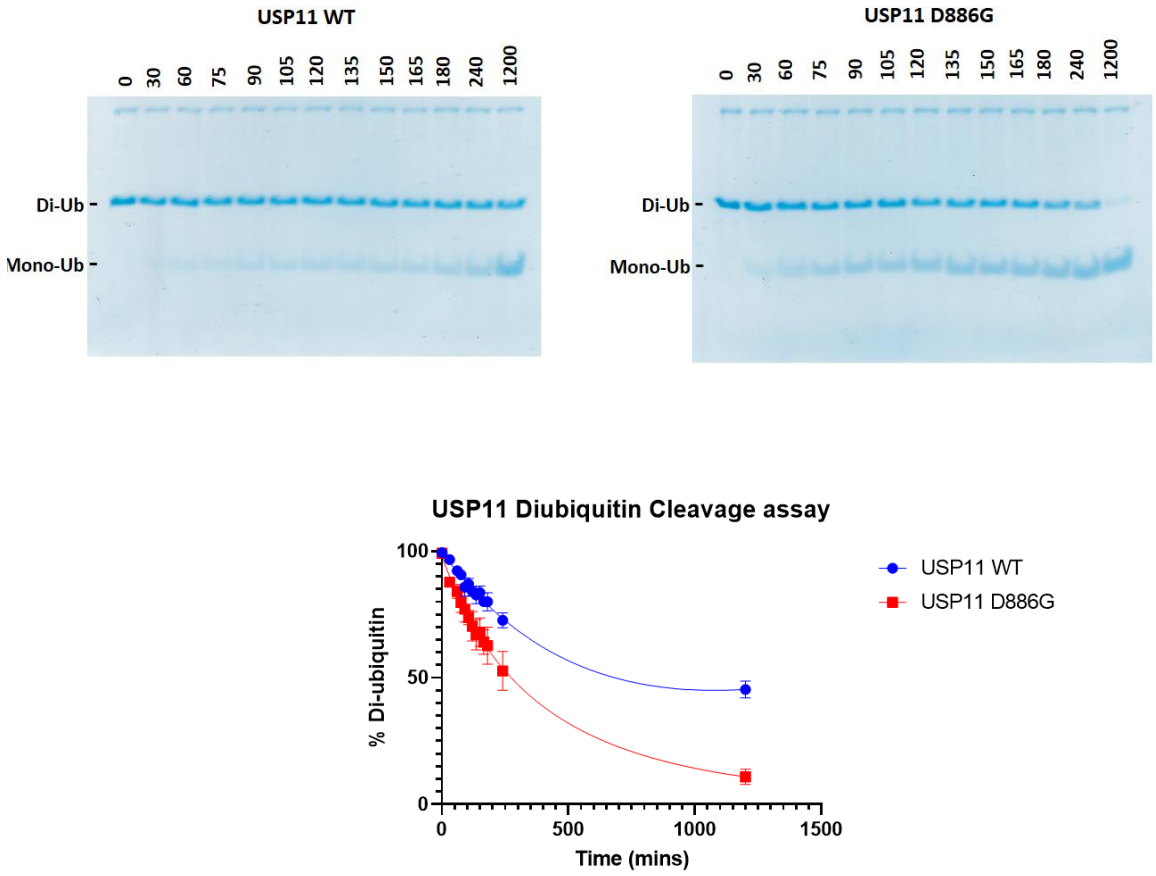

Figure 3.4, USP11 mutants diubiquitin cleavage. The USP11 mutant with the highest difference in Ub-AMC cleavage from the USP11 wildtype was selected for diubiquitin cleavage assays. The graph below the gels shows quantified intensities of the ubiquitin bands against time of reaction. This shows the increased reaction rate for the USP11D886G mutation compared to the wildtype USP11 construct.

The D886G mutant for USP11 caused a decrease in Ub-AMC cleavage, but an increase in linear diubiquitin cleavage. This led to an investigation into an inverse mutation effect, whereby another deubiquitinating enzyme was chosen with a glycine two positions from the catalytic histidine (USP15 G860). This USP was expressed in the same manner as USP11 and was purified with one step affinity chromatography (figure 3.5). The USP15 WT and USP15 G860D was then also used for diubiquitin cleavage assays (figure 3.6). This shows the inverse effect whereby the mutation led to a decrease in reaction rate for the diubiquitin cleavage.

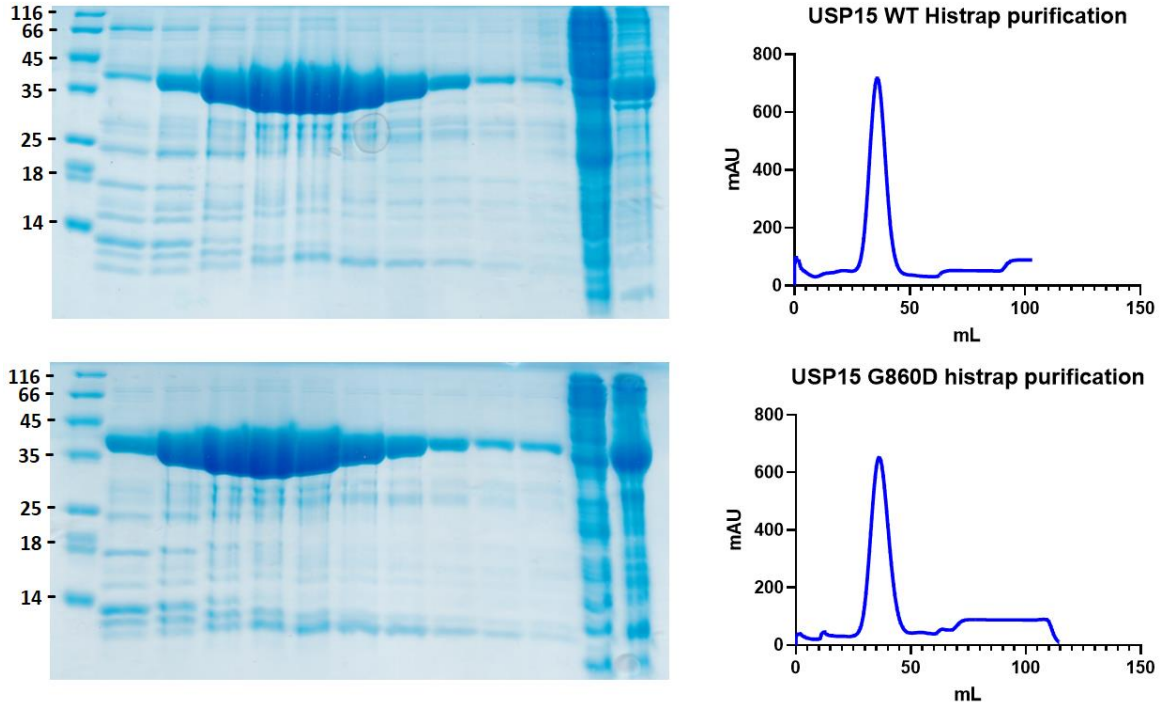

Figure 3.5, USP15 expression analysis. The USP15 constructs selected and expressed were analysed using 18 % acrylamide gels and nickel ion affinity chromatography. Both constructs showed similar expression profiles with similar yields and purities. Fractions with thick bands of the correct molecular weight were pooled and concentrated, flash frozen with liquid nitrogen and stored at -20 °C for future experiments. The ladder on the leftmost lane of both gels is labelled with the MW.

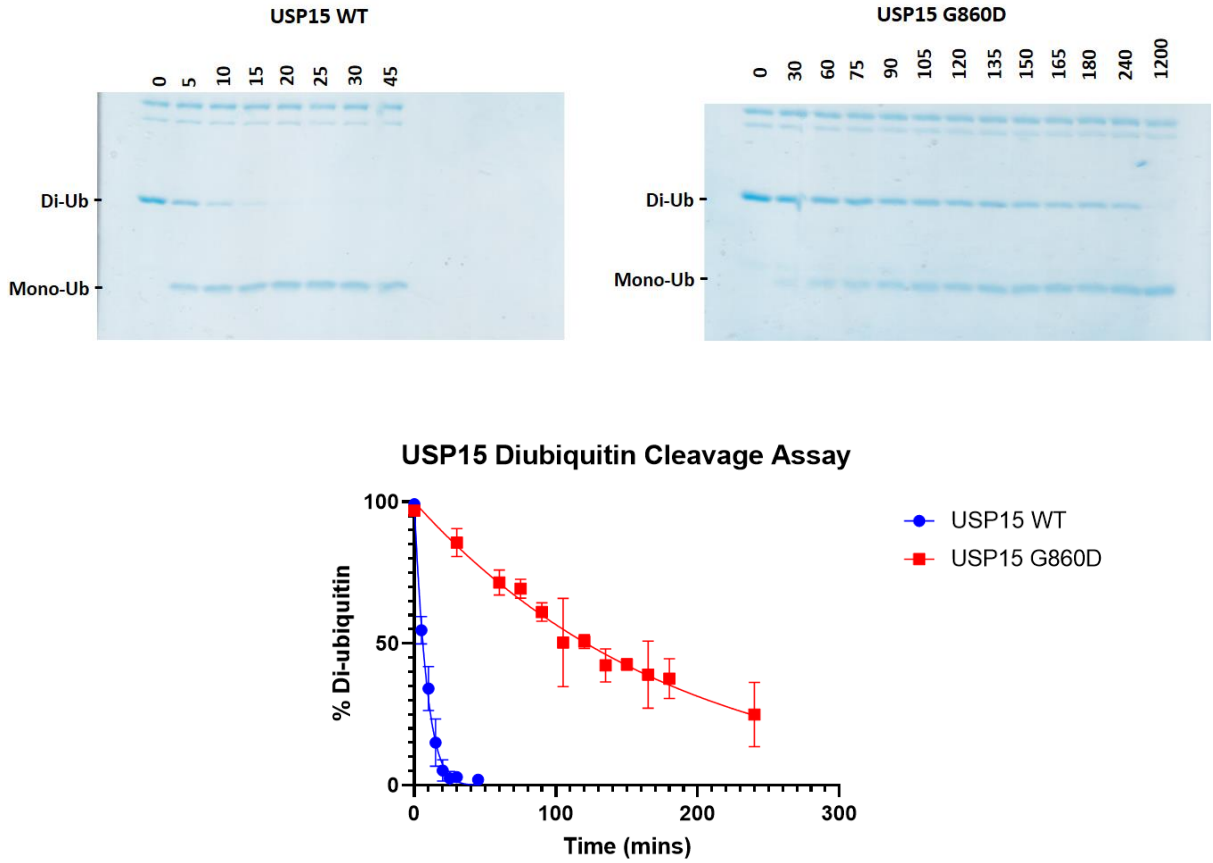

Figure 3.6, USP15 mutants diubiquitin cleavage. The inverse mutation of USP15 for the USP11D886G mutant was compared to the USP15 wildtype to assess impact on diubiquitin cleavage. The above graph shows the quantified intensity of the bands against time of reaction. This shows the USP15 mutant as having lower activity for cleaving diubiquitin compared to the wildtype.

The results show that a glycine residue instead of an aspartic acid residue in the specified location leads to an increase in linear diubiqiutin cleavage. With USP11 showing the increase in diubiquitin cleavage and USP15 showing the decrease through mutations from their wild type sequences. The combination of these two cases led to further analysis utilising different methods to more closely examine the effect of this mutation on substrate specificity.

3.3 ITC and MST analysis of USP11 and USP15

Following from the results previously, more analytical methods were selected to investigate the changes in binding from the G/D mutation. Isothermal titration calorimetry (ITC) was chosen as the first method for comparison between the USP proteins and different target substrates. Figure 3.7 shows the first comparisons between USP11 and USP11D886G and their interactions with diubiquitin, C318S mutations were also introduced to prevent cleavage of (but not binding to) the substrates. USP15 constructs also included mutations in the catalytic C269S residue to prevent cleavage of but not binding to substrates.

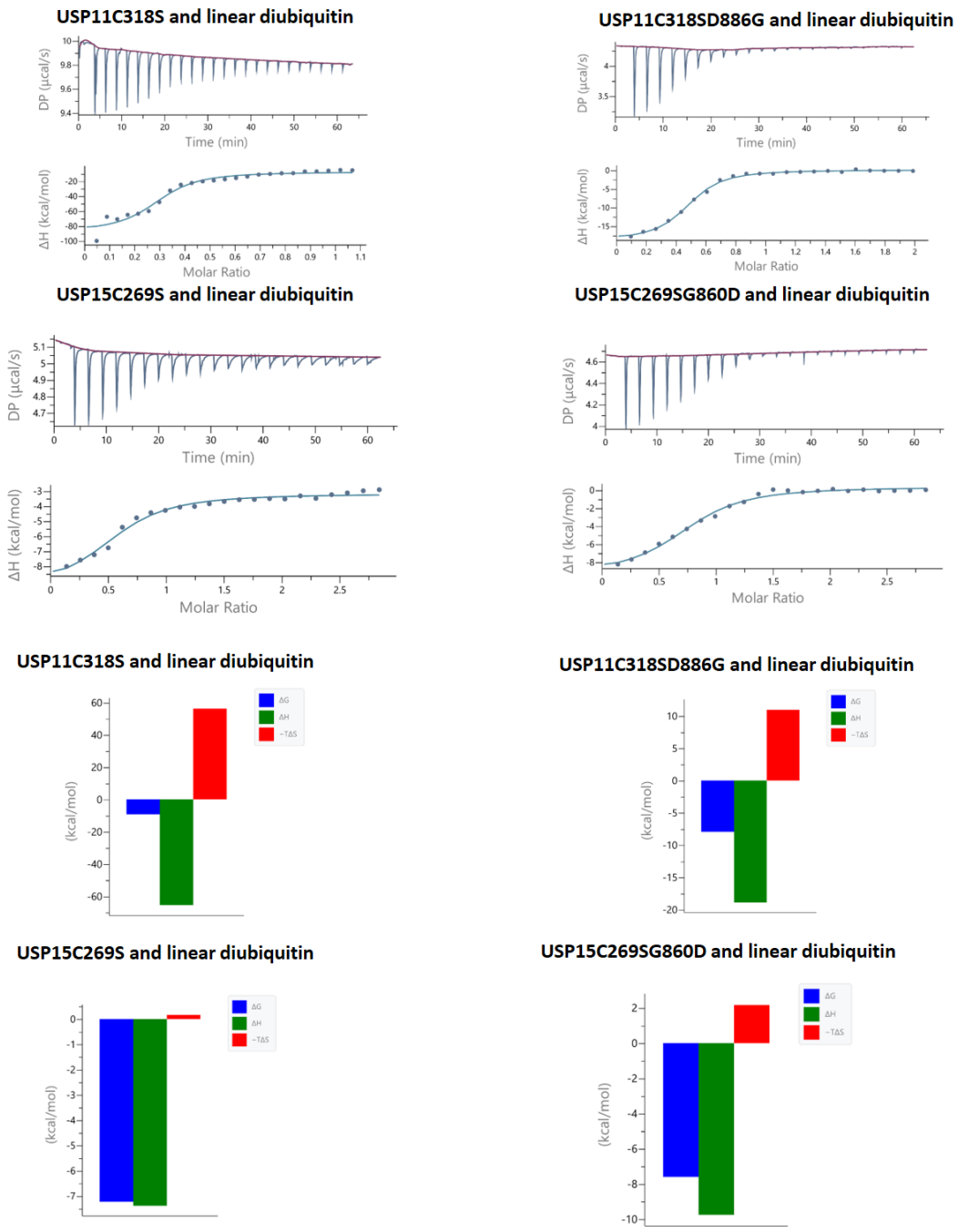

Figure 3.7, USP mutants diubiquitin ITC. Isothermal titration calorimetric readings of USP11 and USP15 mutants compared with diubiquitin. This shows the traces of the reactions, with a fixed concentration of USP and an increasing molar ratio of added linear diubiquitin. The graphs above show the traces and the graphs below show the energy (kcal/mol) of the ΔG (blue), ΔH (green), and –TΔS (red). The traces show relatively similar reactions despite noticeably higher –TΔS results for the USP11 reactions. These are representative graphs from at least 3 repeat experiments per condition.

The reactions between inactivated USP and linear diubiquitin are similar when comparing the graphs of the ITC runs. The output data table (table 3.2) shows relatively similar Kd constants, the differences between Kd and mutations is not always consistent with the trends seen in the ubiquitin gel cleavage assays (figure 3.4, figure 3.6). Due to the expression of USP11 and the unclear outcome of previous data, further ITC assays were completed with different ubiquitin variants as substrates. The results of these assays are seen in figure 3.8 below. These ubiquitin constructs include ubiquitin with an additional glycine residue (Ub-G), ubiquitin with three additional glycine residues (Ub-GGG), and ubiquitin with an additional methionine residue (Ub-M).

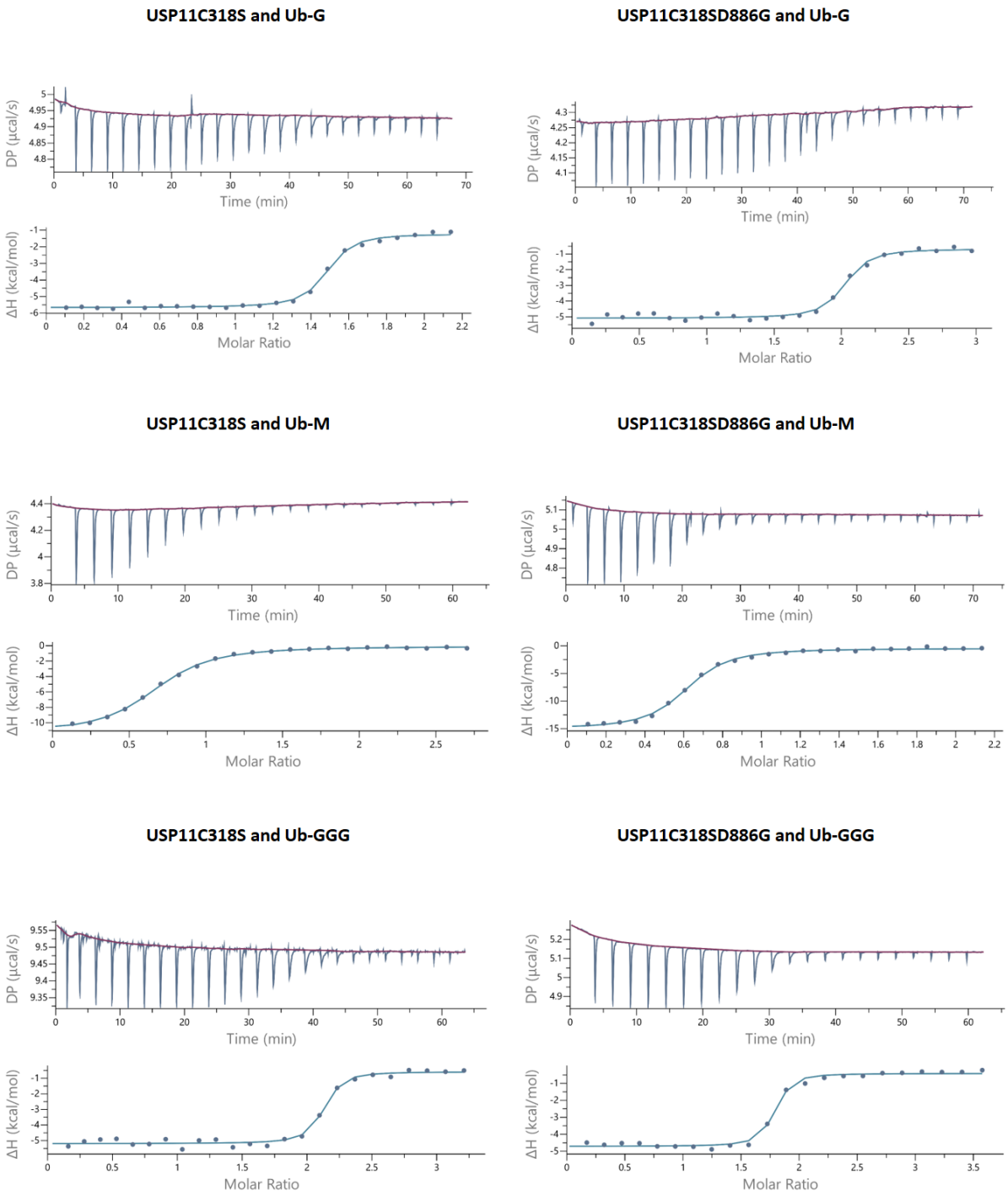

Figure 3.8, ITC of USP11 constructs with different ubiquitin substrates. USP11 C318S catalytically inactive mutant compared with catalytically inactive USP11 construct with the additional D886G mutation. Multiple small constructs were made of single ubiquitin molecules with additional amino acid(s). The differences between constructs is minimal when comparing traces of ITC runs, although binding between Ubi-M and USP11 constructs appears to be tightest, with the lowest molar ratio at the reaction end point.

The ITC runs of the USP11 constructs with different ubiquitin substrates does not show a vast difference between USP11 mutants and their binding. This mimics a similar pattern to the USP11 constructs with diubiquitin constructs. Following this, and assessment of the data in table 3.2, another technique was used in an attempt to elucidate the effects of this mutation on binding of not just diubiquitin constructs but also on ubiquitin bound amino acid(s). Microscale thermophoresis was selected to assess binding between USP11 and two ubiquitin substrate targets, Ub-GGG and diubiqiutin. This work shows the differences between the two substrates and their binding with USP11 as well as allowing a comparison between the other techniques used to assess binding. Figure 9 and 10 show the results of these MST experiments.

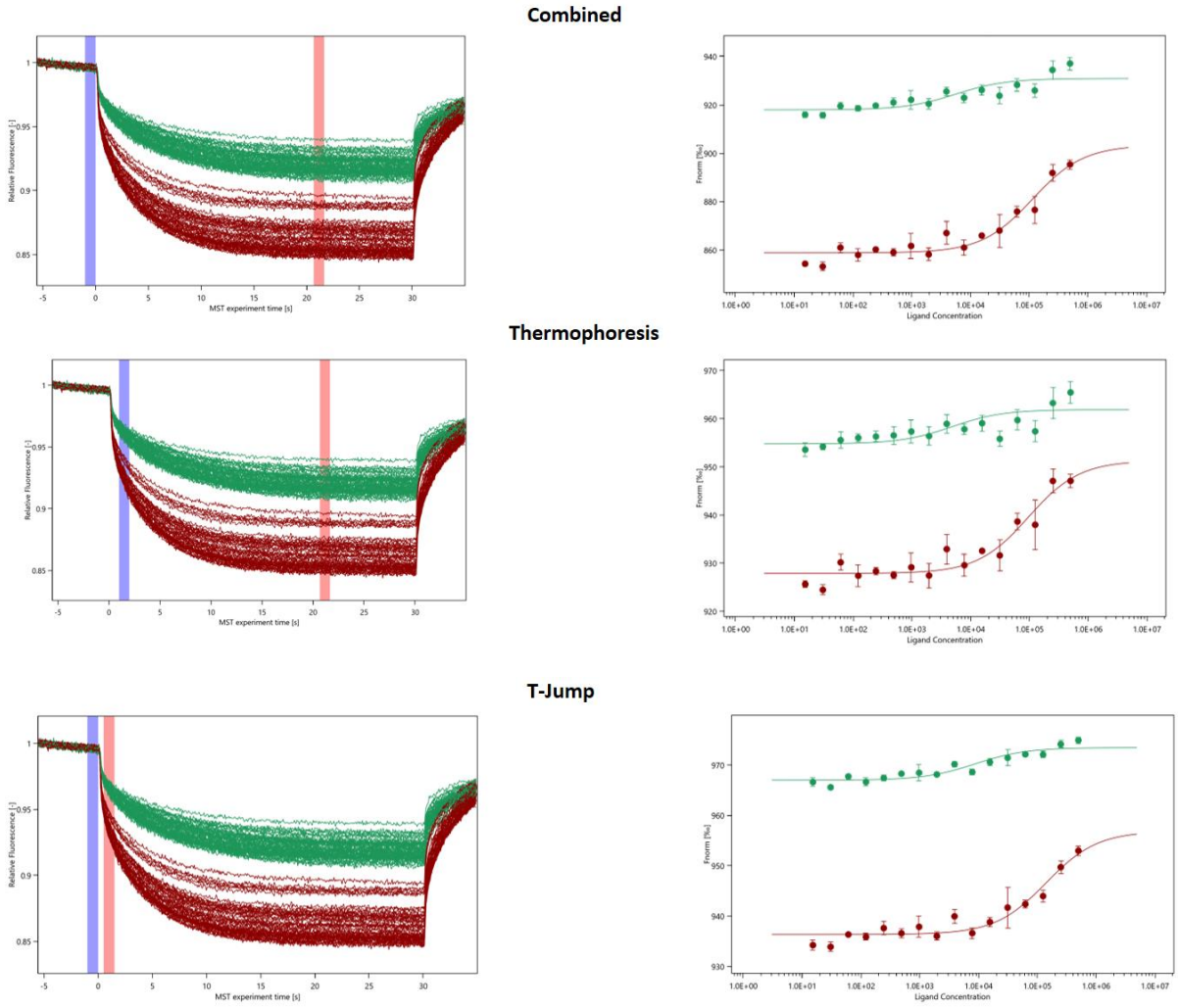

Figure 3.9, MST of USP11wt and Ub-GGG showing different analysis methods for MST data. Comparison for different analysis methods of MST data. “Combined” at the top of the figure shows the difference between Fnorm without MST power and the effect extended MST power has on the sample. Thermophoresis eliminates the initial temperature jump at the initial triggering of MST power and analyses just the difference overtime. The T-Jump shows just the initial change from 5 seconds after MST power has been activated. For the above graphs this shows a similar trend, with higher variability for the combined and thermophoretic samples when compared to the T-Jump analysis method.

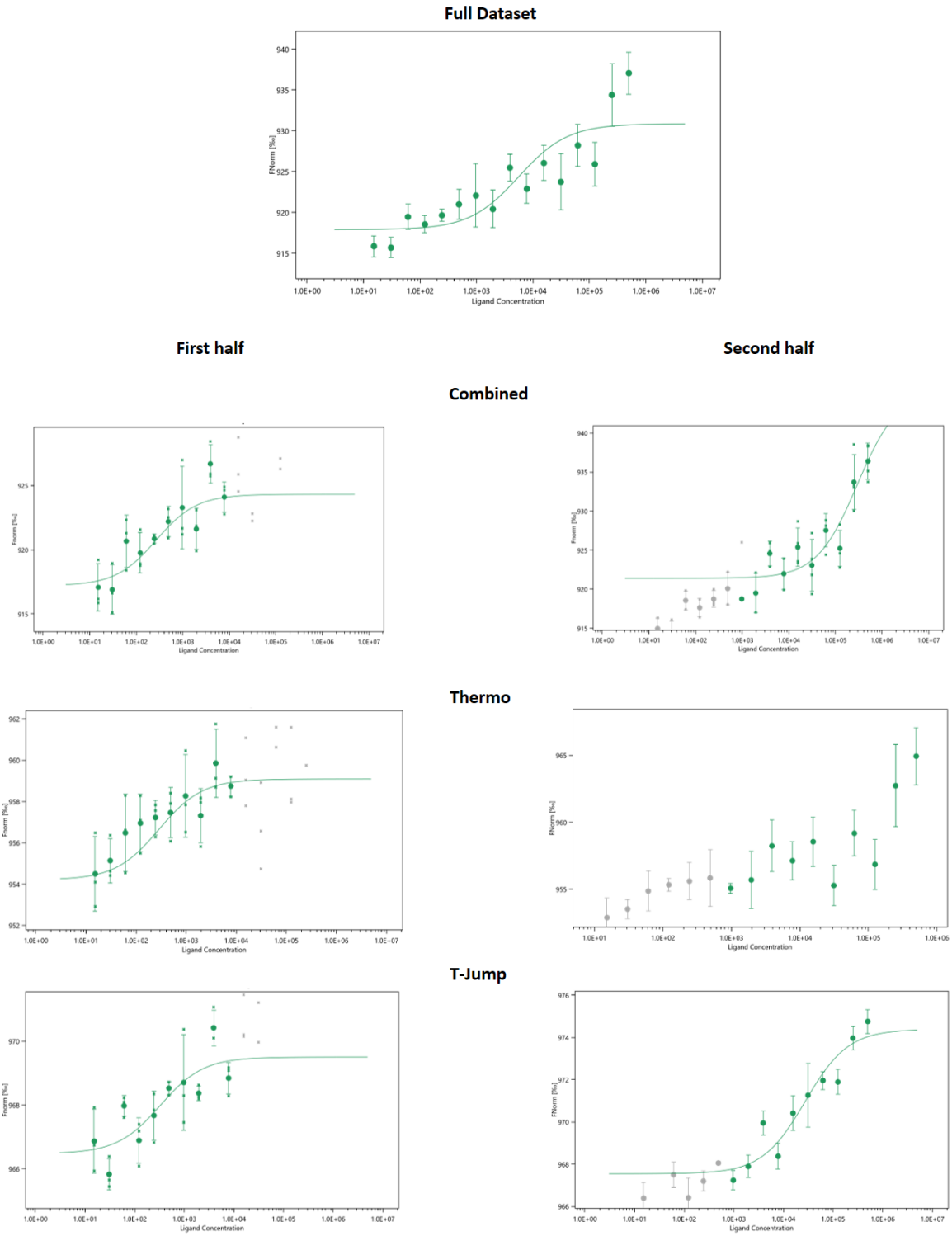

Figure 3.10, MST of USP11wt and Ub-GGG. The analysis of USP11 and UbGGG showed signs of a plateau in the middle of the dataset, whereby the removal of some of the first five or last five data points still yielded sigmoidal curves, but with different end points for the reaction. This is most viewable in the T-Jump dataset. However, the data is still most compelling for the full dataset, with the lowest noise to signal ratio.

Figure 3.9, included for completeness shows examples of the main methods used for analysing MST data, as different trends in different analysis methods holds different implications for the data beyond calculating Kd. Figure 3.10 looked more closely at a dataset for USP11C318S with a monoubiquitin Ub-GGG construct. This also looked closely at the data to investigate not just variability between runs but also if other binding effects were taking place between USP11 and monoubiquitin substrates. Whilst not completely elucidated, this was also included to allow a direct comparison with Figure 3.11, which shows more substantial differences in the progression of an MST run. Figure 3.11 showing USP11C318S and diubiquitin

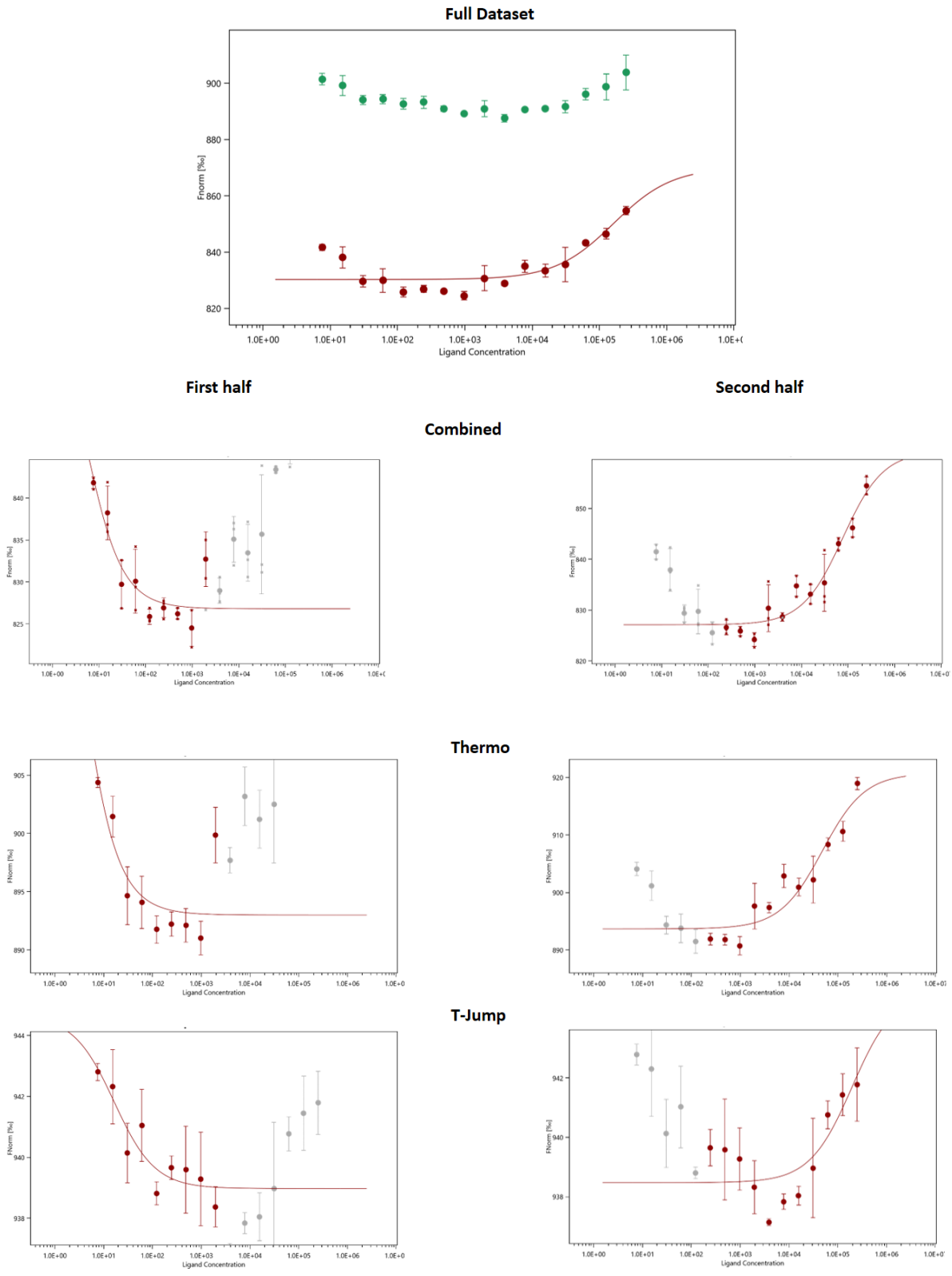

Figure 3.11, MST of USP11C318S and diubiquitin. USP11C318S with diubiquitin showed a high level of curve dimorphism. This is highlighted when the first half and the second half of the graphs is separated and treated as two different datasets. This produces sigmoidal curves with different trends. The higher concentrations of diubiquitin produces the curve with the lowest error to signal ratio. The difference between analysis methods (T-jump, thermal, combined) shows little variation.

The dissociation constants calculated from USP11C318S with the two substrates, UbGGG and diubiquitin, were collected into table 3.2 for analysis and comparison to table 1 which contains the ITC data. Highlighted in table 3.1, and expanded upon with figures 3.10 and 3.11, is the indication of non-monophasic binding. The presence of this data ambiguity led to one final additional substrate for ITC, whereby the diubiqiutin construct had the tail end of the second diubiquitin shortened to reduce substrate flexibility and to attempt to eliminate this tail contributing to a secondary binding event. This is viewable in figure 3.12.

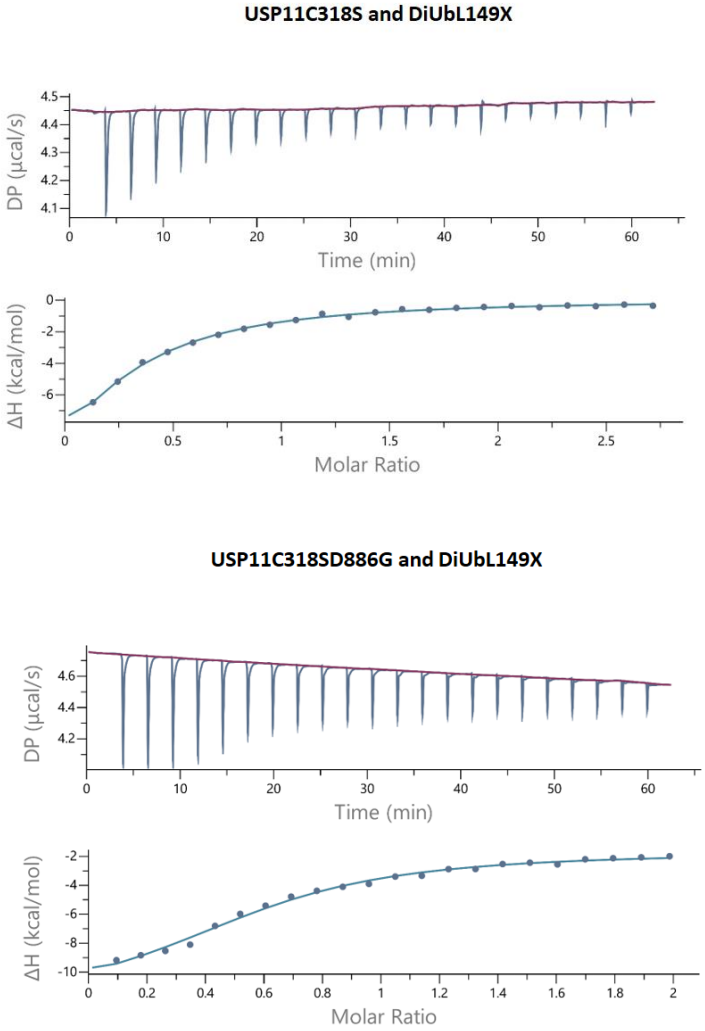

Figure 3.12, ITC of USP11C318S and diubiquitinL149X. ITC of the USP11 mutants and L149X diubiquitin shows no significant differences between the USP11 mutants, with the USP11C318S having a slightly steeper curve. ITC experiments were averaged between three repeats with the Kd and other data deposited in table 3.1.

Sample	K_D (M)	ΔH (kJ/mol)	ΔG (kJ/mol)	-TΔS (kJ/mol)	N (sites)
USP11C318S and UbM	1.43E-06 ±1.54E-07	-14.6 ±3.77	-7.98 ±0.0612	6.60 ±3.75	0.509 ±0.170
USP11C318SD886G and UbM	5.77E-07 ±1.36E-07	-13.0 ±3.24	-8.53 ±0.146	4.52 ±3.13	0.63 ±0.101
USP11C318S and UbG	5.12E-08 ±2.79E-08	-4.17 ±0.836	-10.1 ±0.414	-5.87 ±1.11	1.58 ±0.355
USP11C318SD886G and UbG	1.19E-07 ±1.40E-07	-9.64 ±11.4	-9.77 ±0.706	-0.136 ±12.0	1.53 ±1.05
USP11C318S and UbGGG	2.99E-08 ±1.92E-08	-4.45 ±0.271	-10.4 ±0.530	-5.98 ±0.779	1.90 ±0.106
USP11C318SD886G and UbGGG	3.90E-08 ±1.69E-08	-4.97 ±0.601	-10.1 ±0.272	-5.19 ±0.631	1.93 ±0.219
USP11C318S and diubiquitin	2.21E-06 ±2.94E-06	-18.5 ±12.8	-8.13 ±0.781	10.4 ±12.7	0.844 ±0.617
USP11C318SD886G and diubiquitin	8.90E-07 ±4.66E-07	-24.4 ±6.73	-8.33 ±0.340	16.1 ±6.45	0.387 ±0.143
USP11C318S and DiUbL149X	1.38E-05 ±1.35E-06	-13.4 ±5.79	-6.64 ±0.0546	6.77 ±5.80	0.398 ±0.162
USP11C318SD886G and DiUbL149X	8.37E-06 ±1.63E-06	-7.84 ±3.91	-6.94 ±0.120	0.920 ±4.06	0.631 ±0.0686
USP15C269S and diubiquitin	5.61E-06 ±1.47E-06	-6.88 ±0.729	-7.35 ±0.163	-0.465 ±0.885	0.577 ±0.00424
USP15C269SG860D and diubiquitin	3.32E-06 ±4.10E-07	-9.53 ±0.332	-7.66 ±0.0778	1.88 ±0.403	0.759 ±0.00283

Table 3.1, ITC of USP11 and USP15 mutants with ubiquitin substrates. This table of data shows the K_D, and other information from the ITC runs. Results were kept to 3 significant figures, with the error showing standard deviation. The tightest binding appeared to be between USP11 mutants with UbGGG.

Sample (showing K_D (M))	DataSet	Thermo + T-jump	Thermo	T-Jump
USP11C318S and UbGGG	Complete	2.93E-05 ±3.01E-05	2.01E-05 ±2.67E-05	2.58E-05 ±4.24E-05
	First Half	2.7E-07 ±3.01E-07	2.87E-07 ±3.06E-07	3.28E-07 ±4.03E-07
	Second Half	2.94E-04 ±2.252E-04	no fit	2.74E-05 ±2.79E-05
USP11C318S and diubiquitin	Complete	6.23E-05 ±8.65E-05	3.48E-05 ±4.6E-05	4.07E-06 ±6.97E-06
	First Half	4.19E-09 ±6.76E-09	2.48E-09 ±7.27E-09	1.51E-08 ±1.5E-08
	Second Half	7.71E-05 ±3.42E-05	4.77E-05 ±2.81E-05	1.97E-04 ±3.51E-04

Table 3.2, MST of USP11C318S with ubiquitin substrates. This table of data shows the K_D of MST repeats. Results were kept to 3 significant figures, with the error showing standard deviation. The table shows not only the full dataset but additionally the dataset split into two curves. These data show a large error between experimental repeats of the MST experiments but with USP11C318S showing a slightly tighter K_D with UbGGG as opposed to linear diubiquitin.

Tables 3.1 and 3.2 showing the collective data for ITC and MST runs both indicate a higher affinity K_D for USP11C318S with UbGGG over linear diubiquitin. The difference is however much larger for the ITC runs compared to the MST experiments. The MST experiments also showed different K_D when analysed using different thermophoretic analyses, however these also all showed a tighter binding between USP11C318S and UbGGG compared with the diubiquitin substrate.

To confirm the physiological relevance of the D886G mutation, a final series of diubiqiutin cleavage assays were performed with USP11 physiological target K63-linked diubiquitin (figure 3.13) (Guo, Tang, Yuan, Zhang, & Wang, 2022). This shows the USP11WT cleaving the K63-diubiquitin substrate more efficiently than the D886G mutant, which is contrast to the linear diubiquitin cleavage with the two USP11 mutants. This indicates a substrate specificity between diubiquitin chains for this mutation.

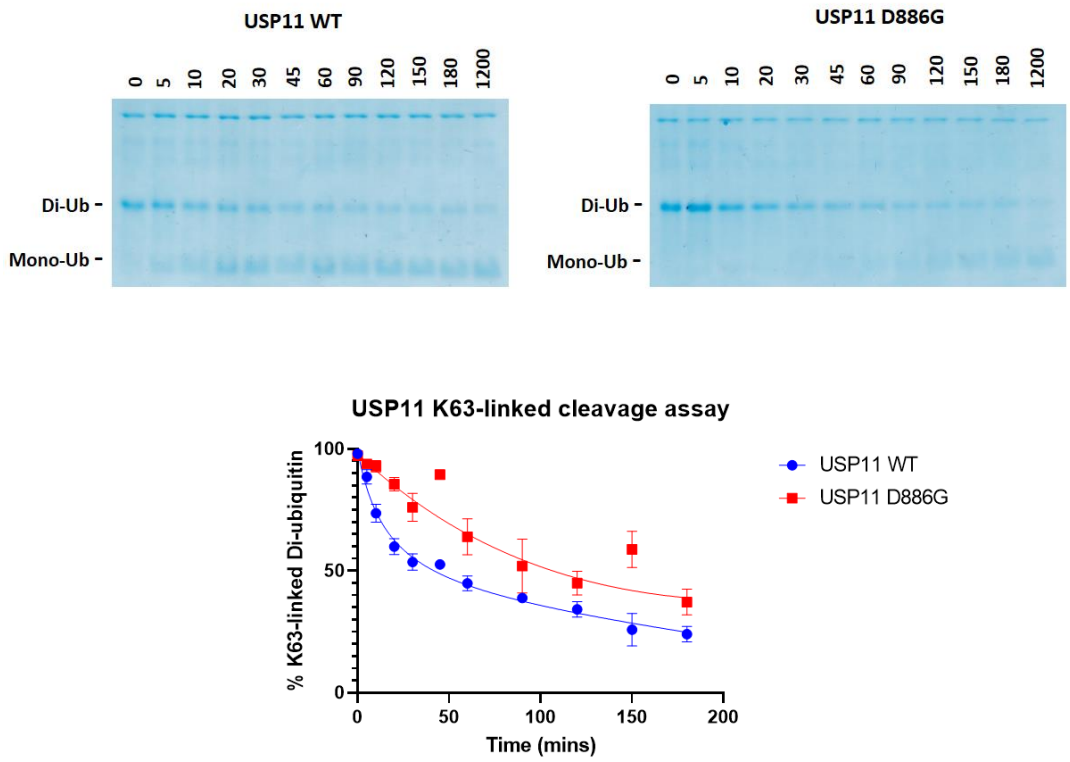

Figure 3.13, diubiquitin cleavage assay of K63 linked diubiquitin. Showing the cleavage of K63-linked diubiquitin by USP11 and the USP11D886G mutant. This shows USP11WT cleaving more efficiently from a repeat of three experiments.

3.4 Expression and Purification of USP17 constructs

In an attempt for further categorise and elucidate the effects of the mutation on USP substrate binding, an additional USP was chosen with the similar aspartic acid residue in the same relative position (two residues before the catalytic histidine). The chosen protein USP17 has relatively little structural data presently available but has been implicated in many disease states (Ducker & Shaw, 2021). This investigation was used to determine the usability of USP17 as a study subject through expression and purification attempts to obtain pure and high yield samples. Figure 3.14 shows the initial expression experiments performed using a bacterial expression system.

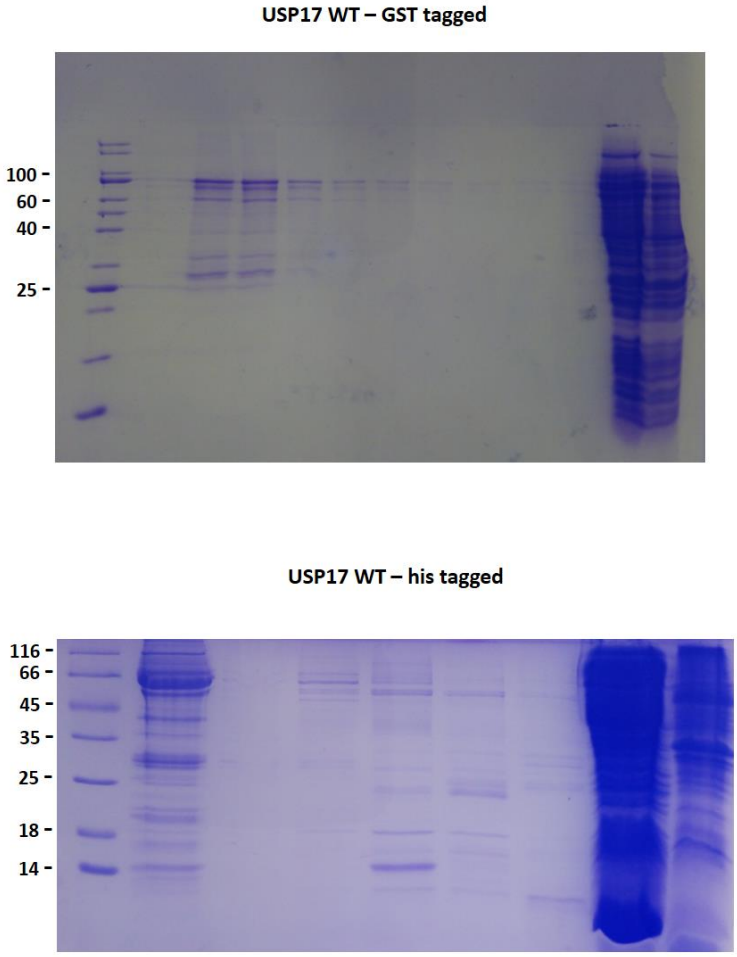

Figure 3.14, USP17 initial expression gel analysis. Initial expression of USP17 through GST and histrap tagged constructs. This shows the low expression with low purity for both USP17-GST (85.3 kDa) and USP17-his (60.4 kDa) constructs.

Initial expression trials of USP17 constructs showed poor yield (<0.1 mg) and low purity. This led to additional expression trials changing both constructs and expression conditions detailed in figure 3.15, including with additional tags. The 1BKR tag was used in conjunction with truncating the USP17 protein to its catalytic core. The full length construct was cloned into the P.Cold1 vector for lower temperature protein expression. This was also coupled with additional additives for slower/most stable protein production.

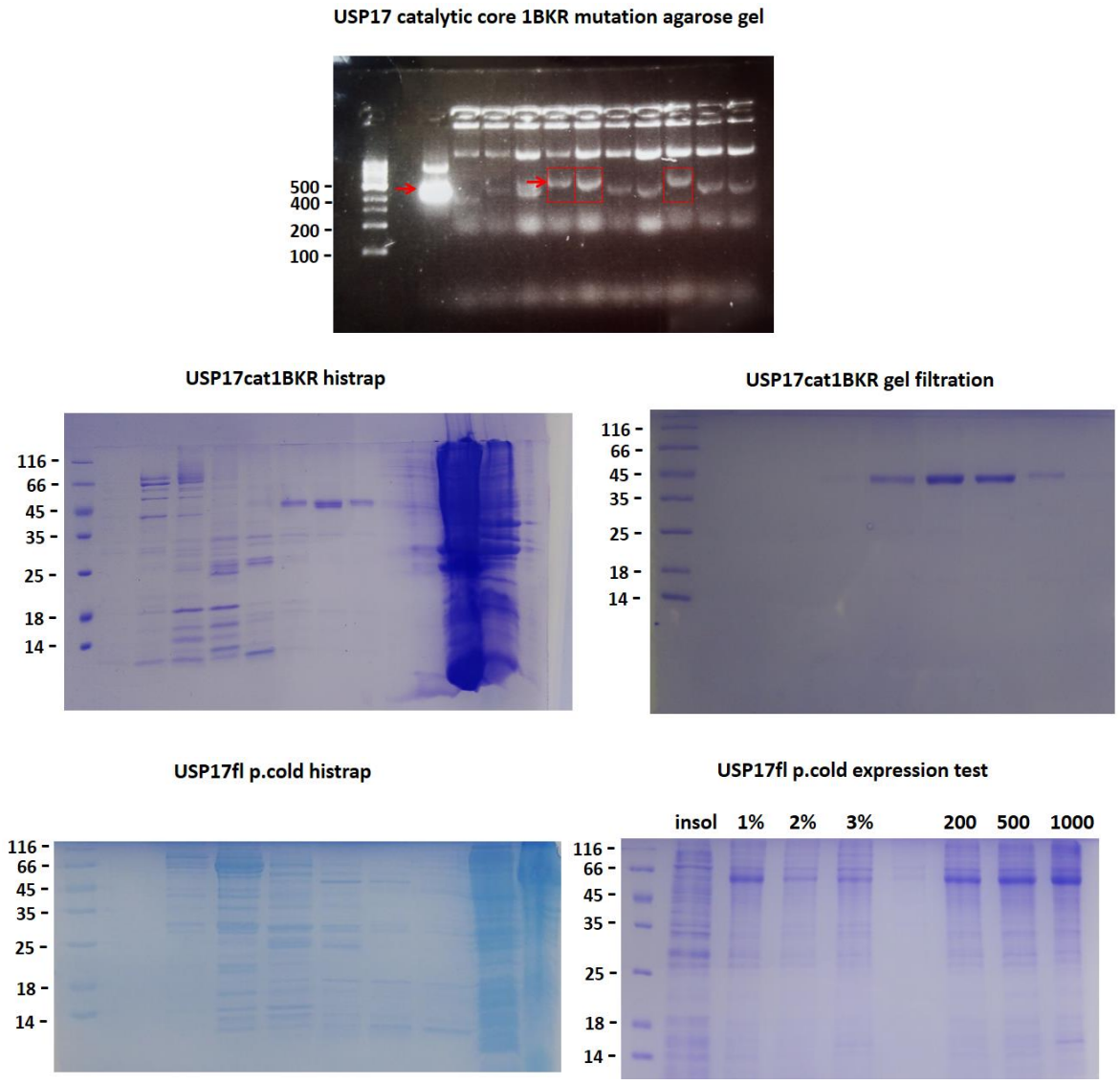

Figure 3.15, USP17 further expression condition trials. The agarose gel shows different clones from the 1bkr insertion experiments, whereby clones with larger molecular weights were selected for sequencing (which confirmed presence of the tag). This catalytic usp17 construct (USP17cat1BKR) was purified and shows suitable purity from SDS-PAGE assessment. Below the USP17 full length (USP17fl) construct was expressed using the p.cold vector. This vector was paired with additional supplements, the wells showing 1, 2, and 3 % indicate ethanol added to the media during expression. The wells showing 200, 500, and 1000 indicate the mM concentration of sorbitol present during expression media. The supplements and plasmid changes indicate an increased protein production.

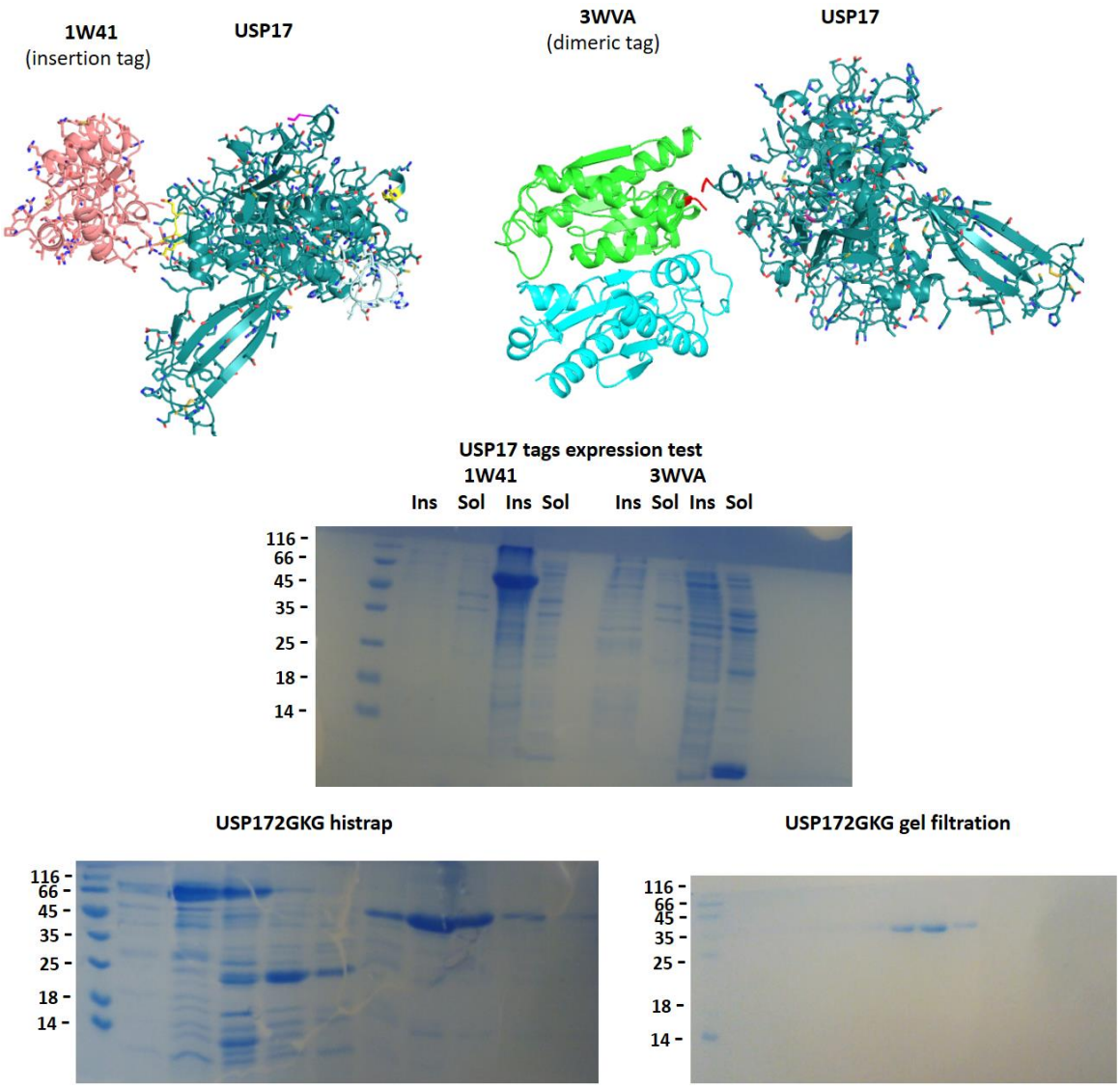

Figure 3.16, USP17 expression conditions with additional protein tags. SDS-PAGE analysis was carried out on the expression of USP17 with different tags. 1W41 and 2GKG were designed as loop stabilising insertion tags, 3WVA was designed as a dimer tag. Homology modelling was used to demonstrate where the tags were cloned into the USP17 sequence. SDS-PAGE analysis shows the insoluble (ins) and soluble (sol) fractions of the 1W41 and 3WVA clones. Below SDS-PAGE shows the purification of the 2GKG insertion tag USP17 construct, which maintains the highest full length USP17 production and purity.

Figure 3.16 shows the analysis of different USP17 tag constructs. Insertion loop and dimer constructs were used in an attempt to increase stability and production of protein. The 2GKG insertion loop tag produced the highest level of protein compared to the other constructs. Despite low initial production, this amount of protein was increased through the use of autoinduction media (figure 3.17). However, regardless of the increases made in protein production and purity, the final amount of protein produced was still of a low total (~0.2 mg from 2.0 L of culture). This led to crystallisation trials which did not yield enough promising hits for optimisation trials. Due to this, and the low production level, USP17 was not taken further through investigation.

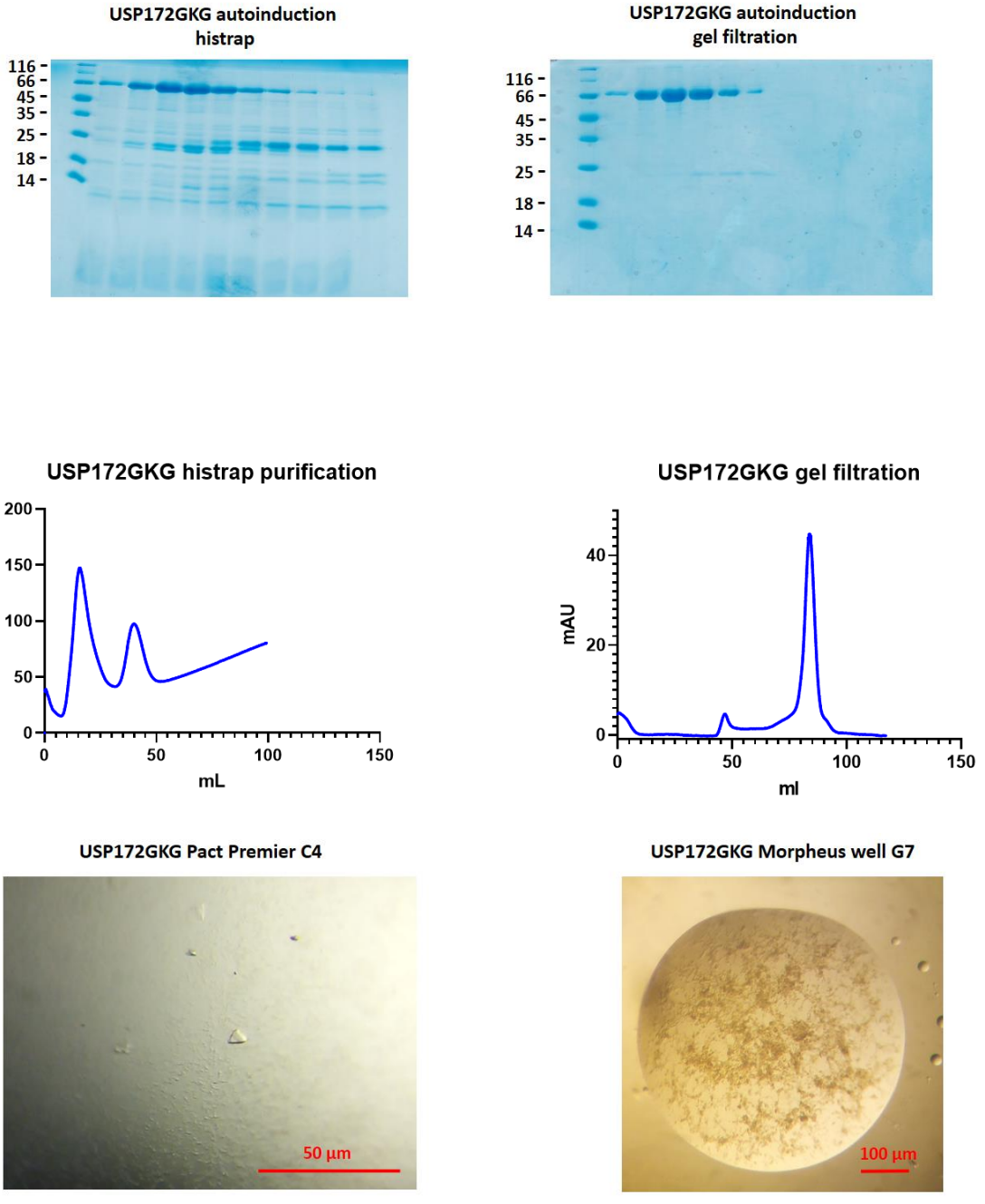

Figure 3.17, USP17 expression with 2GKG and autoinduction media. SDS-PAGE analysis and akta traces of USP172GKG autoinduction expression purification. This shows that relatively pure USP17 was produced with relatively small amounts of protein. Protein was added to spase matrix crystal screens in which two aggregate conditions were observed, Pact Premier C4 (25% PEG 1500, 0.1M PCTP, pH 7.0), and Morpheus I G7 (0.1 M carboxylic acid mix (0.2M Sodium formate; 0.2M Ammonium acetate; 0.2M Sodium citrate tribasic dihydrate; 0.2M Potassium sodium tartrate tetrahydrate; 0.2M Sodium oxamate), 0.1M sodium Hepes-MOPS, 20 % (v/v) glycerol, 10 % (w/v) PEG 4000, pH 7.5).

3.5 X-ray Crystallography of USP11 and linear diubiquitin

USP11 was co-expressed with linear diubiquitin. The USP11 construct was USP11D886GC318S_2GKG to prevent diubiquitin cleavage. The D886G mutant was chosen to encourage USP11 and diubiquitin binding. Additionally the USP11 construct was expressed with the 2GKG insertion stabilising tag. This was due to the previous success of the tag used when crystallisation.

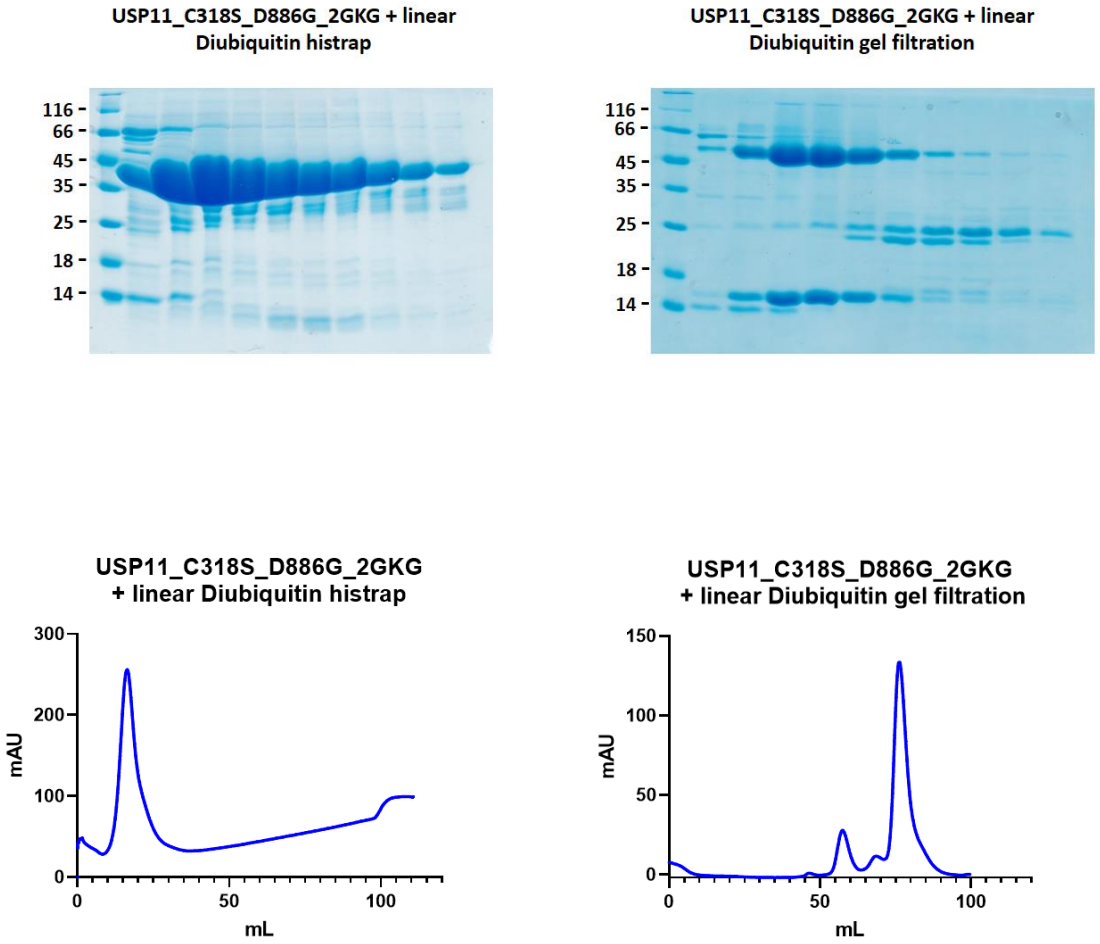

Figure 3.18, USP11_C318S_D886G_2GKG expression with linear diubiquitin. The gels for the histrap and gel filtration size exclusion chromatography for the USP11_C318S_D886G_2GKG construct co-expressed with linear diubiquitin are on the top row. These show the large USP11 band just above the 45 kDa ladder mark. The linear diubiquitin band co elutes with the USP11 band, with the reducing SDS-PAGE gel separating the two proteins with the linear diubiquitin band showing around the 14 kDa ladder mark. The chromatograms below shows the AKTA purifications of the USP11 construct, with small contaminants excluded from the gels.

The expression of the USP11 crystallisation construct produced a sufficient yield of roughly 0.7 mg of USP11 and diubiquitin complex per 2.4L expression culture. This construct was concentrated to roughly 10 mg/mL, after which a mosquito crystal robot was used to pipette 0.2 nL per well into a 96 well plate with different sparse matrix screens, with added zinc chloride. The hits from these were identified primarily through optical assessment with a Leica microscope in conjunction with JANSI UVEXp UV microscope. Optimised screens were developed from potential hits but these yielded no discernible crystals. Due to the hits generated, crystals were collected and taken to the Diamond light source Synchrotron. The highest quality dataset was collected from the Morpheus II H3 condition (40 mM Polyamine mix [0.1M Spermine tetrahydrochloride, 0.1M Spermidine trihydrochloride, 0.1M 1,4-Diaminobutane dihydrochloride, 0.1M DLOrnithine monohydrochloride], 0.1 M 6.5 MOPSO Bis-Tris, pH 6.5, 10% w/v PEG 8000, 20% v/v 1,5-Pentanediol), with figure 3.19 also showing an example of other precipitate from the LMB screen A1 (25 % 1,4-Butanediol, 0.1 M Tris-HCl pH 8.0). As stated in the Methods section, all conditions included the final well concentration of 0.1 µM zinc chloride.

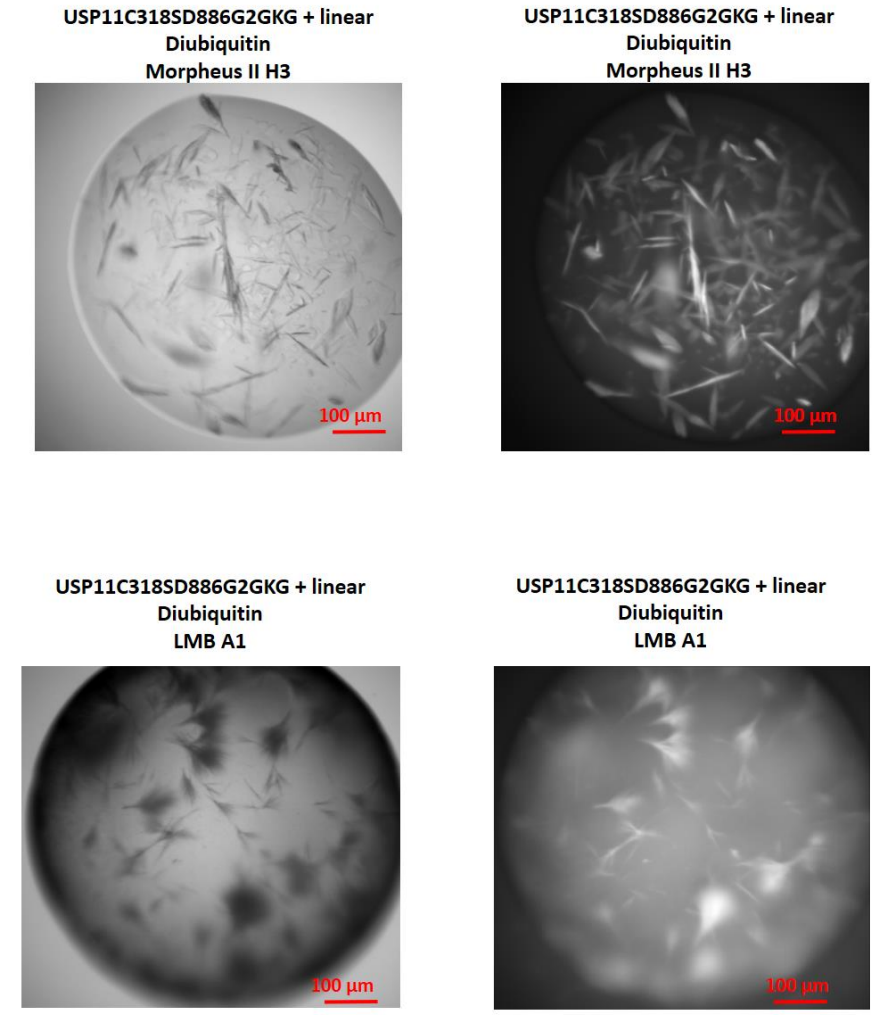

Figure 3.19 USP11_C318S_D886G_2GKG and linear diubiquitin co-crystallisation trials. These example crystallisation hits from the crystallisation screening show light microscope images on the left, with the comparative UV images on the right. This shows the crystalline objects as protein containing as opposed to salt or additive false positives. The bottom image shows diffraction data collected from a crystal collected from the Morpheus II H3 condition.

This crystal conditions Morpheus II H3 produced the highest quality dataset. This crystal was cryocooled before data was collected with 0.2° oscillations at 100K. Downstream processing was performed by Xia2 3dii, the data from this is displayed in table 3.3. The resolution of the data was cut off at 2.4 Å, with a $CC_{1/2}$ of 1.0. This dataset was then phased with the Phenix suite Phaser.MR (molecular replacement) using the previous work structure of USP11_C318S_2GKG and ubiquitin as a search model (figure 1.6)(not yet deposited) (Maurer S. , 2021). After this, Phenix.autobuild was used to generate an initial model before iterative rounds of Phenix.Refine, Coot, and Pymol were used to assess the fit of the model to the data. This elucidated the location of the USP11 active site to not be between the two ubiquitin molecules but at the terminal end (figure 3.20), with the final structure analysis available in table 3.3.

Space Group	C 1 2 1
A, B, C	139.97, 52.51, 106.94
α, β, γ	90.00, 98.66, 90.00
Resolution(Å)	2.413
Observations	63,424
Unique Observations	31,826
R_{meas}	0.8636
I/σ(I)	6.43
$CC_{1/2}$	0.995
Completeness (%)	98.38
Multiplicity	2.0

R_{work}/R_{free}	0.2467 /0.3315
RMSD bond lengths	0.010
RMSD bond angles	1.68
Average B-factor	70.51

Table 3.3 USP11_C318S_D886G_2GKG and linear diubiquitin data processing and refinement statistics.

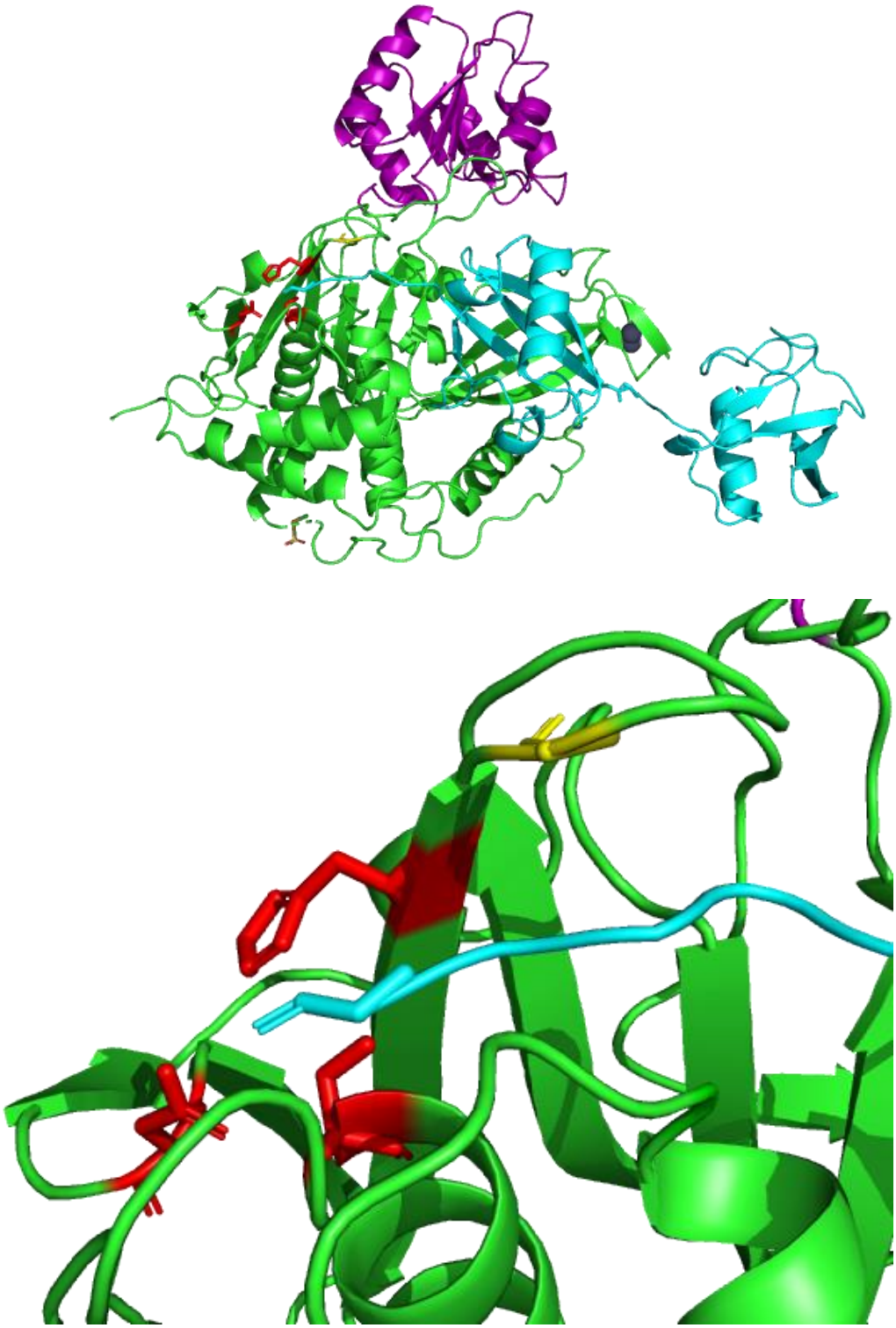

Figure 3.20 USP11_C318S_D886G_2GKG and linear diubiquitin co-crystallisation structure. This top structure shows the 2GKG tag (purple), the main USP11 sequence (green) with the diubiquitin (blue) associating with the catalytic triad (red) through the terminus of the diubiquitin molecule. This is as opposed to the catalytic triad interacting with the centre point between the two ubiquitin molecules. Below shows the electron density of the active site, where the catalytic triad residues (red) and the mutated D886G (yellow) interact with the substrate diubiquitin c terminal glycine residue(light blue).

4. Chapter 4: Discussion

4.1 Expression and Purification of USP11 proteins

Using previously defined *e.coli* recombinant expression protocol, the USP11 sequence was cloned into a pET28b plasmid and expressed according to protocol in methods. This produced a high purity sample with a large yield, enough to support future experiments. Following this, analysis of a previous structure along with sequence analysis (figure 1.7) was used to identify amino acid residues in the S1' site of USP11 to suggest amino acid which direct substrate specificity. These included the histidine H339 of the USP11 switching loop (SL), which had charged side chains for the poor diubiquitin cleavers USP11 and USP17 (H, E respectively) but polar uncharged amino acids for the efficient diubiquitin cleavers USP15 and USP4 (Q, Q). Other investigated targets included the D886 amino acid in the USP11 blocking loop 2 (BL2), and R885 in BL2.

These USP11 mutants also expressed with the same protocol, the amino acid changes not having an impact on bacterial recombinant production that warranted a change in protocol. These mutants when analysed through the Ub-AMC cleavage assay showed little difference with many of these mutants, which also included the USP11_2GKG sequence to assess whether the presence of this tag affected binding and substrate cleavage. However, the mutant with the biggest difference from the wildtype was the D886G mutant. This mutant, the proximally close R885G mutant, and the wildtype USP11 were selected for an increasing Ub-AMC concentration experiment to determine K_{cat}

from the initial reaction rates of these experiments. These determined that USP11D886G cleaved Ub-AMC quantitatively less efficiently than USP11wt. This was compared to linear di-ubiqiutin cleavage assays, which historically has been a poor substrate for USP11wt.

Quantitation of the diubiquitn and monoubiquitin bands (achieved through ImageJ analysis) revealed a semi-quantitative difference between USP11wt and USP11D886G. This difference is substantial due to USP11wt being a more efficient protease of Ub-AMC but USP11D886G being a more efficient protease of linear diubiqiutin. This was a clear indicator of the aspartic acid residue D886 (two amino acids from the catalytic histidine residue in the blocking loop 2 region) being a director of substrate specificity.

USP11D886G was one of the mutants with the lowest yield from recombinant expression, however it is unlikely this can be used as an indicator of downstream effects. Clean and clear peaks from the size exclusion gel filtration analysis indicates well-formed monomers were produced. The catalytic activity also indicates the protein was not denatured and that the amino acid residues were not interfering with the structural integrity of the proteins. This could have been confirmed either through structural analysis of X-ray crystallography or cryoEM, however these techniques often have many difficulties associated with them. Mass spec analysis could have been performed to confirm identity of these mutants instead but this option was not exercised due to the activity assay results.

To confirm the necessity of the aspartic acid residue, an analogue was chosen with a glycine residue in the same relative position. Such as USP15 which has a glycine residue in the same comparative position in blocking loop 2 (G860).

USP15 also showed relatively good yield and purity from the same bacterial recombinant expression protocol. Despite having slightly lower expression than USP11WT, the expression of USP15 was similar to several USP11 mutants and so further optimisation of USP15 expression was not deemed necessary. The low level of contaminant bands caused the use of 1 step histrap purification to be used (as opposed to the two step purification used for most other constructs). This did not cause any discernible issues with downstream analysis of USP15 or the USP15G860D construct. As the cleavage of linear diubiquitin was the substrate of interest, the decision to just perform USP15 diubiquitin cleavage assays was made. This semi quantitative assessment showed clearly that the reverse mutation showed the reverse results. i.e. for USP15, the wt sequence cleaved diubiquitin more efficiently than the USP15G860D mutant. This is contrast to USP11 where the USP11D886G mutant increases diubiquitin cleavage reaction rate.

It is from the combination of these results, the substrate selectivity of the USP11D886G mutation, and the inverse mutation for USP15 producing complementary data, that the hypothesis was formed: The residue in the S1' site (two amino acids from the catalytic histidine) has a direct impact on the substrate specificity of USP cleavage. This is related specifically to the two proteins in question (USP11 and USP15) but also potentially many other USPs in the family of proteins.

4.2 ITC and MST Analysis of USP11 and USP15

For further analysis and quantification of USP and ubiquitin substrate analysis, ITC runs and MST experiments were performed. This allowed calculation of K_D and standard deviation errors of these reactions. Additionally, all USP constructs used within ITC and MST were catalytic cysteine mutants to prevent cleavage of, but not interaction with, ubiquitin substrates. This was done to ensure signal clarity.

Initially, ITC was performed on the interaction between catalytically inactivated USP11 constructs (USP11C318S, USP11C318SD886G) and linear diubiquitin. This was compared to USP15 catalytically inactivated constructs (USP15C269S, USP15C269SG860D) and linear diubiqiutin. These graphs, showing examples taken from one of the 3 repeats, reveal little discernible differences (figure 3.7). However table 3.1 shows the output of K_D calculated using the Malvern microcal PeaQ analysis software.

The result of this, assessing just the interaction between USPs and linear diubiquitin showed USP11C318S having a lower affinity than USP11C318SD886G for diubiquitin cleavage, 2.2 µM ± 2.94 µM and 890 nM ± 466 nM respectively. The large errors for these results does call into question the validity of these results, but this does follow the trend of previous data which supports the hypothesis. However, the inverse mutation in USP15 does not support the hypothesis as concisely. This data showed USP15C269S and diubiquitin had a lower affinity for linear diubiquitin than USP15C269SG860D, 5.6 µM ± 1.47 µM and 3.32 µM ± 410 nM. This data had relatively large standard

deviation errors as well, but the errors were much lower proportionally than the USP11 counterparts.

The USP11 constructs were also assayed for their interaction with single ubiquitin constructs paired to amino acids. The USP11 mutants with ubiquitin-G, ubiquitin-GGG, and ubiquitin-M also showed conflicting results. When comparing the inactive mutant (USP11C318S) with the additional substrate specific mutant (USP11C318SD886G), the substrate UbM showed a lower affinity for the catalytic inactive only mutant (1.43 µM ± 154 nM against 577 nM ± 136 nM respectively). But the UbG substrate (51.2 nM ± 27.9 nM against 119 nM ± 140 nM) and the UbGGG substrate (29.9 nM ± 19.2 nM against 39.0 nM ± 16.9 nM) both showed a higher affinity for the inactive mutant against the inactive mutant with the additional substrate mutation (USP11C318S against USP11C318SD886G respectively). The USP11 interactions with UbM were the most reliable runs with the lowest proportional standard deviation error amongst the experiments, with the other substrates having a standard deviation error of at least 50 % of the overall K_D.

In addition to this, USP11C318S had the lowest affinity for linear diubiquitin (2.2 µM) with the closest in affinity being USP11C318S with UbM (1.43 µM). This falls within USP11 having a poor cleavage of linear diubiquitin, however the difference between these K_D constants does not seem as drastic as the results from the diubiqiutin cleavage assays, where USP11wt does not completely break down linear diubiquitin even after 1,200 minutes of reaction time (20 hours/overnight reaction). There are several possible reasons for this. ITC utilises catalytically inactive protein as it just measures the temperature change in the sample from association of the two components (protein and substrate).

To prevent secondary binding and on/off interactions creating additional temperature changes and causing ambiguity in the graphs, catalytically inactive proteins were produced. This means that the ITC measured association of substrate, whereas the diubiquitin cleavage and Ub-AMC assays measured cleavage of substrate. In this way, USP11 could have a higher affinity for binding with linear diubiquitin than it does for cleavage of linear diubiqiutin. This could explain the non linear relationship between the results of these two groups of experiments. The proximity of the D/G amino acid in question to the catalytic histidine residue may not prevent association of the substrate to the protein, but could still interfere with docking specifically to the catalytically critical residues.

Regardless of these arguments, looking solely at the diubiquitn cleavage of the two assays (ITC and diubiquitin cleavage). The K_D from USP11C318S (2.2 µM) and USP11C318SD886G (890 nM) shows a difference by a factor of roughly 2.5 times. Taking the quantification of the diubiquitin cleavage assay bands, at timepoint 240 minutes, the % of remaining diubiquitin was ~72.7% for the USP11wt and ~ 52.7% for the USP11D886G mutant. This is a difference factor of roughly 1.4 times. By these calculations the ITC predicts a larger difference in substrate cleaved than observed from the gel cleavage assays. The other substrates all having a higher affinity for USP11 than linear diubiquitin also supports that this mutant allows targeting for specific substrates. However, the differences between these substrates and the D/G mutations show still substantial differences, with the possible exception of the UbGGG where both results are overlayed by their standard deviation error margins. This suggests the D/G mutation could even interact with small substrates depending on their

composition, affecting some as drastically as the linear diubiquitin substrates (such as UbM), and affecting some far less (UbGGG).

It is possible that, for the USP15 experimentation, which had a larger difference than USP11 when it came to the effect of the D/G mutation on cleavage of linear diubiquitin, that ITC does not accurately reflect substrate cleavage. In the ITC experiments, the USP15 catalytic only mutant had a lower affinity than the reverse substrate specific mutant. This is adverse to the results of the linear diubiquitin cleavage assays whereby USP15 showed higher cleavage activity compared to its USP15G860D mutant. The combination of these two groups of experiments suggests that the G/D mutant has a larger effect on substrate cleavage than it does on substrate association, especially when taking into account the low standard deviation errors on both experimental results.

The MST data additionally generated results with different implications. The average of the complete data set, taking into account the temperature jump (t-jump) and thermophoresis, showed USP11C318S having tighter binding with UbGGG than with linear diubiquitin (29.3 µM ± 30.1µM and 62.3 µM ± 86.5 µM respectively). However, similar to the ITC results, these results contain large proportional standard deviation errors. These results were the result of three experimental repeats which each contained three sample repeats. These results were also largely different from their ITC counterparts by a factor of at least 10. These results would ideally give the same K_D results as the dissociation constant should not change due to the method of calculation. MST analyses the fluorescence of a temperature sensitive fluorophore (usually

decreasing intensity with increasing temperature), and calculates through movement of fluorophore out of the path of the excitation laser, the molecular size of the attached target protein of interest (relatively to itself with increasing addition of substrates). Through this, MST can be used to calculate dissociation constants in different ways, including how the initial drop in read fluorescence (t-jump) can give different results when comparing this to extended thermophoresis reading. T-jump usually indicates the impact of the close vicinity to the fluorophore, whereas thermophoresis has an impact associated with the wider environment. Comparing these two analysis methods can be used to determine binding interactions and stoichiometry. This is evident in figure 3.9 showing the increased variation in standard deviation error for thermophoresis analysis. However, the trend for both methods are very similar in these instances, and it was thought that ubiquitin constructs had a 1:1 binding stoichiometry with both USP11 and USP15 mutants.

Additional analysis comes in the form of what is suspected to be two steps in the associated graph. This is highlighted in the figures 3.10 and 3.11 (with data in table 3.2). These showed that removal of a third of the data points can change the plateau point of the sigmoidal curve for USP11C318S and UbGGG. However, this data is likely not as reliable as the complete data set as the error/signal ratio is significantly higher, the thermophoresis analysis can no longer fit to some of the data, and the removal of data points lowers the reliability and validity of most experimentation (without an argument for excluding datapoints as outliers).

The question of different curves from one dataset becomes more prevalent when assessing the USP11C318S analysis of linear diubiquitin. These points,

when analysed show that two different curves can be observed with a plateau where they meet in the centre of the data set. This curved nature of the binding curve could be indicative of biphasic binding (or at the least the absence of monophasic binding). Whereby the creation of protein folding intermediates can disrupt or compete with the protein-substrate complex formation (references insert). The presence of this phenomenon would explain the low affinity K_D (compared to ITC) as well as the high standard deviation error of the final constant. It was also decided from the presence of this data to focus on different analysis methods which could produce more reliable dissociation constants.

The MST data also led to further ITC study. It was noted during experimentation (and seen in table 3.1) that the binding sites (n) of the USP11/USP15 interactions with ubiquitin substrates was rarely at the expected n of 1. With binding sites ranging from 0.5 (USP11 and UbM) to 1.5 (USP11 and UbG) and even up to 1.9 (USP11 and UbGGG). It was thought one possible explanation of this was the binding of additional amino acid tails to the proteases. To investigate this, a diubiquitin construct was created with the removal of the ubiquitin tail from the second ubiquitin, diUbL149X. This was run with ITC to investigate the impact of this extraneous ubiquitin tail on the binding dissociation constant and the binding sites, when binding with USP11 constructs (table 3.1).

This met with mixed results, the USP11C318S mutant with diUbL149X giving a K_D of 13.8 µM ± 1.35 µM. USP11C318SD886G with diUbL149X gives a K_D of 8.37 µM ± 1.63 µM. These experiments give a lower affinity than the linear diubiquitin counterparts (2.2 µM ± 2.94 µM and 890 nM ± 466 nM respectively). With a higher K_D by a factor of roughly 10. The diUbL149X ITC experiments

had a lower standard deviation error than the linear diubiquitin experiments, indicating the higher reliability of these K_D outputs from the L149X constructs. This could mean that the extended tail of the second ubiquitin on the diubiquitin construct could aid with association and binding between diubiquitin and USPs. However, the removal of the N-terminal amino acids did little to elucidate the binding sites in the ITC experiments and their difference from the predicted n of 1. USP11C318S and diubiquitin had sites n 0.844 ± 0.617. USP11C318S and diUbL149X had sites n 0.398 ± 0.162. USP11C318SD886G and diubiquitin had sites n 0.387 ± 0.143. USP11C318SD886G and diUbL149X had sites n 0.631 ± 0.0686. This shows a wide variation, with the largest n sites and largest standard deviation error for USP11C318S and diubiquitin with n sites 0.844. The lowest n sites for USP11C318SD886G and diubiquitin. However there is no definitive relationship dividing diubiquitin and the diubiquitin L149X construct in terms of binding sites predicted by ITC.

This data indicates further ITC or MST would not further elucidate the binding mechanisms between USPs and their ubiquitin substrates without changes in experimental planning. Perhaps with different ubiquitin constructs to engineer a 1:1 binding ratio. Additional methods to determine the nature of the binding between USPs and ubiquitin substrates, SPR or NMR could be used to better analyse the precise mechanisms. However, these were not performed as the impact of the D/G mutant had already been clearly demonstrated for diubiquitin substrates. Diubiquitin showed an increased affinity with USP11 when the D886G mutation was present.

There still remained ambiguity about the action of the USP11 and diubiquitin binding. An additional substrate was used to determine whether the specificity

of the D/G mutation was due entirely to size of the substrate creating stoichiometric interference. K63 linked diubiquitin (K63diUb) was decided on as a substrate for diubiqiutin cleavage assays. This assay was chosen for unambiguous cleavage data, and substrate requirements, the cleavage assays requiring the smallest volume and concentration of the assessment experiments. Additionally, K63diUb has been previously identified as a physiological substrate of USP11 (Guo, Tang, Yuan, Zhang, & Wang, 2022). The cleavage of this USP11 substrate was decreased when the D/G mutation was introduced, in the inverse trend to the USP11 D/G mutational effect with linear diubiquitin. This conforms to the hypothesis that the D/G mutation creates effects which are substrate specific, even within diubiquitin targets. This hypothesis is supported through the different chains of ubiquitin which are differentially targeted and cleaved by different members of the USP family.

4.3 USP17 Expression and Purification

Another USP was chosen to assist confirmation of the D/G mutational effect. This was decided to be USP17 which had a singular catalytic domain (as opposed to D1D2 domains separated by a linker region) and has a glycine residue in the sequence two amino acids before the catalytic histidine residue. USP17 had additionally previously shown to be a poor cleaver of linear diubiquitin (figure 1.7).

This began with a GST-tagged construct. However, the expression of this construct left a large number of contamination bands in the SDS-PAGE analysis, so further optimisation was undertaken. A smaller his-tagged construct was used which produced similar issues.

Further avenues were investigated, an *e.coli* phosphokinase PDB.ID:1BKR was used as a tag in an attempt to increase protein production and stability. This was cloned into a USP17 catalytic core construct (figure 3.15). This met with sufficient results for purity with still low protein yield. The full length construct was still investigated for protein production. This began with using the pcold1 vector to allow for protein production over extended time periods (three days, details in methods). Further media supplements were used to increase protein production. These included increasing a percentage v/v of ethanol in the media which has been hypothesised as increasing protein yield by causing activation of heat shock proteins in bacterial cell systems (Chhetri, Kalita, & Tripathi, 2015). Sorbitol was added to the media as well as a common substrate for decreasing *E.coli* growing time, which has been used to increase yield and stability of cytoplasmic and membrane proteins (Azadi, Mahboubi, Naghdi, Solaimanian, & & Mortazavi, 2017). This caused an increase in band thickness but not a substantial protein yield increase.

With the highest success for protein yield being obtained from protein tags. This was further investigated with other protein tags identified in the previous work section (Caulton, 2016) (Maurer S. , 2021). These tags, named after their PDB-IDs, were the 1W41 and 2GKG loop stabilising insertion tags, and the 3WVA dimeric tag (figure 3.16). From these, the 2GKG insertion tag produced the highest yield and purity.

This yield was still relatively low (compared to protein requirements and USP11/USP15 protein production). Other attempts to increase the production yield of bacterial recombinant protein expression of USP17 led to investigation of media composition. This final trial of autoinduction media, with a pET28D

plasmid, the USP172GKG construct, and low temperature induction led to the highest yield (~0.2 mg of protein from 1.5 L of media). The high purity of this sample (figure 3.17) was deemed appropriate enough to attempt crystallisation trials. This created some conditions with noticeable aggregation of sample, but still required a large percentage of protein produced to create. These aggregation wells were not formed well enough to warrant further optimisation trials. This, coupled with the low production yield of USP17 led to USP11 and USP15 being the main targets of this investigation. However future USP17 investigation is still possible.

4.4 X-Ray Crystallography of USP11 and linear diubiqiutin

The construct chosen for crystallographic studies was picked based on previous work. USP11 was picked as the focus due to previous structures and the direct comparison that could be drawn between a monoubiquitin and a di-ubiquitin complex structure. 2GKG as a crystallisation tag was picked for the previous success in aiding crystallisation (Maurer S. , 2021). The C318SD886G mutations were picked to prevent diubiquitin cleavage and to assist in USP11 diubiquitin binding, leading to the full complex being USP11C318SD886G_2GKG.

Coexpression of this USP11 construct with linear diubiquitin led to a sufficient yield (0.7 mg/2.4L culture) for sparse matrix screening trays. The SDS-PAGE gels showed that both the USP11 construct and the diubiquitin were co-eluting from the purification steps with sufficient purity.

Sparse matrix screens were picked for a variety of pre-detemined conditions. Extra zinc was added to every sparse matrix screen condition due to the zinc

binding domain present in USP11 as well as the cadmium ion identified in the previous structure (Maurer S. , 2021). This eventually yielded crystals which were visible in the visual light and UV light spectrum (figure 3.19). This indicates that they were protein crystals due to their fluorescence, attributed to the aromatic ring residues in the sequences. There are some conditions in sparse matrix commercial screens which have compounds fluorescent under the same conditions, which means there is still the possibility for non-protein crystals in some wells. However, this was not the case in the conditions which yielded crystalline objects.

Figure 3.19 shows the crystal forms observed from some of the most promising wells. Whilst some of these conditions had crystals which were picked and taken to the synchrotron at Diamond light source, other conditions had optimisation screens set up, such as the LMB screen, well A1. However, in response to the quality of the data set received from the H3 condition of the Morpheus II screen, and the lack of precipitate forming in optimisation screens, the optimisation screens were not continued. Instead the data was analysed. The data set with the highest quality, a diffraction image of which is shown in figure 3.19, did contain an obvious ice ring. This can cause problems for data analysis or even prevent structure solution (Thorn, et al., 2017). However, utilising the software pipelines of CCP4 and Phenix, a model was still produced.

The results of this in figure 3.20 revealed a different binding form than expected. The placement of the terminal glycine residue of the linear diubiquitin construct was in the USP11 catalytic domain. This was as opposed to the glycine-methionine link between the two ubiquitin molecules being present in the USP11 active site.

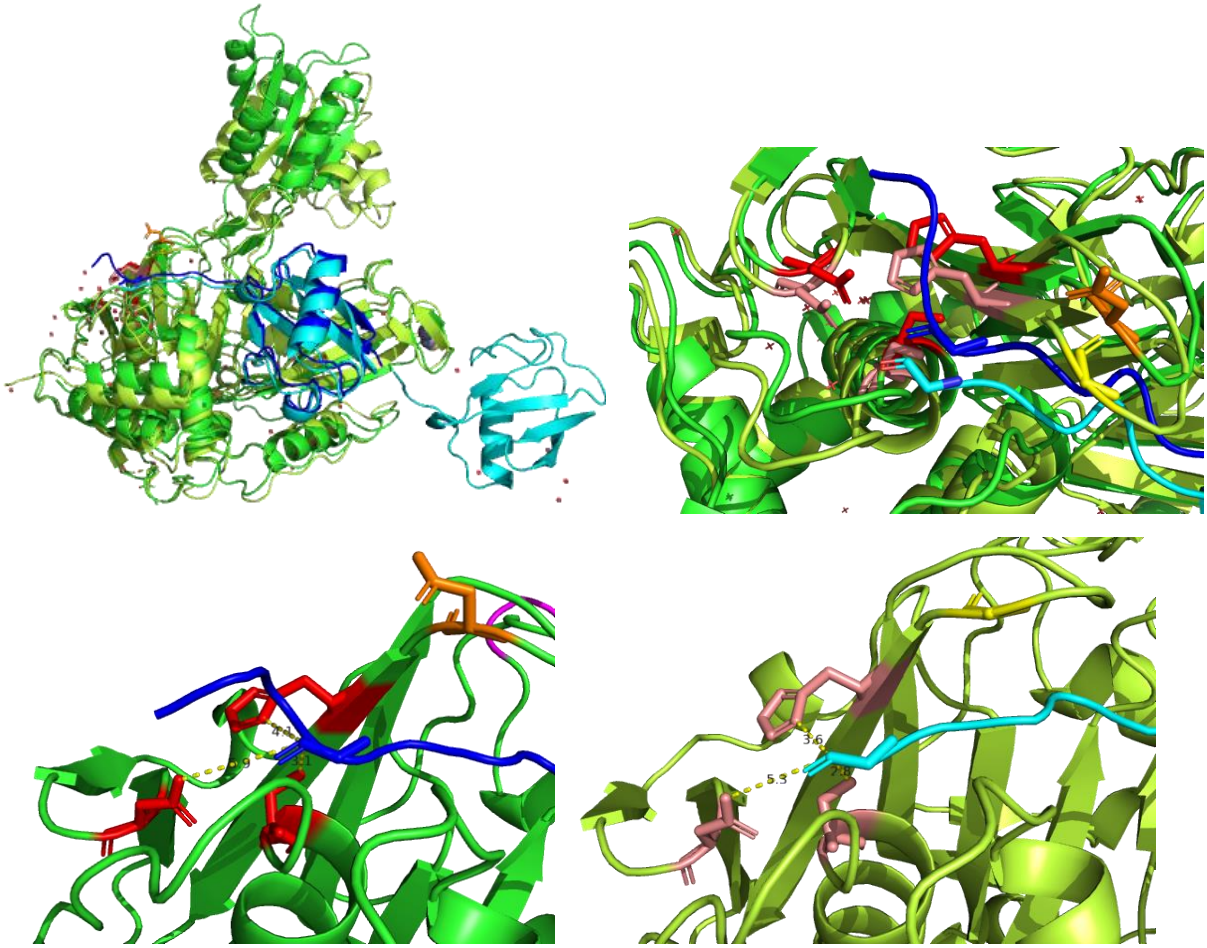

Figure 4.1 USP11_C318S_D886G_2GKG and linear diubiquitin structure compared with previous USP11_C318S_2GKG and mono ubiquitinGGG. The top two structures features an overlay of the novel USP11 + diubiquitin structure with the previous USP11 + monoubiquitin structure. The main USP11 sequences are in green (dark green previous structure, light green novel), with the ubiquitin molecules in blue (dark blue previous structure, light blue novel), the catalytic triad residues in red (red for previous structure, pink for novel), with the 886 mutation subject residue in orange (orange for previous structure, yellow for novel). The bottom images show the structures in isolation, comparing their catalytic domains. With bottom left in darker colours showing the previous monoubiqiutin structure, the bottom right showing the novel diubiquitin structure. The USP11 diubiquitin structure with the D886G mutation shows a closer association between the catalytic residues and the substrate, with 3.6 Å, 2.8 Å, and 5.3 Å (for catalytic histidine, cysteine/serine mutant, and aspartic acid residues respectively) compared to 4.1 Å, 3.1 Å, and 5.9 Å in the monoubiquitin D886 structure.

Figure 4.1 shows a comparison between this novel structure of USP11_C318S_D886G_2GKG with diubiquitin and a previously elucidated structure of USP11_C318S_2GKG with monoubiquitinGGG. Whilst the extra glycine resiudes were added to aid binding with monoubiquitin (Maurer S. , 2021), the diubiquitin structure shows the ability to bind to the terminal glycine residue without additional amino acids. Whilst the diubiquitin structure with the D886G mutation does show a closer conformation between the catalytic triad residues and the diubiquitin substrate, it is possible that the extra glycine tail on the monoubiquitin structure is responsible for this. Otherwise, there is a remarkable similarity between the two structures. This also assists in justifying the model produced of the USP11 construct with diubiquitin structure. As in table 3.3, the R_{work} factor is 0.2467. This is relatively high with most structures aim for an R factor of <0.2. The R factor is a measure of how well the model fits the data, a higher number indicates a “worse” fit.

The structural information should be compared to the previous assay information. For example the diubiquitin cleavage assays indicated a different binding form. This is due to the SDS-PAGE gel indicating cleavage of the diubiquitin into two almost equal monoubiquitin molecules. This was supported by the pattern present in the SDS-PAGE gels, whereby one large molecular weight band (~14 kDa) over time became one lower molecular weight band (~7 Kda) (figure 3.4). The presence of no more than two bands at any timepoint does not support an idea of amino acids being trimmed from the distal ubiquitin molecule. This indicated a binding conformation where the USP11 construct can interact at the link between the two ubiquitin molecules. But this did not eliminate the possibility of binding at the distal glycine residue of the second

ubiquitin molecule, nor the possibility of amino acids being trimmed from the distal end. This is due to the lack of sensitivity for SDS-PAGE gels to detect small molecular weight changes. Additionally this assay only showed cleavage of the diubiquitin molecules and did not present binding information between the two protein constructs.

The ITC and MST experiments presented new data with novel challenges . This included ambiguous results in terms of diubiquitin binding and the effect of different USP11 constructs. The ITC showed differences from what was expected most predominantly in the number of binding sites. This was present across all ubiquitin substrates, with the monoubiquitin substrates having a higher number of sites, (>1, up to 1.9). This indicates that up to two monoubiquitin molecules could bind to one USP molecule for these experiments. This is contrary to many diubiquitin substrates which had a lower number of binding sites than expected (<1, down to 0.5). This would indicate that each USP11 could bind two diubiquitin molecules, and would reach the saturation faster than expected.

There is not enough information to state definitively, however the possibility remains that if there are two binding sites on one diubiquitin molecule, then two USP11 proteins could bind it simultaneously. This would explain the large variation in N-sites received from the ITC experiments (deviating away from the 1:1 binding ratio expected).

The linear diubiquitin construct L149X, terminates the distal ubiquitin flexible C-terminus. This was engineered to investigate this possibility. However the results did not change significantly from the full length diubiquitin construct. The

only monoubiquitin substrate which presented with a lower N-site result was the Ubi-M construct. This terminal residue may have more impact on the ability of USP11 to bind the substrate than the length of the ubiquitin chain. Either a more severe removal of the C terminus, or a less hospitable amino acid mutation should be investigated.

The presence of these multiple binding sites could also explain the imprecision of the MST data, as the technique is better suited to more simplified binding interactions without multiple binding phases. Other experiments such as SPR should be used in future to investigate this.

4.5 Future Work

This investigation consists of data which contributes to the scientific community. An example of this is validation of the effect of the DtoG mutation. This would be completed through investigation with diubiquitin constructs engineered to inhibit binding through the terminal glycine residues. Either with L149X, removing the terminal glycine residues, or mutating the terminal residues to amino acids which will prevent this reaction such as L149M_R150X. This construct would be used with the current USP11 constructs in ITC to be compared to previous results. Additionally SPR experiments with immobilised USP11 could be used to give more detailed binding data (Tiwari, et al., 2021). Additionally, given the success of the 2GKG tag to aid crystallisation, this could be used with the diubiquitin L149X and USP11 constructs for structure determination. Further work would also involve USP17 constructs with the equivalent mutation for gel based diubiquitin

cleavage assays and other ITC experimentation. This would help confirm this mutation effect for USPs without the UBL insertion between the catalytic domains. Additional future work would involve using ubiquitinated targets specific to USP11, USP15, or USP17. Such work would include mutating a ubiquitinated lysine residue into a cysteine to introduce cysteine bonding to mimic ubiquitination and interaction with USPs.

4.6 Conclusion

This thesis supports that there are key residues directing substrate specificity for USPs. The aspartic acid residue two amino acids before the catalytic histidine can direct the cleavage of linear diubiquitin, with the reverse effect on K63 linked diubiquitin. This would help direct drug molecules targeting specific interactions between USPs and single targets, with the potential for fewer off target effects than targeting the catalytic triad completely.

However, more investigation needs to be completed to elucidate the interaction between USP11 and USP15 with linear diubiquitin. The structure solution of an unexpected distal ubiquitin tail binding site for the USPs reveals the need for more specialised ubiquitin constructs. Other techniques such as SPR for binding mechanistic studies would also help elucidate the binding mechanism between catalytically inactive USPs and ubiquitin substrates, which presented as more complexed than previously thought.

Other USPs can also be investigated, such as the USP17 construct expressed and purified in this work.

Appendix 1: Sparse Matrix Screen Conditions Morpheus (I)

Table 1: List of PDB ligands in Morpheus®

PDB Ligand name(s)	Class	PDB ID(s)	Number of Structures*
1,2-Ethanediol (ethylene glycol)	Precipitant	EDO, EGL	1081
1,2-Propanediol (enantiomers R and S)	Alcohols	PGO, PGR	41
1,3-Propanediol	Alcohols	PDO	7
1,4-Butanediol	Alcohols	BU1	11
1,6-Hexanediol	Alcohols	HEZ	19
1-Butanol	Alcohols	1BO	7
2-(N-Morpholino)-ethane sulfonic acid (MES)	Buffer	MES	315
2-Amino-2-hydroxymethyl-propane-1,3-diol (Tris)	Buffer	TRS	334
2-Methyl-2,4-pentanediol (MPD, enantiomers R and S)	Precipitant	MPD, MRD	504
3-Morpholinopropane-1-sulfonic acid (MOPS)	Buffer	MPO	21
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Buffer	EPE	201
Acetic acid, acetate, acetyl	Carboxylic acids	ACY, ACT, ACE	1890
(*S*)-2-Aminopropanoic acid (Alanine, (enantiomers L and D)	Amino acids	ALA, DAL	35
Amino, Ammonia, Ammonium	multiple	NH2, NH3, NH4	582
N,N-bis(2-hydroxyethyl)glycine (Bicine)	Buffer	BCN	13
Bromide	Halogens	BR	120
Calcium	Divalents	CA	3959
Chloride	Multiple	CL	2842
Citric acid, citrate	Carboxylic acids	CIT, FLC	384
D-Galactose (anomers α and β)	Monosaccharides	GAL, GLA	86
D-Glucose (anomers α and β)	Monosaccharides	GLC, BGC	206
Glutamic acid (enantiomers L and D)	Precipitant	GLU, DGL	75
Di(Hydroxyethyl)ether (Di-Ethyleneglycol)	Ethylene glycols	PEG	209
D-Mannose (anomers α and β)	Monosaccharides	MAN, BMA	178
D-Xylopyranose (anomers α and β)	Monosaccharides	XYL, XYP	41
Fluoride	Halogens	F	16
Formic acid	Carboxylic acids	FMT	267
Glycerol	Amino acids	GOL	2884
Glycine	Buffer	GLY	50
Imidazole	Halogens	IMD	154
Iodide	Alcohols	IOD	178
Isopropyl alcohol (iso-propanol, 2-Propanol)	Monosaccharides	IPA, IOH	174
L-Fucose (anomers α and β)	Amino acids	FUC, FUL	62
Lysine (enantiomers L and D)	Amino acids	LYS, DLY	36
Magnesium	Divalents	MG	3991
N-Acetyl-d-glucosamine (anomers α and β)	Monosaccharides	NAG,NBG	1150
Nitrate	NPS	NO3	156
Oxamic acid	Carboxylic acids	OXM	17
Penta(hydroxyethyl)ether (Penta-Ethyleneglycol)	Ethylene glycols	1PE	91
Phosphates	NPS	PO4, PI, 2HP	1687
Potassium	Carboxylic acids	K	720
Serine (enantiomers L and D)	Amino acids	SER, DSN	38
Sodium	multiple	NA	1926
Sulfate	NPS	SO4	5793
Tartaric acid (enantiomers R and S)	Carboxylic acids	TAR, TLA	113
Tetra(hydroxyethyl)ether (Tetra-Ethyleneglycol)	Ethylene glycols	PG4	194
Tri(Hydroxyethyl)ether (Tri-Ethyleneglycol)	Ethylene glycols	PGE	107
		SUM	32956

Continued on Next Page

Table 2: Mixes of additives used in Morpheus®

Mix name	Composition	Catalogue Number (100 mL)	Catalogue Number (250 mL)
Divalents	0.3M Magnesium chloride hexahydrate; 0.3M Calcium chloride dihydrate	MD2-100-70	MD2-250-70
Halogens	0.3M Sodium fluoride; 0.3M Sodium bromide; 0.3M Sodium iodide	MD2-100-71	MD2-250-71
NPS†	0.3M Sodium nitrate, 0.3 Sodium phosphate dibasic, 0.3M Ammonium sulfate	MD2-100-72	MD2-250-72
Alcohols	0.2M 1,6-Hexanediol; 0.2M 1-Butanol 0.2M 1,2-Propanediol; 0.2M 2-Propanol; 0.2M 1,4-Butanediol; 0.2M 1,3-Propanediol	MD2-100-73	MD2-250-73
Ethylene glycols	0.3M Diethylene glycol; 0.3M Triethylene glycol; 0.3M Tetraethylene glycol; 0.3M Pentaethylene glycol	MD2-100-74	MD2-250-74
Monosaccharides	0.2M D-Glucose; 0.2M D-Mannose; 0.2M D-Galactose; 0.2M L-Fucose; 0.2M D-Xylose; 0.2M N-Acetyl-D-Glucosamine	MD2-100-75	MD2-250-75
Carboxylic acids	0.2M Sodium formate; 0.2M Ammonium acetate; 0.2M Sodium citrate tribasic dihydrate; 0.2M Potassium sodium tartrate tetrahydrate; 0.2M Sodium oxamate	MD2-100-76	MD2-250-76
Amino acids	0.2M DL-Glutamic acid monohydrate; 0.2M DL-Alanine; 0.2M Glycine; 0.2M DL-Lysine monohydrochloride; 0.2M DL-Serine	MD2-100-77	MD2-250-77

Table 3: Buffer systems used in Morpheus®

Mix name	Conc.	pH @ 20°C	Composition	Catalogue Number (100 mL)	Catalogue Number (250 mL)
Buffer System 1	1.0M	6.5	Imidazole; MES monohydrate (acid)	MD2-100-100	MD2-250-100
Buffer System 2	1.0M	7.5	Sodium HEPES; MOPS (acid)	MD2-100-101	MD2-250-101
Buffer System 3	1.0M	8.5	Tris (base); BICINE	MD2-100-102	MD2-250-102

Table 4: Mixes of Precipitants used in Morpheus®

Mix name	Old Mix Name	Composition	Catalogue Number (100 mL)	Catalogue Number (250 mL)
60% Precipitant Mix 1	P500MME_P20K	40% v/v PEG 500* MME; 20 % w/v PEG 20000	MD2-100-81	MD2-250-81
60% Precipitant Mix 2	EDO_P8K	40% v/v Ethylene glycol; 20 % w/v PEG 8000	MD2-100-82	MD2-250-82
60% Precipitant Mix 3	GOL_P4K	40% v/v Glycerol; 20% w/v PEG 4000	MD2-100-83	MD2-250-83
~~75% Precipitant Mix 4~~	~~MPD_P1K_P3350~~	25% v/v MPD; 25% PEG 1000; 25% w/v PEG 3350	~~MD2-100-84~~	~~MD2-250-84~~

*The PEG 550 MME that was originally used in this screen has been discontinued and replaced with PEG 500 MME.

Continued on Next Page

Screen ID	Well #	Conc	Ligands	Conc	Buffer	pH	Conc	Precipitant
1-1	A1	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 1
1-2	A2	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 2
1-3	A3	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 3
1-4	A4	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	37.5 % v/v	Precipitant Mix 4
1-5	A5	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 1
1-6	A6	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 2
1-7	A7	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 3
1-8	A8	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	37.5 % v/v	Precipitant Mix 4
1-9	A9	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 1
1-10	A10	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 2
1-11	A11	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 3
1-12	A12	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	37.5 % v/v	Precipitant Mix 4
1-13	B1	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 1
1-14	B2	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 2
1-15	B3	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 3
1-16	B4	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	37.5 % v/v	Precipitant Mix 4
1-17	B5	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 1
1-18	B6	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 2
1-19	B7	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 3
1-20	B8	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	37.5 % v/v	Precipitant Mix 4
1-21	B9	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 1
1-22	B10	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 2
1-23	B11	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 3
1-24	B12	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	37.5 % v/v	Precipitant Mix 4
1-25	C1	0.09 M	NPS	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 1
1-26	C2	0.09 M	NPS	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 2
1-27	C3	0.09 M	NPS	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 3
1-28	C4	0.09 M	NPS	0.1 M	Buffer System 1	6.5	37.5 % v/v	Precipitant Mix 4
1-29	C5	0.09 M	NPS	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 1
1-30	C6	0.09 M	NPS	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 2
1-31	C7	0.09 M	NPS	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 3
1-32	C8	0.09 M	NPS	0.1 M	Buffer System 2	7.5	37.5 % v/v	Precipitant Mix 4
1-33	C9	0.09 M	NPS	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 1
1-34	C10	0.09 M	NPS	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 2
1-35	C11	0.09 M	NPS	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 3
1-36	C12	0.09 M	NPS	0.1 M	Buffer System 3	8.5	37.5 % v/v	Precipitant Mix 4
1-37	D1	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 1
1-38	D2	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 2
1-39	D3	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 3
1-40	D4	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	37.5 % v/v	Precipitant Mix 4
1-41	D5	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 1
1-42	D6	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 2
1-43	D7	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 3
1-44	D8	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	37.5 % v/v	Precipitant Mix 4
1-45	D9	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 1
1-46	D10	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 2
1-47	D11	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 3
1-48	D12	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	37.5 % v/v	Precipitant Mix 4

Continued on Next Page

Tube	Well	Conc	Ligands	Conc	Buffer	pH	Conc	Precipitant
2-1	E1	0.12 M	Ethylene glycols	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 1
2-2	E2	0.12 M	Ethylene glycols	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 2
2-3	E3	0.12 M	Ethylene glycols	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 3
2-4	E4	0.12 M	Ethylene glycols	0.1 M	Buffer System 1	6.5	37.5 % v/v	Precipitant Mix 4
2-5	E5	0.12 M	Ethylene glycols	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 1
2-6	E6	0.12 M	Ethylene glycols	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 2
2-7	E7	0.12 M	Ethylene glycols	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 3
2-8	E8	0.12 M	Ethylene glycols	0.1 M	Buffer System 2	7.5	37.5 % v/v	Precipitant Mix 4
2-9	E9	0.12 M	Ethylene glycols	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 1
2-10	E10	0.12 M	Ethylene glycols	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 2
2-11	E11	0.12 M	Ethylene glycols	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 3
2-12	E12	0.12 M	Ethylene glycols	0.1 M	Buffer System 3	8.5	37.5 % v/v	Precipitant Mix 4
2-13	F1	0.12 M	Monosaccharides	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 1
2-14	F2	0.12 M	Monosaccharides	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 2
2-15	F3	0.12 M	Monosaccharides	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 3
2-16	F4	0.12 M	Monosaccharides	0.1 M	Buffer System 1	6.5	37.5 % v/v	Precipitant Mix 4
2-17	F5	0.12 M	Monosaccharides	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 1
2-18	F6	0.12 M	Monosaccharides	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 2
2-19	F7	0.12 M	Monosaccharides	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 3
2-20	F8	0.12 M	Monosaccharides	0.1 M	Buffer System 2	7.5	37.5 % v/v	Precipitant Mix 4
2-21	F9	0.12 M	Monosaccharides	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 1
2-22	F10	0.12 M	Monosaccharides	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 2
2-23	F11	0.12 M	Monosaccharides	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 3
2-24	F12	0.12 M	Monosaccharides	0.1 M	Buffer System 3	8.5	37.5 % v/v	Precipitant Mix 4
2-25	G1	0.1 M	Carboxylic acids	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 1
2-26	G2	0.1 M	Carboxylic acids	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 2
2-27	G3	0.1 M	Carboxylic acids	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 3
2-28	G4	0.1 M	Carboxylic acids	0.1 M	Buffer System 1	6.5	37.5 % v/v	Precipitant Mix 4
2-29	G5	0.1 M	Carboxylic acids	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 1
2-30	G6	0.1 M	Carboxylic acids	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 2
2-31	G7	0.1 M	Carboxylic acids	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 3
2-32	G8	0.1 M	Carboxylic acids	0.1 M	Buffer System 2	7.5	37.5 % v/v	Precipitant Mix 4
2-33	G9	0.1 M	Carboxylic acids	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 1
2-34	G10	0.1 M	Carboxylic acids	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 2
2-35	G11	0.1 M	Carboxylic acids	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 3
2-36	G12	0.1 M	Carboxylic acids	0.1 M	Buffer System 3	8.5	37.5 % v/v	Precipitant Mix 4
2-37	H1	0.1 M	Amino acids	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 1
2-38	H2	0.1 M	Amino acids	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 2
2-39	H3	0.1 M	Amino acids	0.1 M	Buffer System 1	6.5	30 % v/v	Precipitant Mix 3
2-40	H4	0.1 M	Amino acids	0.1 M	Buffer System 1	6.5	37.5 % v/v	Precipitant Mix 4
2-41	H5	0.1 M	Amino acids	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 1
2-42	H6	0.1 M	Amino acids	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 2
2-43	H7	0.1 M	Amino acids	0.1 M	Buffer System 2	7.5	30 % v/v	Precipitant Mix 3
2-44	H8	0.1 M	Amino acids	0.1 M	Buffer System 2	7.5	37.5 % v/v	Precipitant Mix 4
2-45	H9	0.1 M	Amino acids	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 1
2-46	H10	0.1 M	Amino acids	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 2
2-47	H11	0.1 M	Amino acids	0.1 M	Buffer System 3	8.5	30 % v/v	Precipitant Mix 3
2-48	H12	0.1 M	Amino acids	0.1 M	Buffer System 3	8.5	37.5 % v/v	Precipitant Mix 4

Appendix 2: Sparse Matrix Screen Conditions Morpheus (II)

Table 1: List of PDB ligands in Morpheus II

PDB ligand name	Class	PDB ID (main)	No. of structures*
Lithium sulfate	Common salt	LI	51
Sodium chloride	Common salt	NA	4726
Potassium sulfate	Common salt	K	1638
Manganese chloride tetrahydrate	Divalent cation	MN	1938
Cobalt chloride hexahydrate	Divalent cation	CO	474
Nickel chloride hexahydrate	Divalent cation	NI	699
Zinc acetate dihydrate	Divalent cation	ZN	8413
Barium acetate	Alkali	BA	91
Cesium acetate	Alkali	CS	75
Rubidium chloride	Alkali	RB	34
Strontium acetate	Alkali	SR	101
Sodium chromate tetrahydrate	Oxometalate	CR	7
Sodium molybdate dihydrate	Oxometalate	MOO	20
Sodium orthovanadate	Oxometalate	VO4	73
Sodium tungstate dihydrate	Oxometalate	WO4	47
Erbium (III) chloride hexahydrate	Lanthanide	ER3	2
Terbium (III) chloride hexahydrate	Lanthanide	TB	11
Ytterbium (III) chloride hexahydrate	Lanthanide	YB	57
Yttrium (III) chloride hexahydrate	Lanthanide	YT3	33
Xylitol	Monosaccharide	XYL	25
D-(-)-fructose	Monosaccharide	FRU; FUD	36; 4
D-sorbitol	Monosaccharide	SOR	12
Myo-inositol	Monosaccharide	INS	16
L-rhamnose monohydrate	Monosaccharide	RAM	43
DL-threonine	Amino-acid	DTH; THR	23;n/a
DL-histidine, HCl, H2O	Amino-acid	DHI; HIS	24; n/a
DL-5-hydroxylysine, HCl	Amino-acid	n/a; LYZ	0; 7
Trans-4-hydroxy-L-proline	Amino-acid	HYP	149
Spermine, 4HCl	Polyamine	SPM	103
Spermidine, 3HCl	Polyamine	SPD	32
1,4-diaminobutane, 2HCl	Polyamine	PUT	22
DL-ornithine, HCl	Polyamine	ORD; ORN	3; 56
NDSB 256	Surfactant	DMX	4
NDSB 195	Surfactant	NDS	7
Bis-tris	Buffer	BTB	114

*No of structures as determined by a query of the pdb carried out in December 2014

Continued on next page

Table 2: Mixes of additives used in Morpheus II

Mix name	Composition	Catalogue Number (100 ml)	Catalogue Number (250 ml)
0.9 M LiNaK	0.3 M Lithium sulfate, 0.3 M Sodium sulfate, 0.3 M Potassium sulfate	MD2-100-231	MD2-250-231
0.02M Divalents II	0.005M Manganese(II) chloride tetrahydrate, 0.005M Cobalt(II) chloride hexahydrate , 0.005M Nickel(II) chloride hexahydrate, 0.005M Zinc acetate dihydrate	MD2-100-232	MD2-250-232
0.04 M Alkalis	0.01M Rubidium chloride, 0.01M Strontium acetate, 0.01M Cesium acetate, 0.01M Barium acetate	MD2-100-233	MD2-250-233
0.02 M Oxometalates	0.005M Sodium chromate tetrahydrate, 0.005M Sodium molybdate dihydrate, 0.005M Sodium tungstate dihydrate, 0.005M Sodium orthovanadate	MD2-100-234	MD2-250-234
0.02M Lanthanides	0.005M Yttrium(III) chloride hexahydrate, 0.005M Erbium(III) chloride hexahydrate, 0.005M Terbium(III) chloride hexahydrate, 0.005M Ytterbium(III) chloride hexahydrate	MD2-100-235	MD2-250-235
1M Monosaccharides II	0.2M Xylitol, 0.2M *Myo*-Inositol, 0.2M D-(-)-Fructose, 0.2M L-Rhamnose monohydrate, 0.2M D-Sorbitol	MD2-100-236	MD2-250-236
1M Amino acids II	0.2M DL-Arginine hydrochloride, 0.2M DL-Threonine, 0.2M DL-Histidine monohydrochloride monohydrate, 0.2M DL-5-Hydroxylysine hydrochloride, 0.2M *trans*-4-hydroxy-L-proline	MD2-100-237	MD2-250-237
0.4 M Polyamines (provided as powder for 10mL kits)*	0.1M Spermine tetrahydrochloride, 0.1M Spermidine trihydrochloride, 0.1M 1,4-Diaminobutane dihydrochloride, 0.1M DL-Ornithine monohydrochloride	MD2-100-238	MD2-250-238

Table 3: Buffer systems used in Morpheus II

Mix name*	Conc.	pH @ 20°C	Composition	Catalogue Number (100 ml)	Catalogue Number (250 ml)
Buffer System 4	1.0M	6.5	MOPSO, Bis-Tris	MD2-100-243	MD2-250-243
Buffer System 5	1.0M	7.5	BES, Triethanolamine (TEA)	MD2-100-244	MD2-250-244
Buffer System 6	1.0M	8.5	Gly-Gly, AMPD	MD2-100-245	MD2-250-245

*Buffer systems 1, 2 & 3 are allocated to the original Morpheus screen.

Continued on next page

Table 4: Mixes of Precipitants used in Morpheus II

Mix name*	Composition	Catalogue Number (100 ml)	Catalogue Number (250 ml)
72% Precipitant Mix 5	30% w/v PEG 3000, 40% v/v 1, 2, 4-Butanetriol, 2% w/v NDSB 256	MD2-100-239	MD2-250-239
65% Precipitant Mix 6	25% w/v PEG 4000, 40% w/v 1,2,6-Hexanetriol	MD2-100-240	MD2-250-240

60%Precipitant Mix 7	20% w/v PEG 8000, 40% v/v 1,5-Pentanediol	MD2-100-241	MD2-250-241
62% Precipitant Mix 8	10% w/v PEG 20000, 50% w/v Trimethylpropane, 2% w/v NDSB 195	MD2-100-242	MD2-250-242

*precipitant Mixes 1, 2, 3 & 4 are allocated to the original Morpheus screen.

Continued on next page

Screen ID	Well #	Conc.		Additives (PDB ligands)	Conc.		Buffer	pH	Conc.		Precipitant
1-1	A1	90	mM	LiNaK	0.1	M	Buffer System 4	6.5	36	% v/v	Precipitant Mix 5
1-2	A2	90	mM	LiNaK	0.1	M	Buffer System 4	6.5	32.5	% v/v	Precipitant Mix 6
1-3	A3	90	mM	LiNaK	0.1	M	Buffer System 4	6.5	30	% v/v	Precipitant Mix 7
1-4	A4	90	mM	LiNaK	0.1	M	Buffer System 4	6.5	31	% v/v	Precipitant Mix 8
1-5	A5	90	mM	LiNaK	0.1	M	Buffer System 5	7.5	36	% v/v	Precipitant Mix 5
1-6	A6	90	mM	LiNaK	0.1	M	Buffer System 5	7.5	32.5	% v/v	Precipitant Mix 6
1-7	A7	90	mM	LiNaK	0.1	M	Buffer System 5	7.5	30	% v/v	Precipitant Mix 7
1-8	A8	90	mM	LiNaK	0.1	M	Buffer System 5	7.5	31	% v/v	Precipitant Mix 8
1-9	A9	90	mM	LiNaK	0.1	M	Buffer System 6	8.5	36	% v/v	Precipitant Mix 5
1-10	A10	90	mM	LiNaK	0.1	M	Buffer System 6	8.5	32.5	% v/v	Precipitant Mix 6
1-11	A11	90	mM	LiNaK	0.1	M	Buffer System 6	8.5	30	% v/v	Precipitant Mix 7
1-12	A12	90	mM	LiNaK	0.1	M	Buffer System 6	8.5	31	% v/v	Precipitant Mix 8
1-13	B1	2	mM	Divalents II	0.1	M	Buffer System 4	6.5	36	% v/v	Precipitant Mix 5
1-14	B2	2	mM	Divalents II	0.1	M	Buffer System 4	6.5	32.5	% v/v	Precipitant Mix 6
1-15	B3	2	mM	Divalents II	0.1	M	Buffer System 4	6.5	30	% v/v	Precipitant Mix 7
1-16	B4	2	mM	Divalents II	0.1	M	Buffer System 4	6.5	31	% v/v	Precipitant Mix 8
1-17	B5	2	mM	Divalents II			none		36	% v/v	Precipitant Mix 5
1-18	B6	2	mM	Divalents II			none		32.5	% v/v	Precipitant Mix 6
1-19	B7	2	mM	Divalents II			none		30	% v/v	Precipitant Mix 7
1-20	B8	2	mM	Divalents II			none		31	% v/v	Precipitant Mix 8
1-21	B9	2	mM	Divalents II	0.1	M	Buffer System 6	8.5	36	% v/v	Precipitant Mix 5
1-22	B10	2	mM	Divalents II	0.1	M	Buffer System 6	8.5	32.5	% v/v	Precipitant Mix 6
1-23	B11	2	mM	Divalents II	0.1	M	Buffer System 6	8.5	30	% v/v	Precipitant Mix 7
1-24	B12	2	mM	Divalents II	0.1	M	Buffer System 6	8.5	31	% v/v	Precipitant Mix 8
1-25	C1	4	mM	Alkalis	0.1	M	Buffer System 4	6.5	36	% v/v	Precipitant Mix 5
1-26	C2	4	mM	Alkalis	0.1	M	Buffer System 4	6.5	32.5	% v/v	Precipitant Mix 6
1-27	C3	4	mM	Alkalis	0.1	M	Buffer System 4	6.5	30	% v/v	Precipitant Mix 7
1-28	C4	4	mM	Alkalis	0.1	M	Buffer System 4	6.5	31	% v/v	Precipitant Mix 8
1-29	C5	4	mM	Alkalis	0.1	M	Buffer System 5	7.5	36	% v/v	Precipitant Mix 5
1-30	C6	4	mM	Alkalis	0.1	M	Buffer System 5	7.5	32.5	% v/v	Precipitant Mix 6
1-31	C7	4	mM	Alkalis	0.1	M	Buffer System 5	7.5	30	% v/v	Precipitant Mix 7
1-32	C8	4	mM	Alkalis	0.1	M	Buffer System 5	7.5	31	% v/v	Precipitant Mix 8
1-33	C9	4	mM	Alkalis	0.1	M	Buffer System 6	8.5	36	% v/v	Precipitant Mix 5
1-34	C10	4	mM	Alkalis	0.1	M	Buffer System 6	8.5	32.5	% v/v	Precipitant Mix 6
1-35	C11	4	mM	Alkalis	0.1	M	Buffer System 6	8.5	30	% v/v	Precipitant Mix 7
1-36	C12	4	mM	Alkalis	0.1	M	Buffer System 6	8.5	31	% v/v	Precipitant Mix 8
1-37	D1	2	mM	Oxometalates	0.1	M	Buffer System 4	6.5	36	% v/v	Precipitant Mix 5
1-38	D2	2	mM	Oxometalates	0.1	M	Buffer System 4	6.5	32.5	% v/v	Precipitant Mix 6
1-39	D3	2	mM	Oxometalates	0.1	M	Buffer System 4	6.5	30	% v/v	Precipitant Mix 7
1-40	D4	2	mM	Oxometalates	0.1	M	Buffer System 4	6.5	31	% v/v	Precipitant Mix 8
1-41	D5	2	mM	Oxometalates	0.1	M	Buffer System 5	7.5	36	% v/v	Precipitant Mix 5
1-42	D6	2	mM	Oxometalates	0.1	M	Buffer System 5	7.5	32.5	% v/v	Precipitant Mix 6
1-43	D7	2	mM	Oxometalates	0.1	M	Buffer System 5	7.5	30	% v/v	Precipitant Mix 7
1-44	D8	2	mM	Oxometalates	0.1	M	Buffer System 5	7.5	31	% v/v	Precipitant Mix 8
1-45	D9	2	mM	Oxometalates	0.1	M	Buffer System 6	8.5	36	% v/v	Precipitant Mix 5
1-46	D10	2	mM	Oxometalates	0.1	M	Buffer System 6	8.5	32.5	% v/v	Precipitant Mix 6
1-47	D11	2	mM	Oxometalates	0.1	M	Buffer System 6	8.5	30	% v/v	Precipitant Mix 7
1-48	D12	2	mM	Oxometalates	0.1	M	Buffer System 6	8.5	31	% v/v	Precipitant Mix 8

Screen should be stored between 10-18°C and gently mixed before use

Continued on next page

Screen ID	Well #	Conc.		Additives (PDB ligands)	Conc.		Buffer	pH	Conc.		Precipitant
2-1	E1	2	mM	Lanthanides*	0.1	M	Buffer System 4	6.5	36	% v/v	Precipitant Mix 5
2-2	E2	2	mM	Lanthanides*	0.1	M	Buffer System 4	6.5	32.5	% v/v	Precipitant Mix 6
2-3	E3	2	mM	Lanthanides*	0.1	M	Buffer System 4	6.5	30	% v/v	Precipitant Mix 7
2-4	E4	2	mM	Lanthanides*	0.1	M	Buffer System 4	6.5	31	% v/v	Precipitant Mix 8
2-5	E5	2	mM	Lanthanides*	0.1	M	Buffer System 5	7.5	36	% v/v	Precipitant Mix 5
2-6	E6	2	mM	Lanthanides*	0.1	M	Buffer System 5	7.5	32.5	% v/v	Precipitant Mix 6
2-7	E7	2	mM	Lanthanides*	0.1	M	Buffer System 5	7.5	30	% v/v	Precipitant Mix 7
2-8	E8	2	mM	Lanthanides*	0.1	M	Buffer System 5	7.5	31	% v/v	Precipitant Mix 8
2-9	E9	2	mM	Lanthanides*	0.1	M	Buffer System 6	8.5	36	% v/v	Precipitant Mix 5
2-10	E10	2	mM	Lanthanides*	0.1	M	Buffer System 6	8.5	32.5	% v/v	Precipitant Mix 6
2-11	E11	2	mM	Lanthanides*	0.1	M	Buffer System 6	8.5	30	% v/v	Precipitant Mix 7
2-12	E12	2	mM	Lanthanides*	0.1	M	Buffer System 6	8.5	31	% v/v	Precipitant Mix 8
2-13	F1	100	mM	Monosaccharides II	0.1	M	Buffer System 4	6.5	36	% v/v	Precipitant Mix 5
2-14	F2	100	mM	Monosaccharides II	0.1	M	Buffer System 4	6.5	32.5	% v/v	Precipitant Mix 6
2-15	F3	100	mM	Monosaccharides II	0.1	M	Buffer System 4	6.5	30	% v/v	Precipitant Mix 7
2-16	F4	100	mM	Monosaccharides II	0.1	M	Buffer System 4	6.5	31	% v/v	Precipitant Mix 8
2-17	F5	100	mM	Monosaccharides II	0.1	M	Buffer System 5	7.5	36	% v/v	Precipitant Mix 5
2-18	F6	100	mM	Monosaccharides II	0.1	M	Buffer System 5	7.5	32.5	% v/v	Precipitant Mix 6
2-19	F7	100	mM	Monosaccharides II	0.1	M	Buffer System 5	7.5	30	% v/v	Precipitant Mix 7
2-20	F8	100	mM	Monosaccharides II	0.1	M	Buffer System 5	7.5	31	% v/v	Precipitant Mix 8
2-21	F9	100	mM	Monosaccharides II	0.1	M	Buffer System 6	8.5	36	% v/v	Precipitant Mix 5
2-22	F10	100	mM	Monosaccharides II	0.1	M	Buffer System 6	8.5	32.5	% v/v	Precipitant Mix 6
2-23	F11	100	mM	Monosaccharides II	0.1	M	Buffer System 6	8.5	30	% v/v	Precipitant Mix 7
2-24	F12	100	mM	Monosaccharides II	0.1	M	Buffer System 6	8.5	31	% v/v	Precipitant Mix 8
2-25	G1	100	mM	Amino acids II	0.1	M	Buffer System 4	6.5	36	% v/v	Precipitant Mix 5
2-26	G2	100	mM	Amino acids II	0.1	M	Buffer System 4	6.5	32.5	% v/v	Precipitant Mix 6
2-27	G3	100	mM	Amino acids II	0.1	M	Buffer System 4	6.5	30	% v/v	Precipitant Mix 7
2-28	G4	100	mM	Amino acids II	0.1	M	Buffer System 4	6.5	31	% v/v	Precipitant Mix 8
2-29	G5	100	mM	Amino acids II	0.1	M	Buffer System 5	7.5	36	% v/v	Precipitant Mix 5
2-30	G6	100	mM	Amino acids II	0.1	M	Buffer System 5	7.5	32.5	% v/v	Precipitant Mix 6
2-31	G7	100	mM	Amino acids II	0.1	M	Buffer System 5	7.5	30	% v/v	Precipitant Mix 7
2-32	G8	100	mM	Amino acids II	0.1	M	Buffer System 5	7.5	31	% v/v	Precipitant Mix 8
2-33	G9	100	mM	Amino acids II	0.1	M	Buffer System 6	8.5	36	% v/v	Precipitant Mix 5
2-34	G10	100	mM	Amino acids II	0.1	M	Buffer System 6	8.5	32.5	% v/v	Precipitant Mix 6
2-35	G11	100	mM	Amino acids II	0.1	M	Buffer System 6	8.5	30	% v/v	Precipitant Mix 7
2-36	G12	100	mM	Amino acids II	0.1	M	Buffer System 6	8.5	31	% v/v	Precipitant Mix 8
2-37	H1	40	mM	Polyamines†	0.1	M	Buffer System 4	6.5	36	% v/v	Precipitant Mix 5
2-38	H2	40	mM	Polyamines†	0.1	M	Buffer System 4	6.5	32.5	% v/v	Precipitant Mix 6
2-39	H3	40	mM	Polyamines†	0.1	M	Buffer System 4	6.5	30	% v/v	Precipitant Mix 7
2-40	H4	40	mM	Polyamines†	0.1	M	Buffer System 4	6.5	31	% v/v	Precipitant Mix 8
2-41	H5	40	mM	Polyamines†	0.1	M	Buffer System 5	7.5	36	% v/v	Precipitant Mix 5
2-42	H6	40	mM	Polyamines†	0.1	M	Buffer System 5	7.5	32.5	% v/v	Precipitant Mix 6
2-43	H7	40	mM	Polyamines†	0.1	M	Buffer System 5	7.5	30	% v/v	Precipitant Mix 7
2-44	H8	40	mM	Polyamines†	0.1	M	Buffer System 5	7.5	31	% v/v	Precipitant Mix 8
2-45	H9	40	mM	Polyamines†	0.1	M	Buffer System 6	8.5	36	% v/v	Precipitant Mix 5
2-46	H10	40	mM	Polyamines†	0.1	M	Buffer System 6	8.5	32.5	% v/v	Precipitant Mix 6
2-47	H11	40	mM	Polyamines†	0.1	M	Buffer System 6	8.5	30	% v/v	Precipitant Mix 7
2-48	H12	40	mM	Polyamines†	0.1	M	Buffer System 6	8.5	31	% v/v	Precipitant Mix 8

Screen should be stored between 10-18°C and gently mixed before use

Appendix 2: Sparse Matrix Screen Conditions LMB

Tube #	Conc	Units	Precipitant	Conc	Units	Buffer	pH	Conc	Units	Salt
1-1	25	% v/v	1,4-Butanediol	0.1	M	Tris	8.0			
1-2	30	% v/v	2-Propanol	0.1	M	Sodium cacodylate	4.6	0.1	M	Sodium citrate tribasic dihydrate
1-3	9	% v/v	2-Propanol	0.1	M	MES	6.2	0.2	M	Calcium acetate hydrate
1-4	15	% v/v	2-Propanol	0.2	M	Imidazole	7.6			
1-5	4	M	Ammonium acetate	0.1	M	Bis-Tris propane	7.0			
1-6	1.4	M	Ammonium phosphate monobasic	0.1	M	Tris	7.5			
1-7	1.85	M	Ammonium sulfate	0.1	M	Sodium acetate	4.5	0.4	M	Potassium thiocyanate
1-8	1.2	M	Ammonium sulfate	0.1	M	Ammonium acetate	5.5			
1-9	1	M	Ammonium sulfate	0.1	M	Sodium citrate	5.5			
1-10	2.5	M	Ammonium sulfate	0.1	M	MES	5.6	0.2	M	Sodium chloride
1-11	1.2	M	Ammonium sulfate	0.1	M	MES	5.9			
1-12	2	M	Ammonium sulfate	0.05	M	MES	6.0	5	mM	Magnesium acetate tetrahydrate
1-13	1.6	M	Ammonium sulfate	0.1	M	MES	6.5			
1-14	1.9	M	Ammonium sulfate	0.1	M	MES	6.5			
1-15	1.95	M	Ammonium sulfate	0.1	M	Sodium citrate	6.5			
1-16	1.7	M	Ammonium sulfate	0.1	M	Tris	8.0	0.2	M	Sodium chloride
1-17	1	M	Ammonium sulfate/	0.1	M	Tris	8.5			
	12	% v/v	Glycerol							
1-18	1	M	Ammonium sulfate/	0.1	M	CHES	9.5	0.2	M	Sodium chloride
	6	% v/v	Glycerol							
1-19	38	% v/v	1,4-Dioxane							
1-20	14	% v/v	Ethanol	0.1	M	ADA	6.0			
1-21	18	% v/v	Ethanol	0.1	M	Bis-Tris propane	7.1			
1-22	24	% v/v	Ethanol	0.1	M	HEPES	7.8	0.04	M	Magnesium chloride hexahydrate
1-23	17.5	% v/v	Ethanol							
1-24	1.3	M	Lithium chloride	0.1	M	HEPES	7.4			
1-25	1.5	M	Lithium chloride/	0.1	M	BICINE	8.4			
	10	% w/v	PEG 6000							
1-26	1.6	M	Lithium sulfate	0.4	M	Tris	7.5			
1-27	1.6	M	Magnesium sulfate heptahydrate/	0.1	M	HEPES	7.2			
	6	% v/v	Glycerol							
1-28	30	% v/v	MPD	0.1	M	Imidazole	7.0			
1-29	70	% v/v	MPD	0.1	M	HEPES	7.5			
1-30	18	% w/v	PEG 1000	0.1	M	Sodium/potassium phosphate	6.2	0.2	M	Sodium chloride
1-31	9.5	% w/v	PEG 2000	0.1	M	Sodium acetate	4.6	1.5	mM	Magnesium chloride hexahydrate
1-32	8	% w/v	PEG 2000	0.1	M	Imidazole	7.7	0.3	M	Calcium acetate hydrate
1-33	26	% w/v	PEG 2000 MME	0.1	M	Bis-Tris	5.8			
1-34	15	% w/v	PEG 2000 MME	0.1	M	Bis-Tris propane	6.9			
1-35	8	% w/v	PEG 20,000/	0.1	M	Sodium acetate	4.5	0.25	M	Potassium bromide
	8	% v/v	PEG 550 MME							
1-36	8	% w/v	PEG 20,000/	0.1	M	Sodium acetate	5.5	0.2	M	Potassium thiocyanate
	8	% v/v	PEG 550 MME							
1-37	7.5	% w/v	PEG 20,000/	0.1	M	Tris	7.5	0.08	M	Sodium formate
	7.5	% v/v	PEG 550 MME							
1-38	8	% w/v	PEG 20,000/	0.05	M	Tris	8.5	1.2	M	Sodium formate
	8	% v/v	PEG 550 MME							
1-39	28	% v/v	PEG 300	0.1	M	Imidazole	7.0	0.07	M	Calcium acetate hydrate
1-40	30	% v/v	PEG 300	0.1	M	Bis-Tris propane	7.0			
	15	% w/v	PEG 1000							
1-41	26	% w/v	PEG 3000	0.1	M	CHES	9.2			
1-42	21	% w/v	PEG 3350	0.1	M	MES	6.0	0.15	M	Sodium chloride
1-43	15	% w/v	PEG 3350	0.1	M	MES	6.2			
1-44	24	% w/v	PEG 3350	0.05	M	HEPES	6.8	0.15	M	Sodium chloride
	2	% v/v	PEG 400							
1-45	27	% w/v	PEG 3350	0.1	M	Bis-Tris propane	7.0	0.2	M	Lithium sulfate
1-46	12	% w/v	PEG 3350/	0.1	M	HEPES	7.5	5	mM	Cobalt(II) chloride hexahydrate/
								5	mM	Magnesium chloride hexahydrate
								5	mM	Nickel(II) chloride hexahydrate
1-47	20	% w/v	PEG 3350	0.2	M	Tris	8.5			
1-48	18	% w/v	PEG 3350/	0.1	M	CAPSO	9.0	17	% v/v	PEG 400
	4.8	% v/v	2-Propanol							

Continued on next page

Tube #	Conc	Units	Precipitant	Conc	Units	Buffer	pH	Conc	Units	Salt
2-1	6	% w/v	PEG 3350					0.15	M	Sodium chloride/
								0.4	M	Potassium iodide
2-2	18	% w/v	PEG 3350					0.2	M	Sodium acetate trihydrate
2-3	22	% w/v	PEG 3350					0.2	M	Zinc acetate dihydrate
2-4	26	% v/v	PEG 400	0.1	M	Sodium citrate	5.5	0.1	M	Magnesium chloride hexahydrate/
								0.1	M	Sodium chloride
2-5	4	% v/v	PEG 400	0.1	M	Sodium citrate	5.6	0.1	M	Lithium sulfate
2-6	10	% v/v	PEG 400	0.05	M	MES	6.0	0.1	M	Potassium chloride/
								2	mM	Magnesium chloride hexahydrate
2-7	28	% v/v	PEG 400	0.1	M	MOPS	6.5	0.2	M	Sodium chloride
2-8	25	% v/v	PEG 400/	0.07	M	MES	6.6	1.5	mM	Magnesium chloride hexahydrate
	4.5	% v/v	Ethanol							
2-9	40	% v/v	PEG 400	0.1	M	Tris	8.4	0.2	M	Lithium sulfate
2-10	3.5	% w/v	PEG 4000/	0.1	M	Sodium acetate	4.6			
	15	% v/v	Glycerol							
2-11	28	% w/v	PEG 4000	0.1	M	Sodium citrate	5.2	0.2	M	Ammonium acetate
2-12	20	% w/v	PEG 4000/	0.1	M	Sodium citrate	5.5			
	20	% v/v	2-Propanol							
2-13	25	% w/v	PEG 4000	0.1	M	Sodium acetate	5.6	0.2	M	Ammonium sulfate
2-14	16	% w/v	PEG 4000/	0.1	M	Sodium citrate	5.8	0.1	M	Ammonium sulfate
	20	% v/v	Glycerol							
2-15	18	% w/v	PEG 4000	0.1	M	Sodium citrate	6.0	0.2	M	Ammonium acetate
2-16	14	% w/v	PEG 4000	0.1	M	Sodium/potassium phosphate	6.2			
	6	% v/v	MPD							
2-17	29	% w/v	PEG 4000	0.1	M	Sodium citrate	6.5	0.1	M	Magnesium acetate tetrahydrate/
								0.1	M	Ammonium sulfate
2-18	11	% w/v	PEG 4000	0.1	M	Tris	7.8	0.1	M	Magnesium chloride hexahydrate
2-19	20	% w/v	PEG 4000	0.1	M	Tris	8.5	0.2	M	Magnesium chloride hexahydrate
2-20	18	% w/v	PEG 4000	0.1	M	Tris	9.0	0.3	M	Sodium acetate trihydrate
2-21	18	% w/v	PEG 5000 MME	0.1	M	MES	6.5	0.2	M	Ammonium sulfate
2-22	20	% v/v	PEG 600	0.1	M	Sodium cacodylate	5.6	0.15	M	Potassium thiocyanate/
								0.2	M	Sodium chloride
2-23	29	% w/v	PEG 6000	0.2	M	Sodium citrate	3.5			
2-24	20	% w/v	PEG 6000	0.1	M	Sodium citrate	4.0	0.2	M	Lithium chloride
2-25	12	% w/v	PEG 6000	0.1	M	Sodium citrate	5.6	0.1	M	Lithium sulfate
2-26	3.5	% w/v	PEG 6000	0.1	M	Bis-Tris propane	7.1	0.1	M	Potassium chloride
2-27	8	% w/v	PEG 8000	0.1	M	Sodium acetate	4.5	0.05	M	Magnesium chloride hexahydrate
2-28	22	% w/v	PEG 8000	0.1	M	Sodium cacodylate	5.0	0.2	M	Sodium acetate trihydrate
2-29	8	% w/v	PEG 8000	0.08	M	Potassium phosphate	5.6			
2-30	11	% w/v	PEG 8000	0.1	M	MES	6.0	0.24	M	Ammonium sulfate
2-31	11	% w/v	PEG 8000	0.1	M	Sodium cacodylate	6.2	0.4	M	Sodium chloride
2-32	13	% w/v	PEG 8000	0.05	M	Sodium cacodylate	6.5	0.09	M	Ammonium sulfate
2-33	20	% w/v	PEG 8000	0.1	M	MES	6.5	0.2	M	Magnesium acetate tetrahydrate
2-34	10	% w/v	PEG 8000/	0.1	M	HEPES	7.5			
	9	% v/v	Ethylene glycol							
2-35	20	% w/v	PEG 8000	0.1	M	CAPS	9.0	0.2	M	Magnesium chloride hexahydrate
2-36	20	% w/v	PEG 8000	0.1	M	CHES	9.5			
2-37	1	M	Potassium phosphate monobasic/	0.1	M	Sodium cacodylate	6.5			
	3	% v/v	2-Propanol							
2-38	1.4	M	Sodium acetate trihydrate	0.1	M	Sodium cacodylate	6.5			
2-39	0.5	M	Sodium bromide	0.1	M	Tris	7.5			
2-40	2	M	Sodium chloride	0.1	M	Sodium citrate	4.0			
2-41	3	M	Sodium chloride	0.1	M	Tris	7.5			
2-42	1.5	M	Sodium citrate tribasic dihydrate	0.1	M	Sodium citrate	6.5			
2-43	0.85	M	Sodium citrate tribasic dihydrate	0.1	M	Tris	8.0	0.1	M	Sodium chloride
2-44	1.36	M	Sodium formate/	0.1	M	Tris	8.5	19	% v/v	Glycerol
	3	% w/v	PEG 35,000							
2-45	0.45	M	Potassium sodium tartrate tetrahydrate	0.1	M	MES	6.3			
2-46	0.9	M	Potassium sodium tartrate tetrahydrate/	0.05	M	HEPES	7.4			
	20	% v/v	Glycerol							
2-47				1.6	M	Sodium/potassium phosphate	6.0			
2-48				1	M	Sodium citrate	6.2			

5. References

Abdul Rehman, S. A., Kristariyanto, Y. A., Choi, S. Y., Nkosi, P. J., Weidlich, S., Labib, K., . . . & Kulathu, Y. (2016). MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes. *Molecular cell*, 63(1), 146–155.

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., . . . Terwilliger. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 2), 213–221.

Akimana, C., Al-Khodor, S., & & Abu Kwaik, Y. (2010). Host factors required for modulation of phagosome biogenesis and proliferation of Francisella tularensis within the cytosol. *PloS one*, 5(6), e11025.

Al-Hakim, A. K., Zagorska, A., Chapman, L., Deak, M., Peggie, M., & & Alessi, D. R. (2008). Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains. *The Biochemical journal*, 411(2), 249–260.

Arnold, E., & Rossmann, M. G. (1986). Effect of errors, redundancy, and solvent content in the molecular replacement procedure for the structure determination of biological macromolecules. *Proceedings of the National Academy of Sciences of the United States of America*, 83(15), 5489–5493.

Azadi, S., Mahboubi, A., Naghdi, N., Solaimanian, R., & & Mortazavi, S. A. (2017). Evaluation of Sorbitol-Methanol Co-Feeding Strategy on Production of Recombinant Human Growth Hormone in Pichia Pastoris. *Iranian journal of pharmaceutical research : IJPR*, 16(4), 1555–1564.

Bacik, J., Avvakumov, G., Walker, J., Xue, S., Arrowsmith, C., Edwards, A., . . . (SGC), S. G. (2009, 10 20). *Structure 3JYU*. Retrieved from RCSB PDB: https://www.rcsb.org/structure/3JYU

Balaji, V., Pokrzywa, W., & Hoppe, T. (2018). Ubiquitylation Pathways In Insulin Signaling and Organismal Homeostasis. *BioEssays : news and reviews in molecular, cellular and developmental biology,*, 40(5), e1700223.

Bañuelos, S., Saraste, M., & Djinović Carugo, K. (1998). Structural comparisons of calponin homology domains: implications for actin binding. *Structure (London, England : 1993)*, 6(11), 1419–1431.

Baranauskiene, L., Kuo, T. C., Chen, W. Y., & Matulis, D. (2019). Isothermal titration calorimetry for characterization of recombinant proteins. *Current opinion in biotechnology*, 55, 9–15.

Barbarin-Bocahu, I., & Graille, M. (2022). The X-ray crystallography phase problem solved thanks to AlphaFold and RoseTTAFold models: a case-study report. *Acta crystallographica. Section D, Structural biology*, 78(Pt 4), 517–531.

Bartoschik, T., Galinec, S., Kleusch, C., Walkiewicz, K., Breitsprecher, D., Weigert, S., . . . Tschammer, N. (2018). Near-native, site-specific and purification-free protein labeling for quantitative protein interaction analysis by MicroScale Thermophoresis. *Scientific reports*, 8(1), 4977.

Bello, A. I., Goswami, R., Brown, S. L., Costanzo, K., Shores, T., Allan, S., . . . Mohan, R. D. (2022). Deubiquitinases in Neurodegeneration. *Cells*, 11(3), 556.

Bernassola, F., Karin, M., Ciechanover, A., & Melino, G. (2008). The HECT family of E3 ubiquitin ligases: multiple players in cancer development. *Cancer cell*, 14(1), 10–21.

Bonacci, T., & Emanuele, M. J. (2020). Dissenting degradation: Deubiquitinases in cell cycle and cancer. *Seminars in cancer biology*, 67(Pt 2), 145–158.

Bragg, W. L. (1913). The diffraction of short electromagnetic waves by a crystal. *Proceedings of the Cambridge Philosophical Society*, 17: 43-57.

Bui, Q. T., Hong, J. H., Kwak, M., Lee, J. Y., & Lee, P. C. (2021). Ubiquitin-Conjugating Enzymes in Cancer. *Cells*, 10(6), 1383.

Bujalowski, W., & Lohman, T. M. (1989). Negative co-operativity in Escherichia coli single strand binding protein-oligonucleotide interactions. II. Salt, temperature and oligonucleotide length effects. *Journal of molecular biology*, 207(1), 269–288.

Burroughs, A. M., Balaji, S., Iyer, L. M., & Aravind, L. (2007). Small but versatile: the extraordinary functional and structural diversity of the beta-grasp fold. *Biology direct*, 2, 18.

Burrows, J. F., McGrattan, M. J., & Johnston, J. A. (2005). The DUB/USP17 deubiquitinating enzymes, a multigene family within a tandemly repeated sequence. *Genomics*, 85(4), 524–529.

Burrows, J., McGrattan, M. J., Rascle, A., Humbert, M., Baek, K. H., & Johnston, J. A. (2004). DUB-3, a cytokine-inducible deubiquitinating enzyme that blocks proliferation. *The Journal of biological chemistry*, 279(14), 13993–14000.

Burrows, J., Scott, C. J., & Johnston, J. A. (2010). The DUB/USP17 deubiquitinating enzymes: a gene family within a tandemly repeated sequence, is also embedded within the copy number variable beta-defensin cluster. *BMC genomics*, 11, 250.

Buurma, N. J., & Haq, I. (2007). Advances in the analysis of isothermal titration calorimetry data for ligand-DNA interactions. *Methods (San Diego, Calif.)*, 42(2), 162–172.

Cao, S., Engilberge, S., Girard, E., Gabel, F., Franzetti, B., & Maupin-Furlow, J. A. (2017). Structural Insight into Ubiquitin-Like Protein Recognition and Oligomeric States of JAMM/MPN+ Proteases. *Structure*, 25(6), 823–833.e6.

Cappadocia, L., & Lima, C. D. (2018). Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chemical reviews*, 118(3), 889–918.

Carrà, G., Panuzzo, C., Torti, D., Parvis, G., Crivellaro, S., Familiari, U., . . . Morotti, A. (2017). Therapeutic inhibition of USP7-PTEN network in chronic lymphocytic leukemia: a strategy to overcome TP53 mutated/deleted clones. *Oncotarget*, 8(22), 35508–35522.

Castañeda, C. A., Dixon, E. K., Walker, O., Chaturvedi, A., Nakasone, M. A., Curtis, J. E., . . . Fushman, D. (1993). Linkage via K27 Bestows Ubiquitin Chains with Unique Properties among Polyubiquitins. *Structure*, 24(3), 423–436.

Caulton, S. (2016). Characterisation of the deubiquitinating enzyme USP20. PhD Thesis, University of Nottingham.

Chauhan, D., Tian, Z., Nicholson, B., Kumar, K. G., Zhou, B., Carrasco, R., . . . Orlowski, R. (2012). A small molecule inhibitor of ubiquitin-specific protease-7 induces apoptosis in multiple myeloma cells and overcomes bortezomib resistance. *Cancer cell*, 22(3), 345–358.

Chen, R., Zhang, L., Zhong, B., Tan, B., Liu, Y., & Shu, H. B. (2010). The ubiquitin-specific protease 17 is involved in virus-triggered type I IFN signaling. *Cell research*, 20(7), 802–811.

Cheng, C., Niu, C., Yang, Y., Wang, Y., & Lu, M. (2013). Expression of HAUSP in gliomas correlates with disease progression and survival of patients. *Oncology reports*, 29(5), 1730–1736.

Chhetri, G., Kalita, P., & Tripathi, T. (2015). An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli. *MethodsX*, 2, 385–391.

Chou, C. K., Chang, Y. T., Korinek, M., Chen, Y. T., Yang, Y. T., Leu, S., . . . Chiu, C. C. (2017). The Regulations of Deubiquitinase USP15 and Its Pathophysiological Mechanisms in Diseases. *International journal of molecular sciences*, 18(3), 483.

Ciechanover, A., Elias, S., Heller, H., & Hershko, A. (1982). "Covalent affinity" purification of ubiquitin-activating enzyme. *The Journal of biological chemistry*, 257(5), 2537–2542.

Clague, M. J., Barsukov, I., Coulson, J. M., Liu, H., Rigden, D. J., & Urbé, S. (2013). Deubiquitylases from genes to organism. *Physiological reviews*, 93(3), 1289–1315.

Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography*, 50(Pt 5), 760–763.

Colland, F., Formstecher, E., Jacq, X., Reverdy, C., Planquette, C., Conrath, S., . . . Daviet. (2009). Small-molecule inhibitor of USP7/HAUSP ubiquitin protease stabilizes and activates p53 in cells. *Molecular cancer therapeutics*, 8(8), 2286–2295.

Cornelissen, T., Haddad, D., Wauters, F., Van Humbeeck, C., Mandemakers, W., Koentjoro, B., . . . Vandenberghe, W. (2014). The deubiquitinase USP15 antagonizes Parkin-mediated mitochondrial ubiquitination and mitophagy. *Human molecular genetics*, 23(19), 5227–5242.

Coux, O., Tanaka, K., & Goldberg, A. L. (1996). Structure and functions of the 20S and 26S proteasomes. *Annual review of biochemistry*, 65, 801–847.

Dauter, Z. (2017). Collection of X-Ray Diffraction Data from Macromolecular Crystals. *Methods in molecular biology (Clifton, N.J.)*, 1607, 165–184.

Ducker, C., & Shaw, P. E. (2021). USP17-mediated de-ubiquitination and cancer: Clients cluster around the cell cycle. *The international journal of biochemistry & cell biology*, 130, 105886.

Duff, M. R., Grubbs, J., & Howell, E. E. (2011). Isothermal titration calorimetry for measuring macromolecule-ligand affinity. *Journal of visualized experiments : JoVE*, (55), 2796.

Duhr, S., & Braun, D. (2006). Why molecules move along a temperature gradient. *Proceedings of the National Academy of Sciences of the United States of America*, 103(52), 19678–19682.

Dupont, S., Mamidi, A., Cordenonsi, M., Montagner, M., Zacchigna, L., Adorno, M., . . . Piccolo, S. (2009). FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls Smad4 monoubiquitination. *Cell*, 136(1), 123–135.

Eichhorn, P. J., Rodón, L., Gonzàlez-Juncà, A., Dirac, A., Gili, M., Martínez-Sáez, E., . . . Seoane, J. (2012). USP15 stabilizes TGF-β receptor I and promotes oncogenesis through the activation of TGF-β signaling in glioblastoma. *Nature medicine*, 18(3), 429–435.

Elliott, P. R., Liu, H., Pastok, M. W., Grossmann, G. J., Rigden, D. J., Clague, M. J., . . . Barsukov, I. L. (2011). Structural variability of the ubiquitin specific protease DUSP-UBL double domains. *FEBS letters*, 585(21), 3385–3390.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 4), 486–501.

Erlekam, F., Zumbansen, M., & Weber, M. (2022). Parameter estimation on multivalent ITC data sets. *Scientific reports*, 12(1), 13402.

Faesen, A., Luna-Vargas, M. P., & Sixma, T. K. (2012). The role of UBL domains in ubiquitin-specific proteases. *Biochemical Society transactions*, 40(3), 539–545.

Faesen, A., Luna-Vargas, M. P., Geurink, P. P., Clerici, M., Merkx, R., van Dijk, W. J., . . . Sixma, T. K. (2011). The differential modulation of USP activity by internal regulatory domains, interactors and eight ubiquitin chain types. *Chemistry & biology*, 18(12), 1550–1561.

Falconer, J. R. (2016). Applications of isothermal titration calorimetry - the research and technical developments from 2011 to 2015. *Journal of molecular recognition : JMR*, 29(10), 504–515.

Fountain, M. D., Oleson, D. S., Rech, M. E., Segebrecht, L., Hunter, J. V., McCarthy, J. M., . . . Bi. (2019). Pathogenic variants in USP7 cause a neurodevelopmental disorder with speech delays, altered behavior, and neurologic anomalies. *Genetics in medicine : official journal of the American College of Medical Genetics*, 21(8), 1797–1807.

Fraser, J. S., Merlie, J. P., Jr, E. N., Weisfield, S. R., Mignot, T., Wemmer, D. E., . . . Alber, T. (2007). An atypical receiver domain controls the dynamic polar localization of the Myxococcus xanthus social motility protein FrzS. *Molecular microbiology*, 65(2), 319–332.

Freyer, M. W., & Lewis, E. A. (2008). Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods in cell biology*, 84, 79–113.

Fujishiro, T., Kahnt, J., Ermler, U., & Shima, S. (2015). Protein-pyridinol thioester precursor for biosynthesis of the organometallic acyl-iron ligand in [Fe]-hydrogenase cofactor. *Nature communications*, 6, 6895.

Gjonaj, L., Sapmaz, A., Flierman, D., Janssen, G. M., van Veelen, P. A., & Ovaa, H. (2019). Development of a DUB-selective fluorogenic substrate. *Chemical science*, 10(44), 10290–10296.

Grabowski, M., Cooper, D. R., Brzezinski, D., Macnar, J. M., Shabalin, I. G., Cymborowski, M., . . . Minor, W. (2021). Synchrotron Radiation as a Tool for Macromolecular X-Ray Crystallography: a XXI Century Perspective. *Nuclear instruments & methods in physics research. Section B, Beam interactions with materials and atoms*, 489, 30–40.

GraphPad Software, San Diego. (2023). *GraphPad Prism version 8.0.0 for Windows*. Retrieved from www.graphpad.com: www.graphpad.com

Guo, T., Tang, H., Yuan, Z., Zhang, E., & Wang, X. (2022). The Dual Role of USP11 in Cancer. *Journal of oncology*, 2022, 9963905.

Guo, T., Tang, H., Yuan, Z., Zhang, E., & Wang, X. (2022). The Dual Role of USP11 in Cancer. *Journal of oncology*, 9963905.

Harper, S., Gratton, H., Cornaciu, I., Oberer, M., Scott, D. J., Emsley, J., & Dreveny, I. (2014). Structure and catalytic regulatory function of ubiquitin specific protease 11 N-terminal and ubiquitin-like domains. *Biochemistry*, 53(18), 2966–2978.

Harrigan, J. A., Jacq, X., Martin, N. M., & Jackson, S. P. (2018). Deubiquitylating enzymes and drug discovery: emerging opportunities. *Nature reviews. Drug discovery*, 17(1), 57–78.

Hershko, A., & Ciechanover, A. (1998). The ubiquitin system. *Annual review of biochemistry*, 67, 425–479.

Hetfeld, B. K., Helfrich, A., Kapelari, B., Scheel, H., Hofmann, K., Guterman, A., . . . Dubiel, W. (2005). The zinc finger of the CSN-associated deubiquitinating enzyme USP15 is essential to rescue the E3 ligase Rbx1. *Current biology : CB*, 15(13), 1217–1221.

Hjortland, N. M., & Mesecar, A. D. (2016). Steady-state kinetic studies reveal that the anti-cancer target Ubiquitin-Specific Protease 17 (USP17) is a highly efficient deubiquitinating enzyme. *Archives of biochemistry and biophysics*, 612, 35–45.

Hu, B., Deng, T., Ma, H., Liu, Y., Feng, P., Wei, D., . . . Ye, M. (2019). Deubiquitinase DUB3 Regulates Cell Cycle Progression via Stabilizing Cyclin A for Proliferation of Non-Small Cell Lung Cancer Cells. *Cells*, 8(4), 297.

Hu, M., Chen, H., Han, C., Lan, J., Xu, Y., Li, C., . . . Lou, M. (2016). Expression and functional implications of USP17 in glioma. *Neuroscience letters*, 616, 125–131.

Hu, M., Li, P., Li, M., Li, W., Yao, T., Wu, J. W., . . . Shi, Y. (2002). Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell*, 111(7), 1041–1054.

Huang, X., Langelotz, C., Hetfeld-Pechoc, B. K., Schwenk, W., & Dubiel, W. (2009). The COP9 signalosome mediates beta-catenin degradation by deneddylation and blocks adenomatous polyposis coli destruction via USP15. *Journal of molecular biology*, 391(4), 691–702.

Ideguchi, H., Ueda, A., Tanaka, M., Yang, J., Tsuji, T., Ohno, S., . . . Ishigatsubo, Y. (2002). Structural and functional characterization of the USP11 deubiquitinating enzyme, which interacts with the RanGTP-associated protein RanBPM. *The Biochemical journal*, 367(Pt 1), 87–95.

Ilari, A., & Sevano, C. (2008). Protein Structure Determination by X-Ray Crystallography. *Keith, J.M. (eds) Bioinformatics. Methods in Molecular Biology™*, vol 452. Humana Press.

Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H. M., Duhr, S., Baaske, P., & Breitsprecher, D. (2014). MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure*, (1077) 101-113.

Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta crystallographica. Section A, Foundations of crystallography*, 47 (Pt 2), 110–119.

Kategaya, L., Di Lello, P., Rougé, L., Pastor, R., Clark, K. R., Drummond, J., . . . Stiffle. (2017). USP7 small-molecule inhibitors interfere with ubiquitin binding. *Nature*, 550(7677), 534–538.

Kim, R. Q., van Dijk, W. J., & Sixma, T. K. (2016). Structure of USP7 catalytic domain and three Ubl-domains reveals a connector α-helix with regulatory role. *Journal of structural biology*, 195(1), 11–18.

Komander, D., Clague, M., & Urbe, S. (2009). Breaking the chains: Structure and function of the deubiquitinases. *Nature reviews molecular cell biology*, 10, 550-563.

Koulich, E., Li, X., & DeMartino, G. N. (2008). Relative structural and functional roles of multiple deubiquitylating proteins associated with mammalian 26S proteasome. *Molecular biology of the cell*, 19(3), 1072–1082.

Krishnamoorthy, G. K., Alluvada, P., Hameed Mohammed Sherieff, S., Kwa, T., & Krishnamoorthy, J. (2019). Isothermal titration calorimetry and surface plasmon resonance analysis using the dynamic approach. *Biochemistry and biophysics reports*, 21, 100712.

Krojer, T., Talon, R., Pearce, N., Collins, P., Douangamath, A., Brandao-Neto, J., . . . von Delft, F. (2017). The XChemExplorer graphical workflow tool for routine or large-scale protein-ligand structure determination. *Acta crystallographica. Section D, Structural biology*, 73(Pt 3), 267–278.

Kwasna, D., Abdul Rehman, S. A., Natarajan, J., Matthews, S., Madden, R., De Cesare, V., . . . Kulathu, Y. (2018). Discovery and Characterization of

ZUFSP/ZUP1, a Distinct Deubiquitinase Class Important for Genome Stability. *Molecular cell*, 70(1), 150–164.e6.

Lamanna, W. C., Holzmann, J., Cohen, H. P., Guo, X., Schweigler, M., Stangler, T., . . . Schiestl, M. (2018). Maintaining consistent quality and clinical performance of biopharmaceuticals. *Expert opinion on biological therapy*, 18(4), 369–379.

Lange, S. M., Armstrong, L. A., & Kulathu, Y. (2022). Deubiquitinases: From mechanisms to their inhibition by small molecules. *Molecular cell*, 82(1), 15–29.

Lee, C. F., Allen, M. D., Bycroft, M., & Wong, K. B. (2005). Electrostatic interactions contribute to reduced heat capacity change of unfolding in a thermophilic ribosomal protein l30e. *Journal of molecular biology*, 348(2), 419–431.

Lee, E. W., Seong, D., Seo, J., Jeong, M., Lee, H. K., & Song, J. (2015). USP11-dependent selective cIAP2 deubiquitylation and stabilization determine sensitivity to Smac mimetics. *Cell death and differentiation*, 22(9), 1463–1476.

Li, X., Liu, Z., Xia, C., Yan, K., Fang, Z., & Fan, Y. (2022). SETD8 stabilized by USP17 epigenetically activates SREBP1 pathway to drive lipogenesis and oncogenesis of ccRCC. *Cancer letters*, 527, 150–163.

Liao, Y., Zhou, D., Wang, P., Yang, M., & Jiang, N. (2022). Ubiquitin specific peptidase 11 as a novel therapeutic target for cancer management. *Cell death discovery*, 8(1), 292.

Lin, Y., Wang, Y., Shi, Q., Yu, Q., Liu, C., Feng, J., . . . Wu, Y. (2017). Stabilization of the transcription factors slug and twist by the deubiquitinase dub3 is a key requirement for tumor metastasis. *Oncotarget*, 8(43), 75127–75140.

Liu, T., Yu, J., Deng, M., Yin, Y., Zhang, H., Luo, K., . . . … Lou, Z. (2017). CDK4/6-dependent activation of DUB3 regulates cancer metastasis through SNAIL1. *Nature communications*, 8, 13923.

Lu, C. H., Yeh, D. W., Lai, C. Y., Liu, Y. L., Huang, L. R., Lee, A. Y., . . . Chuang, T. H. (2018). USP17 mediates macrophage-promoted inflammation and stemness in lung cancer cells by regulating TRAF2/TRAF3 complex formation. *Oncogene*, 37(49), 6327–6340.

March, E., & Farrona, S. (2018). Plant Deubiquitinases and Their Role in the Control of Gene Expression Through Modification of Histones. *Frontiers in plant science*, 8, 2274.

Maurer, S. (2021). Developing novel fusion tags to facilitate protein production and structure determination. PhD thesis, University of Nottingham.

Maurer, T., & Wertz, I. E. (2016). Length Matters: MINDY Is a New Deubiquitinase Family that Preferentially Cleaves Long Polyubiquitin Chains. *Molecular cell*, 63(1), 4–6.

McFarlane, C., Kelvin, A. A., de la Vega, M., Govender, U., Scott, C. J., Burrows, J. F., & Johnston, J. A. (2010). The deubiquitinating enzyme USP17 is highly expressed in tumor biopsies, is cell cycle regulated, and is required for G1-S progression. *Cancer research*, 70(8), 3329–3339.

McFarlane, C., McFarlane, S., Paul, I., Arthur, K., Scheaff, M., Kerr, K., . . . Johnston, J. A. (2013). The deubiquitinating enzyme USP17 is associated with non-small cell lung cancer (NSCLC) recurrence and metastasis. *Oncotarget*, 4(10), 1836–1843.

Mevissen, T. E., Hospenthal, M. K., Geurink, P. P., Elliott, P. R., Akutsu, M., Arnaudo, N., . . . & Komander, D. (2013). OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell*, 154(1), 169–184.

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

Oestergaard, V. H., Langevin, F., Kuiken, H. J., Pace, P., Niedzwiedz, W., Simpson, L. J., . . . Patel, K. J. (2007). Deubiquitination of FANCD2 is required for DNA crosslink repair. *Molecular cell*, 28(5), 798–809.

Oliveira, R. I., Guedes, R. A., & Salvador, J. A. (2022). Highlights in USP7 inhibitors for cancer treatment. *Frontiers in chemistry*, 10, 1005727.

Oss-Ronen, L., Sarusi, T., & Cohen, I. (2022). Histone Mono-Ubiquitination in Transcriptional Regulation and Its Mark on Life: Emerging Roles in Tissue Development and Disease. *Cells*, 11(15), 2404.

Otwinowski, Z., & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in enzymology*, 276, 307–326.

Pardon, E., Laeremans, T., Triest, S., Rasmussen, S. G., Wohlkönig, A., Ruf, A., . . . Steyaert, J. (2014). A general protocol for the generation of Nanobodies for structural biology. *Nature protocols*, 9(3), 674–693.

Pauli, E. K., Chan, Y. K., Davis, M. E., Gableske, S., Wang, M. K., Feister, K. F., & Gack, M. U. (2014). The ubiquitin-specific protease USP15 promotes RIG-I-mediated antiviral signaling by deubiquitylating TRIM25. *Science signaling*, 7(307), ra3.

Pereg, Y., Liu, B. Y., O'Rourke, K. M., Sagolla, M., Dey, A., Komuves, L., . . . Dixit, V. M. (2010). Ubiquitin hydrolase Dub3 promotes oncogenic transformation by stabilizing Cdc25A. *Nature cell biology*, 12(4), 400–406.

Pierce, M. M., Raman, C. S., & Nall, B. T. (1999). Isothermal titration calorimetry of protein-protein interactions. *Methods (San Diego, Calif.)*, 19(2), 213–221.

Pozhidaeva, A., & Bezsonova, I. (2019). USP7: Structure, substrate specificity, and inhibition. *DNA repair*, 76, 30–39.

Protein Data Bank. (2023, March 31). *PDB Statistics: Growth of Structures from X-ray Crystallography Experiments Released per Year*. Retrieved from https://www.rcsb.org/: https://www.rcsb.org/stats/growth/growth-xray

Qiu, X. B., Ouyang, S. Y., Li, C. J., Miao, S., Wang, L., & Goldberg, A. L. (2006). hRpn13/ADRM1/GP110 is a novel proteasome subunit that binds the deubiquitinating enzyme, UCH37. *The EMBO journal*, 25(24), 5742–5753.

Reimund, M., Kovrov, O., Olivecrona, G., & Lookene, A. (2017). Lipoprotein lipase activity and interactions studied in human plasma by isothermal titration calorimetry. *Journal of lipid research*, 58(1), 279–288.

Reyes, F., Shanks, J., Komander, D., & Wilkinson, K. (2008). Recognition of polyubiquitin isoforms by the multiple ubiquitin binding modules of isopeptidase T. *The Journal of biological chemistry*, 283(28), 19581–19592.

Reyes-Turcu, F. E., & Wilkinson, K. D. (2009). Polyubiquitin binding and disassembly by deubiquitinating enzymes. *Chemical reviews*, 109(4), 1495–1508.

Rossmann, M. (1990). The molecular replacement method. *Acta crystallographica. Section A, Foundations of crystallography*, 46 (Pt 2), 73–82.

Rut, W., Zmudzinski, M., Snipas, S. J., Bekes, M., Huang, T. T., & Drag, M. (2020). Engineered unnatural ubiquitin for optimal detection of deubiquitinating enzymes. *Chemical science*, 11(23), 6058–6069.

Scaglione, K. M., Basrur, V., Ashraf, N. S., Konen, J. R., Elenitoba-Johnson, K. S., Todi, S. V., & Paulson, H. L. (2013). The ubiquitin-conjugating enzyme (E2) Ube2w ubiquitinates the N terminus of substrates. *The Journal of biological chemistry*, 288(26), 18784–18788.

Schauer, N. J., Magin, R. S., Liu, X., Doherty, L. M., & Buhrlage, S. J. (2020). Advances in Discovering Deubiquitinating Enzyme (DUB) Inhibitors. *Journal of medicinal chemistry*, 63(6), 2731–2750.

Schoenfeld, A. R., Apgar, S., Dolios, G., Wang, R., & Aaronson, S. A. (2004). BRCA2 is ubiquitinated in vivo and interacts with USP11, a deubiquitinating enzyme that exhibits prosurvival function in the cellular response to DNA damage. *Molecular and cellular biology*, 24(17), 7444–7455.

Schrödinger, LLC. (2023). *The Pymol Molecular Graphics System Version 2.0*. Retrieved from Pymol.org: https://pymol.org/

Schweitzer, K., Bozko, P. M., Dubiel, W., & Naumann, M. (2007). CSN controls NF-kappaB by deubiquitinylation of IkappaBalpha. *The EMBO journal*, 26(6), 1532–1541.

Shaw, J. A., Page, K., Blighe, K., Hava, N., Guttery, D., Ward, B., . . . Coombes, R. C. (2012). Genomic analysis of circulating cell-free DNA infers breast cancer dormancy. *Genome research*, 22(2), 220–231.

Shi, Y. (2014). A glimpse of structural biology through X-ray crystallography. *Cell*, 159(5), 995–1014.

Shin, J. M., Yoo, K. J., Kim, M. S., Kim, D., & Baek, K. H. (2006). Hyaluronan- and RNA-binding deubiquitinating enzymes of USP17 family members associated with cell viability. *BMC genomics*, 7, 292.

Smith, B. E., Wang, S. L., Jaime-Figueroa, S., Harbin, A., Wang, J., Hamman, B. D., & Crews, C. M. (2019). Differential PROTAC substrate specificity dictated by orientation of recruited E3 ligase. *Nature communications*, 10(1), 131.

Song, M. S., Salmena, L., Carracedo, A., Egia, A., Lo-Coco, F., Teruya-Feldstein, J., & Pandolfi, P. P. (2008). The deubiquitinylation and localization of PTEN are regulated by a HAUSP-PML network. *Nature*, 455(7214), 813–817.

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

Sparks, R. P., & Fratti, R. (2019). Use of Microscale Thermophoresis (MST) to Measure Binding Affinities of Components of the Fusion Machinery. *Methods in molecular biology (Clifton, N.J.)*, 1860, 191–198.

Sparks, R. P., Lawless, W., Arango, A. S., Tajkhorshid, E., & Fratti, R. A. (2022). Use of Microscale Thermophoresis to Measure Protein-Lipid Interactions. *Journal of visualized experiments : JoVE*, (180).

Stegmeier, F., Rape, M., Draviam, V. M., Nalepa, G., Sowa, M. E., Ang, X. L., . . . Elledge, S. J. (2007). Anaphase initiation is regulated by antagonistic ubiquitination and deubiquitination activities. *Nature*, 446(7138), 876–881.

Studier, W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein expression and purification*, 41(1), 207–234.

Sun, J., Shi, X., Mamun, M. A., & Gao, Y. (2020). The role of deubiquitinating enzymes in gastric cancer. *Oncology letters*, 19(1), 30–44.

Swatek, K. N., & Komander, D. (2016). Ubiquitin modifications. *Cell Res*, 26(4): 399–422.

Ten Eyck, L. (1973). Crystallographic fast fourier transforms. *Acta Crystallogr A*, A29:183–91.

Thakur, A. S., Robin, G., Guncar, G., Saunders, N. F., Newman, J., Martin, J. L., & Kobe, B. (2007). Improved success of sparse matrix protein crystallization screening with heterogeneous nucleating agents. *PloS one*, 2(10), e1091.

Thorn, A., Parkhurst, J., Emsley, P., Nicholls, R. A., Vollmar, M., Evans, G., & Murshudov, G. N. (2017). AUSPEX: a graphical tool for X-ray diffraction data analysis. *Acta crystallographica. Section D, Structural biology*, 73(Pt 9), 729–737.

Tiwari, P. B., Bencheqroun, C., Lemus, M., Shaw, T., Kouassi-Brou, M., Alaoui, A., & Üren, A. (2021). SPRD: a surface plasmon resonance database of common factors for better experimental planning. *BMC molecular and cell biology*, 22(1), 17.

Tso, S. C., Chen, Q., Vishnivetskiy, S. A., Gurevich, V. V., Iverson, T. M., & Brautigam, C. A. (2018). Using two-site binding models to analyze microscale thermophoresis data. *Analytical biochemistry*, 540-541, 64–75.

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., . . . Takanen, J. O. (2015). Proteomics. Tissue-based map of the human proteome. *Science (New York, N.Y.)*, 347(6220), 1260419.

Velazquez-Campoy, A., Leavitt, S. A., & Freire, E. (2004). Characterization of protein-protein interactions by isothermal titration calorimetry. *Methods in molecular biology*, 261, 35–54.

Vijay-Kumar, S., Bugg, C. E., & Cook, W. J. (1987). Structure of ubiquitin refined at 1.8 A resolution. *Journal of molecular biology*, 194(3), 531–544.

Vlasschaert, C., Xia, X., Coulombe, J., & Gray, D. A. (2015). Evolution of the highly networked deubiquitinating enzymes USP4, USP15, and USP11. *BMC evolutionary biology*, 15, 230.

Ward, S. J., Gratton, H. E., Indrayudha, P., Michavila, C., Mukhopadhyay, R., Maurer, S. K., . . . Dreveny, I. (2018). The structure of the deubiquitinase USP15 reveals a misaligned catalytic triad and an open ubiquitin-binding channel. *The Journal of biological chemistry*, 293(45), 17362–17374.

Wienken, C. J., Baaske, P., Duhr, S., & Braun, D. (2011). Thermophoretic melting curves quantify the conformation and stability of RNA and DNA. *Nucleic acids research*, 39(8), e52.

Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D., & Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. *Nature communications*, 1, 100.

Wilkinson, K. (1995). Roles of ubiquitinylation in proteolysis and cellular regulation. *Annual review of nutrition*, 15, 161–189.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., . . . W. (2011). Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography*, 67(Pt 4), 235–242.

Xu, P., & Peng, J. (2008). Characterization of polyubiquitin chain structure by middle-down mass spectrometry. *Analytical chemistry*, 80(9), 3438–3444.

Ye, Y., Scheel, H., Hofmann, K., & Komander, D. (2009). Dissection of USP catalytic domains reveals five common insertion points. *Molecular bioSystems*, 5(12), 1797–1808.

Yu, M., Liu, K., Mao, Z., Luo, J., Gu, W., & Zhao, W. (2016). USP11 Is a Negative Regulator to γH2AX Ubiquitylation by RNF8/RNF168. *The Journal of biological chemistry*, 291(2), 959–967.

Yuan, T., Yan, F., Ying, M., Cao, J., He, Q., Zhu, H., & Yang, B. (2018). Inhibition of Ubiquitin-Specific Proteases as a Novel Anticancer Therapeutic Strategy. *Frontiers in pharmacology*, 9, 1080.

Zeng, C., Zhao, C., Ge, F., Li, Y., Cao, J., Ying, M., . . . Zhu, H. (2020). Machado-Joseph Deubiquitinases: From Cellular Functions to Potential Therapy Targets. *Frontiers in pharmacology*, 11, 1311.

Zhang, E., Shen, B., Mu, X., Qin, Y., Zhang, F., Liu, Y., . . . Fan, Y. (2016). Ubiquitin-specific protease 11 (USP11) functions as a tumor suppressor through deubiquitinating and stabilizing VGLL4 protein. *American journal of cancer research*, 6(12), 2901–2909.

Zhang, W., & Sidhu, S. S. (2014). Development of inhibitors in the ubiquitination cascade. *FEBS letters*, 588(2), 356–367.

Zhu, Y., Pless, M., Inhorn, R., B., M.-P., & D'Andrea, A. D. (1996). The murine DUB-1 gene is specifically induced by the betac subunit of interleukin-3 receptor. *Molecular and cellular biology*, 16(9), 4808–4817.

Zou, Q., Jin, J., Hu, H., Li, H. S., Romano, S., Xiao, Y., . . . Sun, S. C. (2014). USP15 stabilizes MDM2 to mediate cancer-cell survival and inhibit antitumor T cell responses. *Nature immunology*, 15(6), 562–570.