Elucidating the catalytic and structural functions of ubiquitin specific proteases (USPs) 2 and 17

Thesis submitted to the University of Nottingham for the degree of Master of Pharmacy, Master of Research.

June 2026

Adam Albera

20335326

Supervisors: Ingrid Dreveny, Jonas Emsley

School of Pharmacy, University of Nottingham.

Contents:

1. Abstract

Ubiquitin is an essential regulatory protein, being the main component of the ubiquitination process, a post-translational modification whereby target proteins (substrates) are sent to the 26s proteasome for degradation. Deubiquitinating enzymes (DUBs) are a category of enzymes which hydrolyse the isopeptide bond between the ubiquitin and substrate, enabling the removal of monoubiquitin or polyubiquitin chains, among the sub-categories of DUBs, ubiquitin specific proteases (USPs) are the largest. USP2 and USP17 are among the most essential USPs, as both regulate the cell cycle progression, specifically through the regulation of cyclin D and c-myc respectively, meaning that cell proliferation is controlled and prevents tumorigenesis. However, aberrant overexpression of these proteins can be the cause of specific cancers such as non-small cell lung cancer and promote metastasis, therefore an effective and highly specific inhibitor is highly in demand, especially for USP17.

The aims for this project are to successfully purify the correct constructs which will be used for further analysis, identify and confirm the substrates that USP17 is able to cleave efficiently, through mutagenesis confirm that USP17 with a D332G mutation is able to cleave linear di-ubiquitin, using PR619 to identify the Ic50 of USP17, and test both proteins against multiple fluorescein bound probes.

The constructs which required expressing, were successfully purified, these include USP2CD and all the USP17 constructs. Through a series of gel cleavage assays, USP2 and the USP17 constructs were tested against 4 different substrates: K6-linked di-ubiquitin mimetic (K6), K11-linked di-ubiquitin mimetic (K11), GG-linked di-ubiquitin (GG), and linear di-ubiquitin. USP2 successfully cleaving all the substrates, and the activated USP17 constructs (non-C89S mutated) successfully cleaved K6, K11, and GG, with linear di-ubiquitin showing no activity, as the literature states. However, with the D332G mutation of the wildtype USP17, the enzyme showed very minor to no discernible improvements in the cleavage rate of linear di-ubiquitin, which was demonstrated using gel cleavage assays. Furthermore, with the aid of repeated fluorescence polarization assays, an Ic50 value for USP17 was determined at 1.549μM when used against PR-619, a small molecule inhibitor. Crystallization was attempted on two moieties: USP17D1D2-2GKG, and USP17D1D2-C89S2GKG/GG complex, both crystallizations utilised a Morpheus screen and a standard gel filtration buffer, yet no crystals were identifiable on both screens.

Overall, with the aid of more repeats, especially in crystallography, more conclusive data can be reached, however the current findings can be utilized to do further assay work, especially to determine the structural functions of USP17.

2. Graphical Abstract

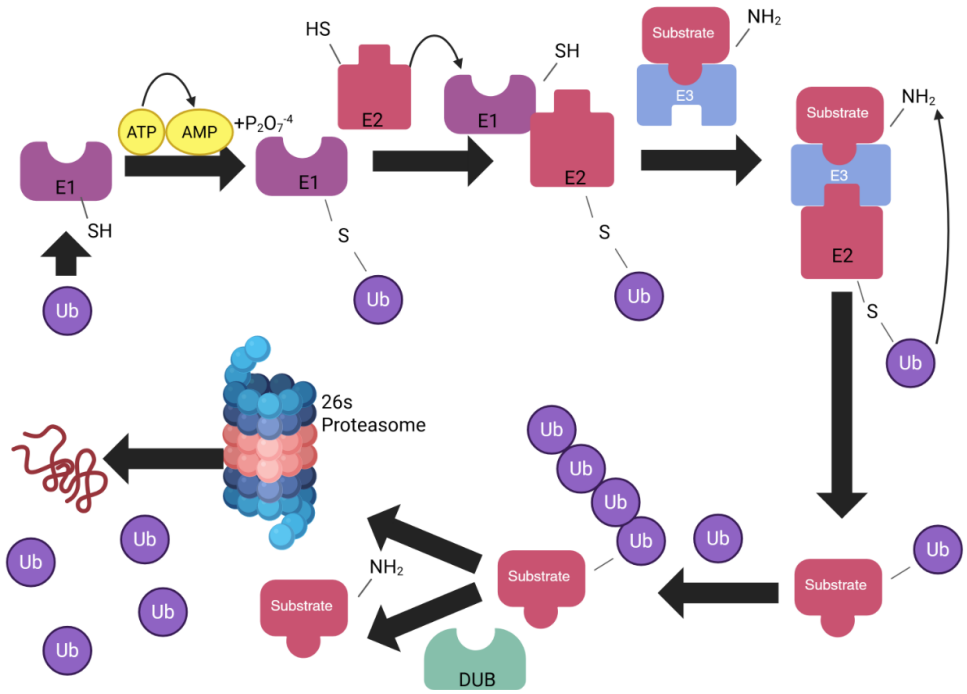


Figure 1- Graphical abstract demonstrating the eukaryotic ubiquitination/deubiquitination process, ubiquitin forms a covalent bond with the thiol group present on the E1 enzyme via the adenylated C-terminus on the ubiquitin (this reaction requires a hydrolysis of ATP to AMP and pyrophosphate). The E2 enzyme is able to acquire ubiquitin via a thioester transfer, while the E3 enzyme binds to the substrate via the active binding domain, with the E2 enzyme binding to the allosteric site of E3, an isopeptide bond is formed between the lysine residue of the substrate and the C-terminus present on the ubiquitin.

3. Introduction

3.1 The ubiquitination process

Ubiquitin is a 76 amino acid regulatory protein first discovered in 1981; its primary functions are related to post translational modifications via covalent bonding to substrates in eukaryotic organisms. Ubiquitin is involved in the process of ubiquitination, which is the bonding process of ubiquitin to a target substrate through the use of a 3-enzyme cascade[1]. Enzyme E1 is used to activate the ubiquitin through the use of ATP, additionally E1 charges E2 enzymes which are responsible for conjugation[2] with E3 ubiquitin ligases recruiting the substrates to the transfer of ubiquitin molecules via the C-terminal glycine to the substrate lysine residue[3]. However in rarer cases such as linear di-ubiquitin the bond is made using the methionine M1 bond present on the ubiquitin.[2] There is one E1 enzyme of its kind, while there are 20 E2 enzymes and hundreds of E3 enzymes.[4] Ubiquitin is also able to form multiple bonds with other ubiquitin moieties through polyubiquitination which can be formed via the methionine M1 or through 7 different lysine residues K6, K11, K27, K29, K33, K48, or K63.[2]

Significant linkages among ubiquitin molecules include K6-linked di-ubiquitin, a complex which has been recently discovered to bind to TAK1-binding protein 2 Npl4-zinc-finger domain (TAB2-NZF), activating transforming growth factor-beta-activated kinase 1 (TAK1)[5], otherwise referred to as MAP3K7. Functionally the protein is able to regulate the genes relating to cell proliferation and inflammation through mitogen-activated protein kinases (MAPK)[5]. Moreover, K6-linked di-ubiquitin has been discovered to bind to parkin proteins, molecules which are produced from the PARK2 gene and contribute directly to the development of Parkinson's disease, however to the contrary of the functionality of ubiquitin complexes, prevents the protein from degradation via the 26s proteasome[6]. However, upon taking into consideration the literature strictly focused on K6-linked di-ubiquitin, there seems to be a real lack of research directly relating to K6, its structure, and its functionality within the human body and how diseases may arise through aberrant expression. Furthermore K6-linked di-ubiquitin targets proteins to be directed to the 26s proteasome and hence degraded, yet with the research accomplished by Durcan et al. the conclusion was made to be that K6 promoted protein survival in Parkinson's disease instead. On the contrary K6 is not understood fully due to the glaring lack of independent research done on the protein, especially when compared to K11 polyubiquitin chains or without the involvement of other lysine residues.

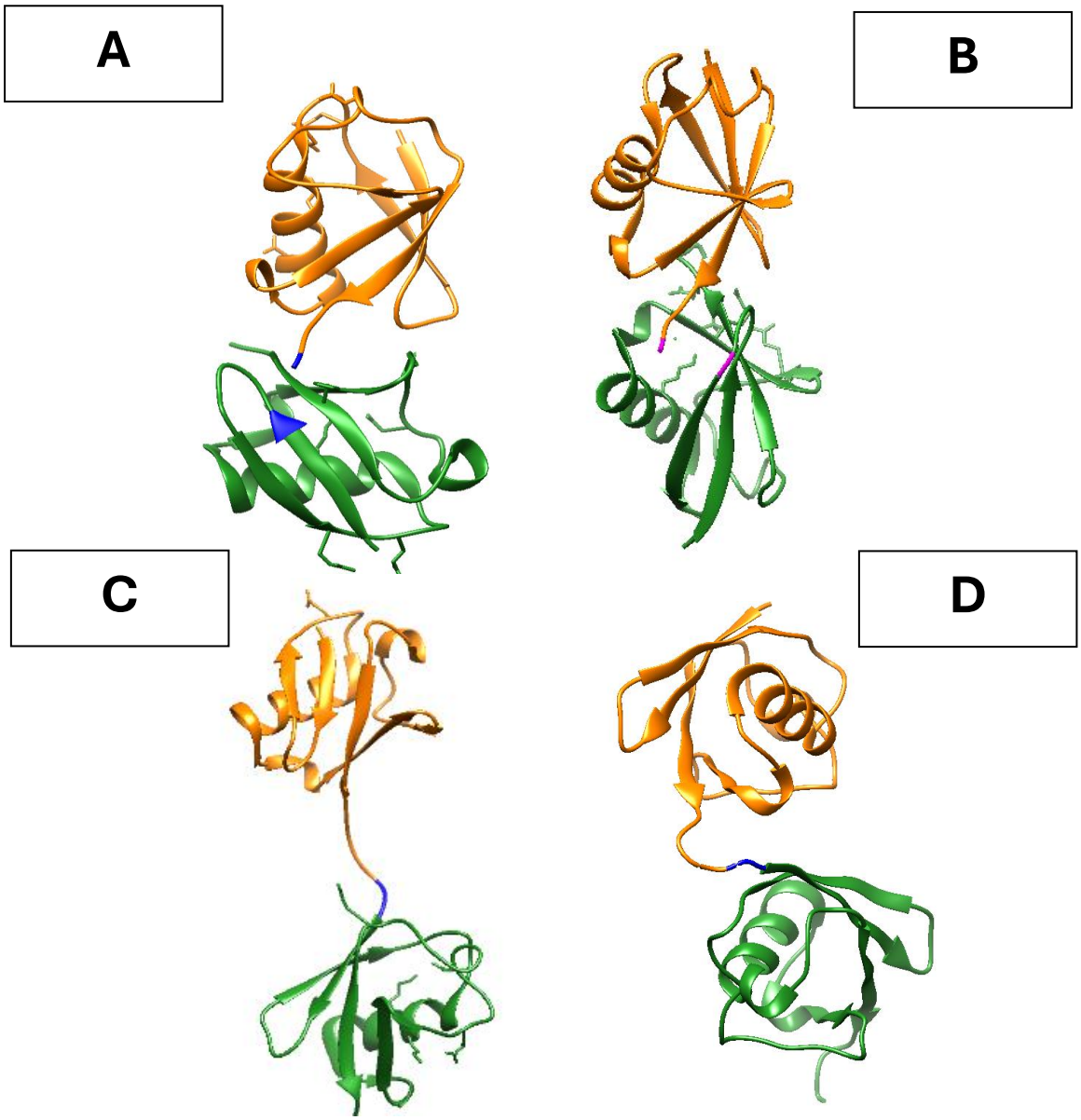


Figure 2- Crystalline structures of the four di-ubiquitin moieties of interest and colour coded using UCSF Chimera. (A): crystalline structure of K6-linked di-ubiquitin, the structure was obtained from protein databank (PDB) with the code 2XK5. (B): crystalline structure of K11-linked di-ubiquitin di-ubiquitin, the structure was obtained from PDB with the code 3NOB. (C): crystalline structure of linear (M1) di-ubiquitin, the structure was obtained from PDB with the code 2W9N. (D) predicted crystalline structure of GG-linked di-ubiquitin, the structure was predicted using Alphafold. Colour legend: orange- ubiquitin 1, green- ubiquitin 2, blue- residues involved in the linkage between ubiquitins, magenta- residues involved in the linkage of K11-linked di-ubiquitin (colour change was decided for better visualisation).

K11-linked di-ubiquitin chains have been involved in various biological processes, especially in yeast as the complex was discovered to be of high abundance along with K48 and K63[7], upon modifications in the environment, including removal of DSK2 (a protein involved in endoplasmic reticulum-associated ubiquitin-dependent protein catabolism[8]). Rad23, a protein responsible for nucleotide excision repair, consisting of a ubiquitin-like domain in its structure[9], removal of these proteins causes the abundance of K11 to increase by nearly triple, indicating the protein is selectively targeted by proteins that are capable of forming a binding. K11 has been noted to promote degradation of proteins, notably ubiquitin ligases such as anaphase promoting

complexes (APC/C), causing a defection of mitotic progression, ultimately leading to cell death through the inability of proliferation[10].

Interestingly, M1-linked (linear) di-ubiquitin is one of two ubiquitin complex which does not form an isopeptide bond with a lysine, instead a true peptide bond is formed between the C-terminus and N-terminal methionine of a second ubiquitin[11] (hence the term 'linear'), with the synthesis of the protein being catalysed by the linear ubiquitin chain assembly complex (LUBAC)[12]. As for the role of linear di-ubiquitin, along with K48, the proteins are able to regulate the nuclear factor- kappa beta (NF-$\kappa\beta$)[13], a transcription factor regulating inflammatory and immunity responses[14], particularly M1 polyubiquitin chains are able to activate the NF- $\kappa\beta$ complex, inducing cell death[13].

The second of the unique ubiquitin complexes is Ub-GG-Ub, sometimes referred to as Ub diglycine, generated through the formation of an isopeptide bond with an amine group present on the R group of a lysine residue to a methionine on a second ubiquitin molecule. The moiety was initially discovered in 1977 on histone A24 by scientists I.L Goldknopf and Harris Busch[15], due to the nature of these proteins being modified variants of linear di-ubiquitin[16], identifying these proteins is difficult despite the bottom-up mass spectrometry strategy[16]. As for the functionality of the protein, USP1 contains a Ub-GG-Ub complex within its structure, the complex enables the enzyme to recognise its substrates, with addition to catalysis[17], regarding any independent functions within a eukaryotic context, the information is scarce and not analysed to a substantial degree.

Substrates which have a ubiquitin molecule or a polyubiquitin chain are recognised by the 26s proteasome, an enzyme complex formed in eukaryotic cells, its main function is to degrade proteins through an ATP-dependant cascade[18]. The proteasome hydrolyses the structure of substrates where the unfolded polypeptides are translocated to a degradation chamber whereby proteolytic cleavage occurs[18]. The proteins which are degraded include a vast range with varying functionalities such as tumour suppressors, transcription factors, and cell cycle regulators[3].

3.2 Deubiquitination: the removal of ubiquitin

However, this process can be regulated through deubiquitination, which reverses the catalysation of ubiquitin to substrates, in humans this process can be achieved with the use of deubiquitinating enzymes (DUBs), there are 99 currently known DUBs split into 7 sub-families, with ubiquitin specific proteases (USPs) being the most prevalent with 56 known USPs[19]. USP2 is among the most abundant USPs in the human proteasome, its main function is to deubiqutinate cyclin D which allows cells to continue mitosis past the G1 phase[20]. Other functions of USP2 include stabilizing a large volume of oncoproteins such as TGF-b threonine/serine kinase receptors and fatty acid synthase[21] through the direct removal of ubiquitin, hence avoiding the degradation occurring via the proteasome.

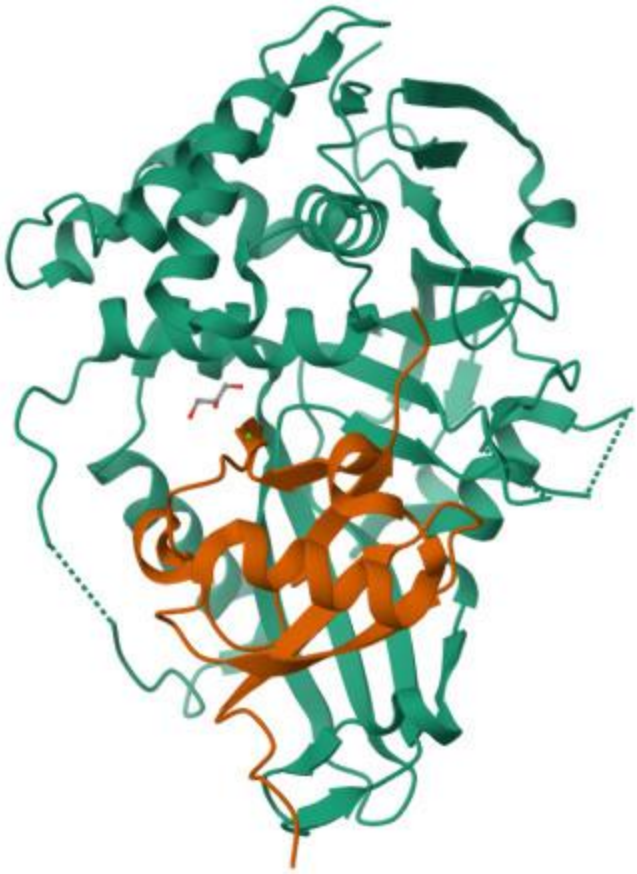


Figure 3- Crystalline structure of USP2WT with a mono-ubiquitin that contains a total of 4 mutations (S17G, N23K, S26T, and H29T), the structure can be found on the protein data bank (PDB) under the code 3V6E.

USP2 has been identified to cleave numerous bonds due to its nonspecific nature[22], including isopeptide bonds successfully, meaning USP2 is able to efficiently hydrolyse the bonds in any polyubiquitin moieties bound to lysine residues (K6, K11, K48), and linear polyubiquitin chains. USP2 is classified as a cysteine protease, due to one of the residues present in the catalytic triad being a cysteine at position 276, as the cysteine acts as a nucleophile, it is able to attack the isopeptide bond forming a thioester intermediate, this intermediate is then hydrolysed, releasing the ubiquitin moieties from either one another or the ubiquitinated protein[23]. This same mechanism of action is used to attack and hydrolyse the peptide bond in linear polyubiquitin chains[24].

However, aberrant expression of USP2 can cause tumorigenesis, mainly due to the stabilisation of fatty acid synthase, a protein present and correlated to malignant prostate cancer, overexpression of this DUB was found in 44% studied prostate tumours[25]. The literature also mentions that USP2 is also able to stabilize squamous cell carcinoma tumours, however the literature referenced by the author, does not mention this conclusion in the findings, hence the most plausible tumour which USP2 has a more significant and clear impact are prostate tumours. On the contrary, when USP2 is wrongly targeted for degradation, other tumours are formed, specifically due to the destabilisation of programmed-death ligand-1 (PD-L1), USP2 is unable to reverse ubiquitination of PD-L1 therefore this protein is degraded by the endoplasmic

reticulum-associated degradation (ERAD) pathway, causing the body to lose immunity to tumorigenesis[26].

USP17 (otherwise known as DUB3) is a DUB enzyme which is expressed during the G1 cell cycle phase and is involved in the Ras and guanosine triphosphate (GTP) enzymes through intracellular localisation[27]. Originally identified in mice, USP17 genes were discovered to be in a tandemly repeated mega-satellite sequence (RS447) with a copy number between 20-103 present on chromosomal sequences 4p16.1 and 8p23.1, with USP17 being expressed most prominently when induced by cytokines such as interleukin 4 and 6, which the article wrongly identifies as interleukin 2 and 3[28]. When USP17 is downregulated, the tumour cell is blocked from accessing the synthesis phase of the cell cycle as the localisation of oncogenic GTPases such as Ras is dysregulated[29].

In regard to the substrate specificity of USP17, the enzyme is most proficient at cleaving K48-linked polyubiquitin chains, the deubiquitination causes the previously ubiquitinated proteins to not be directed to the 26s proteasome, therefore avoiding cell death. There is also evidence of USP17 being able to fully hydrolyse other ubiquitin chains such as K11, K33, and K48 linked di/polyubiquitin chains, partially hydrolysing K6, K27, and K29[30], and incapable of hydrolysing linear (M1 bound) di-ubiquitin chains[31].

Interestingly, the oncogenic characteristics of USP17 have also been reported. The results indicated high level of USP17 in lung, colon, oesophagus and cervix tumour samples to promote cell G1-S transition and cell proliferation[32]. Besides, USP17 was overexpressed in non-small cell lung cancer tissues and patients with high level of USP17 had reduced survival[33]. In osteosarcoma, USP17 was up regulated in tumour tissues and cell lines and USP17 facilitated cell migration and invasion through deubiquitinating and stabilizing SMAD4[34].

The structure of USP17 contains a USP catalytic domain, which is composed of amino acids responsible for the binding, mainly being aspartic acid, glycine, and cysteine, which is located near to the N-terminus, with the hyaluronan (and RNA) binding motifs (HABM) being present at positions 401-409 and 445-453, which is nearer to the C-terminus[35].

3.3 Approved inhibitors for USP2 and USP17

Currently there are several inhibitors which are selective against the two DUBs, for instance ML364 is a selective inhibitor for USP2, which has an Ic50 of 1.1μm, the inhibition of USP2 results in cyclin D1 to not be deubiquitinated causing an increase in its degradation[36]. Therefore the cell cycle is terminated early causing a decrease in the proliferative potential of cancerous cells.

Recently, through synthesis optimisation studies, a derivative of ML364 was synthesized (LLK203), which shows dual-inhibitory effects against USP2 and USP8. The difference between the two compounds is the removal of the thiazol functional group, and the replacement of the ether functional group with a trifluoro methane[37]. A second, orally administered, ACE2 small molecule inhibitor also shows potent inhibitory effects against USP2, being MS102, the targeting of ACE2 is essential as the protein is regularly deubiquitinated by USP2, through inhibition, ACE2 slowly degrades, mitigating the risk of ACE2-dependent COVID-19[38].

While no direct and specific inhibitor for USP17 is available in the market at the moment, palbociclib, a cyclin dependant kinase 4/6 inhibitor[39], has shown promise in identifying USP17 as a potential target, specifically in breast cancer. Since one of the various functions of USP17

is to stabilize a protein known as SNAIL1[40], a transcription factor which binds to E-boxes on DNA sequences to block the activation of particular genes such as E-cadherin[41], via deubiquitination. Palbociclib was also shown to indirectly affect USP17 such as targetting and breakdown USP17 substrates such as melanoma-associated antigen (MAGEA3), a protein when overexpressed, is an indicator of melanoma and lung cancer[42], and bromodomain containing protein 4 (BRD4), a protein which directly affects the phosphorylation of SER2 in RNA polymerase II on the C-terminal domain. With the process being mediated by the positive transcription elongation factor b[43] when aberrantly expressed, BRD4 upregulates the production of oncogenes such as c-MYC, ultimately causing tumorigenesis[44].

3.4 Mimetic usage in research

A protein mimetic, by definition, is a synthetic moiety specifically engineered to capture and replicate the biophysical and chemical properties of the original biological molecule, a successful mimetic is able to have identical protein-protein interactions and tertiary structure[45]. Usually, mimetics are chosen over their counterpart in drug discovery, the specific reasons vary between the proteins including prolonging storage life, decrease production costs, adding essential PTMs to improve characteristics such as hydrophilicity, or adding any tags[45]. Two mimetics were used in this research, being the K6 and K11-linked di-ubiquitin mimetics, as the costs of synthesizing a large volume of these proteins was found to be more efficient. However, both proteins underwent modifications.

Isopeptide bonds are covalent amine bonds, occurring between the carboxylic acid group of residue 1 and the amine group found on the R-group of a lysine residue[46], however due to the high conformation specificity of the bond and the complex localization of both ubiquitin substrates, engineering an isopeptide bond is a difficult procedure[47]. As a result, ubiquitin constructs containing this bond (any lysine bound polyubiquitin) need to be truncated (residues cleaved), therefore the K6 and K11-linked di-ubiquitin constructs had the first 6 and 11 residues respectively truncated to ensure the binding site between the two ubiquitin moieties was near identical to the in vivo counterparts. As the mimetics exhibited the same catalytic properties as described in literature, and the structures remained stable despite the truncation, these mimetics were cleared for use to analyse the catalytic activity of USP17.

3.5 Protein Purification

Proteins are crucial to the human body, with further analysis being crucial to understanding and discovering their catalytic and structural properties. However an appropriate technique is required to be chosen to isolate the protein of interest, as transforming an E. coli strain will result in a sample which contains various cellular components and products which are not desired, these are known as contaminants. Protein purification is a multiple step technique where the end result is a fully purified sample of protein of interest, with the concentration correlating to the success of purification. The chances of a successful purification can increase dramatically on multiple parameters such as the type of column used, the concentration of imidazole used, and through the appropriate selection of fractions collected.

Initially, the desired DNA sample undergoes plasmid transformation using a heat shock method, in conjunction with competent bacteria. In the case of this project NovaBlue cells, a genetically modified variant of E. coli cells specifically engineered to have high transformation efficiency- due to the cells being an E. coli K-12 derivative containing recA and endA mutations which aid in facilitation of foreign plasmid DNA. Once colonies have been successfully

harvested, single colonies are chosen to overexpress- the medium chosen is preferential, with Yeast Extract Tryptone (XYT media), originally designed for rapid M13 bacteriophage propagation[48] provide an increased concentration of tryptone and yeast than Lysogeny Broth (LB). This allows higher density E. coli to increase production of the protein of interest in higher volumes and more efficiently[49], however the rich composition of the broth can cause overwhelming metabolic and genetic stress on the cells as E. coli are not naturally optimised to inhabit such environments[50]. LB media is a more widely, less optimised medium used for overexpression, while still having the essential nutrients required for bacterial growth, since the yeast and tryptone are significantly lower. This lowers the risk of potential stress for the E. coli cells; however the chance of a higher end protein yield is lower as the overall concentration of nutrients is lower than XYT media, therefore LB media is safer, however protein production may be lower than anticipated.

Upon retrieving a pellet from the overexpressed substance, regardless of the medium used, sonication is an essential purpose to ensure the E. coli cell structure is fully lysed. Sonication is a technique which utilizes ultrasound waves to cause cellular barrier disruption through the formation of a cavitation bubble, a vapour void occurring after liquid vapour pressure is higher than the static pressure (this phenomenon is most commonly seen in high velocity propeller-like motion in liquid). E. coli cells present within the vicinity of the cavitation bubble have their plasma membranes and cell walls pulverized[51], meaning that intracellular components can be easily retrieved and analysed.

When the sonicated substance is prepared, it is slowly incorporated into a Histrap column, a chromatography column used for immobilizing proteins tagged with a polyhistidine tag (His-tags), a negatively charged affinity tag which becomes immobilized upon interacting with positively charged transition metal ions[52]. Ni^{2+} displays the highest affinity towards His-tags[53], therefore if the column is purchased with no added component, nickel sulphate ($NiSO_4$) solution is used to recharge the column, on the contrary ethylene diamine tetra-acetic acid (EDTA) is utilized for stripping already used nickel sulphate before recharging. However, an Excel Histrap column is resistant to EDTA stripping, therefore can be reused constantly and be washed with ethanol and deionized water to remove excess sample.

The most crucial equipment is the protein purifier, a chromatography system utilizing collection vessels through 2 active pumps[54], one of the most common and consistent systems being the Äkta Start by Cytiva. Once the column is incorporated into the flow system, appropriate buffer is pumped into the column at a manually chosen rate, the proteins present within the column are then separated from the column and enter fraction tubes- proteins of interest appear on the chromatogram with significantly higher absorbance. In the case of nickel columns and the preferable choice of buffer, buffers with free imidazole are most efficient- imidazole acts as a competitor for the metal ion binding when interacting with histag proteins. Once a high enough concentration is reached, the protein begins to elute from the column and be collected- an essential action to consider is to obtain 2 buffers. One buffer with a low concentration of imidazole to prohibit the binding of naturally occurring histag proteins[55], one with high concentration for the elution of the desired protein, indicative of a peak appearing on the chromatogram, the solution collected in the fraction tubes is then analysed to make absolute certainty that the correct protein was eluted.

3.6 General concepts and methodologies of SDS-PAGE

Identification of the protein collected from purification is essential, as the presence of a peak is often a sufficient indication of protein, to ensure the correct protein was eluted, the most frequent method of identification is through the use of sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The analytical technique was established in 1970 by UK Laemmli, an MRC laboratory of molecular biology in Cambridge[56], with the methodology remaining consistent until present day, initially a protein is mixed thoroughly with an anionic detergent, most often being sodium dodecyl sulphate, allowing movement to occur through a gel matrix towards the cathode. Denaturation is compulsory in electrophoresis as the proteins are required to be unfolded, as otherwise the proteins will be aggregated leading to unclear readings regarding molecular weight[57].

As the name suggests, the most crucial ingredient in gels used for SDS-PAGE is acrylamide, a carcinogenic chemical typically released in food rich with starch upon applying heat and in cigarette smoke, in combination with tetramethylethylenediamine (TEMED) and ammonium persulfate (APS) the acrylamide begins to undergo free radical polymerization. The APS generates the release of these free radicals and TEMED acting as a catalyst to increase the speed of this release[58]. The present free radicals multiply in numbers upon coming into contact with acrylamide monomers (converting into more free radicals), inactive acrylamide monomers begin forming polymer chains once reacting with the free radicals[59], the chains initiate crosslinking between bis-acrylamide particles forming pores whereby the samples are able to navigate through upon being exposed to a positive charge[60]. Complete depletion of acrylamide monomers causes the radicalization to terminate[61], the result being a stable gel lattice.

Once loaded and activated, the proteins begin migration through the lattice, organized by their molecular weight and charge, smaller/ negatively charged proteins being carried more downstream. Identification of the proteins can be made less complex with the use of a protein ladder, a solution composed of a multitude of proteins of different, yet known, molecular weights, highly utilized to correctly identify the target proteins in relation to their position, the ladder is of great use in the estimation of molecular weight of adjacent bands while recognizing unexpected bands as contaminants.

3.7 Small molecule inhibitors

Along with antibody treatment, small molecule inhibitors are one of the two most commonly used therapies which are capable of efficiently targeting tumours while being able to cross the threshold of the blood brain barrier (BBB)[62]. However, small molecule inhibitors possess certain advantages which makes their usage more favourable over antibodies (which are classified as macromolecules), as small molecules tend to offer higher oral administration, bioavailability, cost of manufacturing, and storage of API[63]. Furthermore the distribution of API is considered to be highly effective, as small molecule inhibitors show improved solubility and tissue distribution[64]. Therefore even with the recent rise of interest for monoclonal antibodies, small molecule inhibitors are still a promising and advantageous therapeutic method, with the choice being dependent what the API requires, for guaranteed distribution of drug at a target site for a short duration, small molecule inhibitors are arguably the most efficient choice, however for prolonged/ delayed release, monoclonal antibodies are more reliable.

2,6-diaminopyridine-3,5(thiocyanate) (PR-619) is one such small molecule inhibitor, a DNA topoisomerase II (TOP2) poison which has high specificity towards DUB inhibitors[65]. The mechanism of action being forming a covalent complex with the TOP2 preventing DNA replication due to blocking transcription causing the strand to degrade, causing apoptosis in the

affected cells[66]. There is evidence of microtubules being affected in the presence of PR-619, causing tau deposits to accumulate at the microtubule-organizing center (MTOC), resulting in microtubule stabilization and destabilization of cellular structure, therefore not allowing the cell to proliferate, this is most evident in nerve cells[67].

PR-619 is most evident as an antitumour agent, in a study regarding urothelial carcinoma (caused by aberrant expression of USP14 and USP21), with 20μM of inhibitor showing notably increased apoptotic activity occurring within the cancerous tissue after 2 days. The cells released caspase-4 and phospho-JNK which are endoplasmic reticulum stress-related apoptotic markers[68], the effect of these markers is causing cell arrest before the cell enters the mitotic stage of the cell cycle[69]. Furthermore, PR-619 shows promising results regarding chondrosarcoma, a tumour developing at the site of supporting tissues surrounding healthy bone. This cancer also involves overexpression of USP14 and USP21, 5μM of PR-619 was effective in inducing apoptosis within 2 days due to inhibiting phospho-ERK and phospho-ARK, crucial markers of their respective pathways[70].

3.8 General concepts and methodologies of fluorescence polarization assays

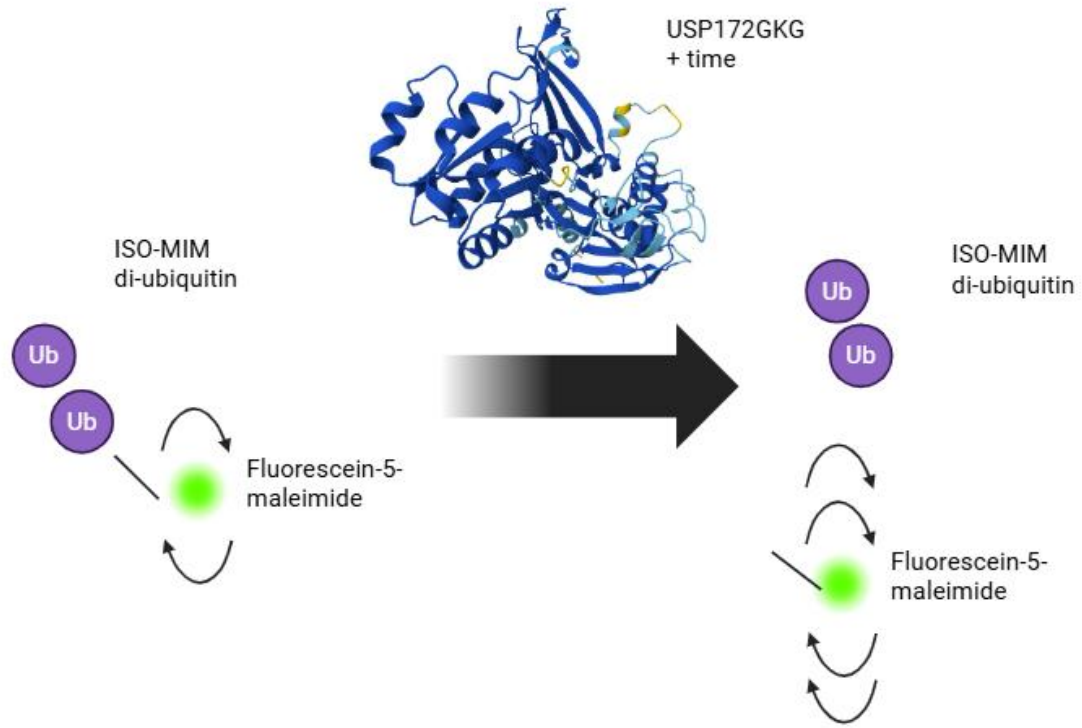


Figure 4- Schematic of the principle of the fluorescence polarization assay, isopeptide bond substrate mimetic (ISO-MIM) di-ubiquitin is linked with a fluorescein-5-maleimide 3 with a thioester bond, the fluorescein maleimide rotates slowly due to being bound to a significantly larger molecule, the emitted light becomes more stable, FP value increases. After being introduced to a deubiquitinating enzyme such as USP17, the thioester bond is severed during the cleaving process freeing the fluorescent moiety, the fluorescein rotates much more rapidly as it is no longer bound to a larger molecule, the emitted light becomes more scattered, FP value decreases.

Assays are analytical techniques which aid in the discovery and analysis of the functionality of a protein of interest- these can range from the measurement of the amount present to interactions between other proteins. In the case of USP17, as a DUB, the determination of enzymatic activity is crucial to understanding the molecular activity between the enzyme and its substrates. Therefore the ideal choice of a high-throughput, quantitative assay allowing the

measure of the cleavage rate of a protein is the fluorescence polarization (FP) assay. Involving a fluorophore bound to a substrate candidate, the fluorophore tumbles slowly when bound to a moiety causing the light to be emitted more consistently in a single direction (polarized). Inversely, an enzyme of interest may be able to cleave the bond between the fluorophore and substrate, the fluorophore tumbles more rapidly due to the liberation, light becomes more scattered (depolarization).

Regarding the science behind FP assays, upon its initial discovery in 1926[71], the technique has undergone numerous changes regarding its usage and overall convenience. Currently the assay functions by using linear plane-polarized light which excites a stationary molecule, causing the molecule to become excited and be able to absorb this light resulting in a 'glow-like' effect, this is known as fluorescence[72]. The value of polarization is inversely proportional to the molecule's rotation and can be calculated using the equation (polarization= intensity of fluorescence parallel to polarized light – (instrument correction factor x intensity of fluorescence perpendicular to polarized light)/ intensity of fluorescence parallel to polarized light + (instrument correction factor x intensity of fluorescence perpendicular to polarized light))[73]. This movement although stimulated by the light, has a random factor of being present in a liquid, which as Brownian motion suggests, fine particles, when in a liquid, move irregularly[74].

A plate reader is equipped with various filters to ensure the correct signal is distributed depending on the fluorophore used, the millipolarization units (mP) are then displayed. Analysis of the data requires the identification of the difference between the mP of the samples (ΔmP) compared to the mP of the probe, with the formula being: ΔmP= -(mP of probe- mP of sample).

3.9 General concepts and methodologies of mutagenesis

By definition, a mutation is a change in the nucleotide sequence on DNA, different variations which often have an overall impact on the final protein structure include insertion (a nucleotide is added to the original sequence), deletion (a nucleotide is removed from the original sequence), or point (a nucleotide is exchanged for a different nucleotide)[75]. Most mutations occur naturally and often spontaneously within an organism's body, or when exposed to exterior factors such as alpha, beta, or gamma radiation. However, mutations can also be accomplished in a controlled environment on isolated DNA, such as polymerase chain reaction (PCR), XL-1 red-cells, and degenerate oligonucleotides-Pfu (DOP)[76].

PCR is a site directed mutagenesis method using paired, double-stranded oligonucleotide primers incorporated with mutation instructions for a specific section of the target DNA[77]. Additional components include template DNA, Thermus Aquaticus (taq) DNA polymerase (used in PCR to withstand the high temperatures in the thermal cycler), and the 4 nucleotides present within DNA[78]. Once the mixture enters the thermal cycler, the DNA is denatured, to separate the double strand allowing the modification to occur, once separated, the temperature is lowered to 50°c-67°c for annealing, allowing the primers to bind to the complementary sequence on the template DNA, once bound the temperature is raised ensuring that the taq polymerase is able to synthesize a single strand of DNA.

3.10 General concepts and methodologies of X-Ray crystallography

Structural determination of a previously unknown molecule is essential to understanding the relation between its molecular structure and its function within the body, this process is especially crucial in protein analysis as the quantity of amino acids can be identified. Hence the

catalytic triad (the three amino acids which are responsible for binding with the substrate) can be determined. The 3D structure can also be discovered using these structural determination techniques, which allows for intramolecular interactions to be identified, this includes tertiary structure and how a protein is able to fold due to the presence of specific amino acids. One such technique is X-ray crystallography (XRC), with XRC allowing to view the protein on its own in a liquid matrix, view the protein along with a known substrate and understand the enzyme-substrate complex, or having multiple proteins in the same matrix and compare the binding differences between the different proteins used[79].

To further enhance the crystallization potential of USP17, the RDFrzS domain (2GKG), a fusion tag, which is incorporated to the loop regions present within USP17[80]. The tag consists of 5 α-helices surrounding a central β-sheet and taking shape of a globular fold[81], with a coiled N-terminal receiver domain and a lagging pole retention C-terminal tail aid in the localization of the tag on the subcellular level. The 2GKG tag possesses qualities of high solubility and crystallizing potential[80], therefore seeing the tag function with great success with USP11 and USP15[80], USP17 was deemed a suitable candidate for the tag due to the domain containing disordered loops, potentially causing a challenge for crystallization.

4. Aims

1. USP2 and USP17, although belonging to the same family of proteins, have different functionalities, and cleave different substrates, this can be determined using gel cleavage assays (for a more qualitative analysis), and FP assays (for a more quantitative analysis). Substrates of interest include the K6-linked di-ubiquitin mimetic, K11-linked di-ubiquitin mimetic, GG-linked di-ubiquitin, and linear di-ubiquitin.
2. The 3D crystalline structure of USP17 is yet to be solved, to increase the odds of crystal formation the 2GKG tag was incorporated into the structure of USP17 and a Morpheus screen (versatile crystal screen), crystals have a chance of formation with subsequent analysis.
3. Inspired by the USP11D1D2D833G mutation (causing an increase in the cleavage rate of linear di-ubiquitin), mutating an aspartate residue to a glycine in the vicinity of the catalytic triad, will theoretically decrease steric hindrance and enhance cleavage of substrates. USP17 has an aspartate residue at the 332 position (only 2 residues away from H334, which is a component of the catalytic triad), would mutating this residue to a glycine enhance the cleavage rate of the previously non-cleavable linear di-ubiquitin? Would the mutation cause the enzyme to not cleave any substrates which can be cleaved by the original enzyme?

5. Materials and Methods

5.1 Constructs used

Construct Name	Molecular weight (kDa)	Residue number	Vector
USP2CD$^{259-605}$ (Addgene)	38.925	605 full length, 346 catalytic domain	pET28a
USP17$^{71-402}$D1D22GKG$^{3-123}$ (insert at A229 and A230)	50.468	530 full length USP17, 331 D1D2, 451 with 2GKG	pET21b
RDFrzS$^{3-123}$	13.667 full length, 13.287 first 3 residues and C-terminus removed,	127 full length, 124 inserted into USP17	*Myxococcus Xanthus*
USP17$^{71-402}$C89S2GKG$^{3-123}$	50.452	451	pET21b
USP17$^{71-402}$D1D2D332GC89S2GKG$^{3-123}$	50.393	451	pET21b
USP17$^{71-402}$D1D2D332G2GKG$^{3-123}$	50.410	451	pET21b
Linear di-ubiquitin	17.111	152	pET26b
K6-linked di-ubiquitin (mimetic) (Azenta)	16.479	148	pET26b
Ubi-GG-Ubi	17.444	154	pET26b
K11-linked di-ubiquitin (mimetic) (Azenta)	16.197	143	pET26b

Table 1- List of all the constructs and protein variations used throughout the methodologies, included are the molecular weights in kilodaltons (kDa), the residue numbers (for both molecular weights and residue numbers, the numbers vary to indicate both the original, full length constructs and the catalytic domain accounting for the truncated non-catalytic regions of the constructs), and the plasmid vector used to express the constructs.

Table 1 shows a comprehensive list of all the constructs to be used throughout this research. Since the entire residue length of USPs are unneeded, and analysis of the catalytic functions of USP2 and USP17 is to be conducted, the residues which do not form the composition of the catalytic domain, were truncated. The 2GKG tag had the first 2 and the last 4 residues truncated and was inserted between A229 and A230 of USP17. Three different versions of USP17 were also used, one with a C89S mutation for inactivation, one with a D332G mutation for research, and one which included both mutations. Non-lysine linked di-ubiquitin constructs did not have any truncation or modifications, K6-linked di-ubiquitin had the first 6 residues on ubiquitin 2 truncated, and K11-linked di-ubiquitin had 11 residues on ubiquitin 2 truncated.

5.2 Buffer and agar plate recipes

All buffers and agar plates composed during the duration of the research used ingredients provided the University of Nottingham medical school.

Buffer A [50mM Tris-HCl, 300mM NaCl, 10% glycerol, 20mM imidazole, pH 7.4]

Buffer B [50mM Tris-HCl, 300mM NaCl, 10% glycerol, 500mM imidazole, pH 7.4]

Gel filtration buffer [50mM Tris-HCl, 300mM NaCl, 1% glycerol, pH 7.2]

Ammonium buffer A: [20mM ammonium acetate, 1L, pH 5.1]

Ammonium buffer B: [500mM ammonium acetate, 1 L, pH 5.1]

Agar plate composition (for competent cells): 15ml LB broth, 15µl chloramphenicol.

Agar plate composition (for pET21b): 15ml LB broth, 15µl ampicillin.

Agar plate composition for control (pET21b): 15ml LB broth, 15µl kanamycin.

5.3 Expression of USP2CD, USP17D1D2-2GKG, and mutagenesis of USP17D1D2-D332G2GKG

Protocol for overnight expression: A single colony was chosen from the plate and aliquoted to 2 tubes with 5ml LB broth, which was then placed in an incubator at 37°c overnight.

Mini prep protocol: Following the overnight expression protocol, the 2 tubes were centrifuged for 10 minutes at 2500rpm, the supernatant was discarded, 200µl cold resuspension solution (with added RNase), 200µl of lysis solution, and 350µl of neutralizing solution was aliquoted and mixed via inversion, the tubes were then centrifuged at 13300rpm (top speed) for 10 minutes. After discarding the flow through, the lysate was transferred to a binding column and spun at top speed for 1 minute (flow-through being discarded), 100ml of ethanol was added to optional wash solution- with 500µl of the solution being aliquoted to both binding columns and spun at top speed for 1 minute (flow-through being discarded). 750µl wash solution was aliquoted to the binding columns, spun at top speed for 1 minute (flow-through being discarded), the spinning process was repeated once more, with the flow-through also being discarded. 100µl elution solution was aliquoted to the binding columns and placed in 1.5ml tubes, and spun for 1 minute at top speed, with the flow-through being stored at -20°c.

Overexpression following mini prep: 5µl of mini prep solution was aliquoted to a tube, agarose was heated using a microwave until fully dissolved, once in the mould, 2µl of SDS gel dye was added with the mixture being homogenized. Loading dye consisted of 500µl distilled water, 500µl glycerol, pipette tip of EDTA crystals, pipette tip worth of bromophenol blue, 2µl of the dye was aliquoted to the mini prep samples, 7µl of the samples were aliquoted to a 15% SDS-PAGE gel, with the gel being run at 80V for 20 minutes. The transformation and overnight protocols were repeated using BL21CP E. coli cells and chloramphenicol being aliquoted to the overnight culture, the remaining parameters and conditions were identical for the overnight preparation.

Mutagenesis protocol: the primer was reconstituted using the appropriate volume of ultrapure water. In a 600µl tube, the following solutions were aliquoted: 10µl of appropriate forward primer, 10µl of appropriate reverse primer, 5µl of reaction buffer, 1µl dNTP, 1µl 50% DMSO, 1µl Phusion DNA polymerase, a calculated volume of template double stranded DNA which equalled to 10ng, an appropriate volume of distilled water made up the final mixture to 50µl. The final volume was split equally between 3 smaller tubes, and inserted into the PCR machine, with the parameters being: initial denaturation at 95°c for 5 minutes, denaturation at 95°c for 30 seconds, annealing at 62°c for 1 minute, extension at 72°c for 3 minutes, with the process running at 40 cycles, and the final extension at 72°c for 10 minutes. 1µl of DpnI enzyme was aliquoted to the tubes after the PCR process, with the mixtures being incubated at 37°c overnight, 4µl of each mixture were then transformed into Novablue cells and followed the plasmid transformation protocol.

DNA sequencing protocol: (protocol was found on the Source Bioscience website[82]) initially samples were prepared to enhance purity and concentration. The T7 promoter and terminator were annealed to the delivered DNA template, chain-terminating nucleotides were incorporated

to produce DNA fragments of various sizes, these fragments were then separated by size, chromatograms were produced of the fragments and sent via e-mail.

Protocol for transformation: insert 1µl of pET21b to (vol) of novablue cells on ice for 10 mins, heat shock the mixture at 42°c for 45 secs, back in ice for 2 mins, 500µl LB broth, shake tube at 37°c for 45 mins, on an ampicillin agar plate 200µl of mixture is inserted, plate is left at 37°c for 18 hours, colonies stored in the fridge at 2°c.

Harvesting cells protocol (USP17): 28ml of overnight culture is added to 800ml LB broth, 800µl ampicillin is additionally inserted, left at 37°c on 150rpm for 2 hours, absorbance (600) is read, if above 0.5, 300µl IPTG is added, culture is placed back in the incubator overnight, temperature lowered to 20°c. Mixture is divided into 4 centrifuge bottles, centrifuged for 30 minutes at 4500rpm at 9°c, the remainder of the mixture is placed into tubes and centrifuged at 6000rpm for 20 minutes at 9°c. The resulting pellet was then scraped using a 25ml pipette, the solution used for the scraping was a mixture of ~20ml supernatant and 30ml buffer A, once all pellet was scraped, the resulting supernatant was centrifuged at 6000rpm for 20 minutes, the supernatant was then discarded, and the resulting pellet was stored at -20°c.

5.4 Purification of USP2CD, USP17D1D2-2GKG, and USP17D1D2-D332G2GKG

Protein purification protocol: 20ml of buffer A was aliquoted to the USP pellets and homogenized using a vortex, the mixture was transferred to a beaker on ice and sonicated for 8 minutes at 10 second intervals with 5 second breaks or until the pellet is completely homogenized. The mixture is then divided between 2 tubes and centrifuged at 12000 rpm for 50 minutes, meanwhile the protein purifier and peristaltic pump with the Histrap excel column were equilibrated using buffer A and B. The supernatant was filtered using a 0.45µM filter, with the pellets and a remainder of the lysate which remained in the syringe being stored in the fridge at 3°c. The lysate was then incorporated into the column using the peristaltic pump at flow rate of 2ml/min, with the flow-through being stored in the fridge, the column was then attached to the protein purifier with the Histrap column protocol- the fractions which showed a significant peak were collected and 1µl of DTT was added to each before being stored in the fridge. The gel samples were prepared in tubes according to the composition: 10µl of lysate, 10µl of flow-through, tip load of both pellets, and 10µl of 5 fractions of interest, 5µl of SDS with 0.01µM of DTT being added to each tube. The samples were heat shocked at 98°c for 10 minutes, 5µl of samples were aliquoted to run on an SDS-PAGE gel for 50 minutes at 200V. In the meantime, an additional tube of 10µl of the protein ladder was prepared, the entirety was then aliquoted on the leftmost well of the gel to be used for comparison against potential bands. 5 samples which showed the most intense bands at the correct MW to the protein of choice were aliquoted to a 10kDa concentrator tube and centrifuged at 6500rpm until the volume reached below 1ml. 50µl samples were aliquoted into smaller tubes and flash frozen using liquid nitrogen and stored in a -80°c freezer.

Gel filtration protocol: this method was achieved using an Äkta Start protein purifier: both pumps were inserted into dH2O at a flow rate of 5ml/min for 2 minutes, with pump B being cleaned first. Pump A was inserted into gel filtration buffer at a flow rate of 5ml/min for 2 minutes, the Excel histrap column was then cleaned at 0.8ml/min for 2 hours. The column was then cleaned with the water using pump B using the same flow rate for 10 minutes. Pump A was inserted into the ammonium buffer A, pump B was inserted into the ammonium buffer B, 10ml

of buffer B followed by 15ml of buffer A were used to equilibrate the column at flow rate 0.8ml/min. 5 fractions acquired from protein purification were spun down in a 10kDa centrifugal filter until the volume was 2ml, the solution was then transferred to 2 tubes on ice which were centrifuged for 10 minutes at 13300rpm. Meanwhile, 20ml dH2O was inserted into the sample loop using a syringe and needle with the switch being set to load sample- 20ml of gel filtration buffer was inserted into the sample loop with the needle being left in the loop. The centrifuged fractions were inserted into a 5ml syringe with the solution being slowly inserted into the loop while the switch is set to load sample.

5.5 Gel cleavage assays using USP2CD, USP17D1D2-2GKG, and USP17D1D2-D332G2GKG

Gel cleavage assay protocol: UbK6Ub (K6) was diluted to 10μM using gel cleavage buffer the gel cleavage assay used the following tube compositions: 2x 10μl 10μM K6 mimetic + 10μl gel cleavage buffer, and 4x 10μl 2μM USP2 + 10μl 10μM K6 mimetic, 10μl of SDS with 0.001mM DTT was aliquoted immediately to one tube of K6 and one tube of K6 mimetic + USP2. The SDS/DTT mixture was added to each USP2 tube at 1 hour, 2 hours, and overnight accordingly, the samples were heat shocked for 10 minutes at 98°c, and 10μl were aliquoted to each well on the 18% SDS-PAGE gel and run for 50 minutes at 200V. The gel was then heated in the microwave for 1 minute while submersed in water, 1 minute submersed in SDS stain, and 1 minute in water.

Gel quantification protocol: imageJ was used to calculate the intensity of the bands presented by the gels. The colours of the gel were inverted using the edit tab, the smallest band was identified, and with the rectangle function activated, the entirety of the band was selected (ensuring that the rectangle does not involve the background gel. To measure a blank, after the rectangle was created, the rectangle was moved to only include the background of the gel, ctrl+M was used to measure the parameters, the integrated density was noted. The rectangle was moved back to its original position and measured; the integrated density of the band was subtracted from the density of the blank. This process was repeated for each well with the exception of the protein ladder, the same rectangle was used for each band and a new blank was measured for each well.

5.6 Fluorescence polarization and competition assays using PR-619/K6 mimetic

Fluorescence polarization assay protocol: the K6-linked di-ubiquitin was diluted by a factor of 10, starting with 1μM decreasing down to 1nM, PBS being used as a dilutant. The probe, which consisted of linear di-ubiquitin linked with fluorescein-5-maleimide was diluted using PBS to 10nM. To denature the USP2, the tube was placed in a 98°c water bath for 30 minutes. A total volume of 50μl was inserted into each well in triplicates: 50μl PBS, 30μl PBS + 20μl 1μM denatured USP2, 30μl PBS + 10μl denatured USP2 + 10μl 1μM K6, 30μl PBS + 10μl USP2 + 1μM K6 (this composition was repeated for every K6 concentration). Before the substrate was added, the microplate reader was initialized using the following parameters: wavelength 494nM, the readings being accomplished every 30 seconds at a constant temperature of 25°c.

Competition fluorescence polarization assay protocol: the assay used a 384 well microplate from Corning™, the well composition used was as followed: wells 1-3 were composed of 35μl PBS + 15μl 350nM USP17D1D2-2GKG, wells 4-6 had an identical volume of PBS with the remaining volume being composed of 5μM PR-619. In triplicates, the following wells were composed of 20μl PBS + 15μl 350nM USP17D1D2-2GKG + a varying concentration of 15μl PR-

619/K6 mimetic, the concentrations used were 5μM, 4μM, 3μM, 2μM, 1.5μM, 1μM, 0.5μM, which when added, was immediately inserted into the microplate reader.

IC50 calculation assay protocol: once the data was transferred from the microplate reader, the data was analysed in an Excel spreadsheet, the rate of change was calculated by subtracting the FP value at 0 minutes from the FP value at 15 minutes for each sample. This data was organised into a table (X-axis- concentration of PR-619, Y-axis- rate of change) , any singular anomalies were removed.

5.7 Structural determination of USP17-2GKG

Crystal tray protocol: following the gel filtration protocol, the 5 fractions which showed the highest peaks were collected in a beaker, the combined fractions being placed in a 10μM fractionation column and centrifuged for at 5000rpm until the volume reached 1-2ml. 5ml gel filtration buffer was aliquoted to each well on a Morpheus crystal screen, while 10μl of USP17D1D2-2GKG fraction mixture was aliquoted to the smaller, protruding wells on the screen. Once aliquoting was accomplished, foil was cut using a scalpel and placed on the crystal screen, the screen was then put in an incubator set at 20°c for 5 weeks.

Alphafold prediction protocol: For USP17D1D2-2GKG, and all other USP17 constructs, the amino acid sequence used was found in protein databank, any truncations, residue modifications, and protein inserts were accounted for and changed in the Alphafold server once copied. After the job has concluded loading, the zip-file was saved, the. model file was saved onto ‘this PC’ for easier access to SFC Chimera editing (any .model file would have sufficed, for consistency purposes,. model_0 was chosen for all predicted structures). The protein structure was then edited in Chimera to remove any ions/ particles not contributing to the raw tertiary structure, and colour coded to highlight key areas of interest such as singular ubiquitin moieties forming a di-ubiquitin and altered residues.

6. Results

6.1 Expression and purification of USP2CD catalytic domain by affinity chromatography.

Initially, the proteins required for analysis, including the substrates, underwent plasmid transformation to Novablue E. coli cells, followed by overnight growth and isolation of DNA through the usage of a plasmid DNA miniprep kit from Thermo Scientific, the resulting DNA was incorporated into Novablue E. coli cells, with a pellet being formed which was sonicated and inserted into a Histrap Excel column- the column was then plugged into an Äkta start 2 protein purifier. All purifications and their corresponding chromatograms were all accomplished using the same protein purifier and the same column. Compared to a standard Histrap column, EDTA does not to be aliquoted to strip the nickel, instead buffer B and ethanol was added to remove the excess protein and clear the column from contaminants. All SDS-PAGE gels was conducted using the vertical electrophoresis tank BK-VET01 by BiobaseTM and the heating blocks used to denature the protein samples was accomplished using the dry block heater- digital by GrantTM.

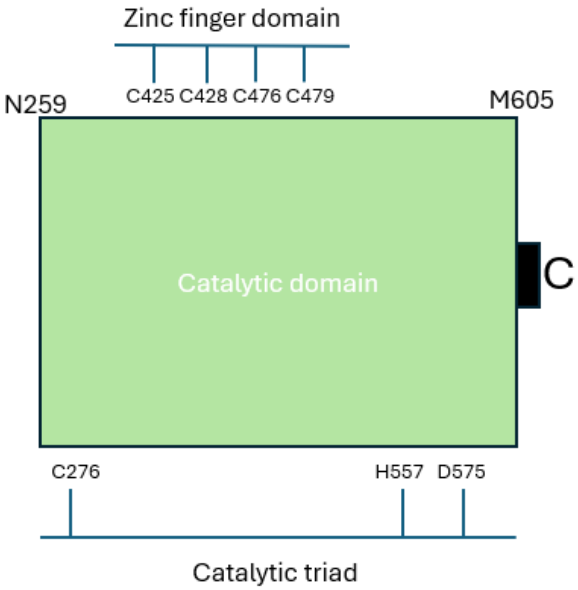


Figure 5- Schematic of the catalytic domain of USP2 used in the research, with the boundaries, catalytic triad, and zinc finger domain indicated with the residues composing of the subdomains

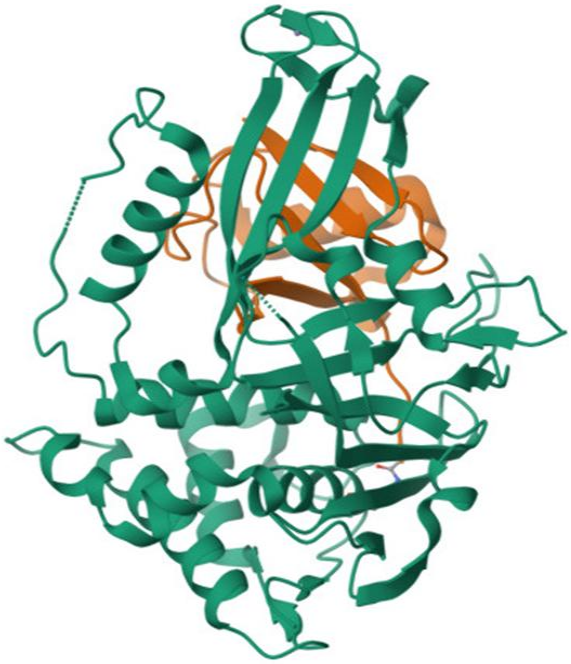


Figure 6- the catalytic domain of USP2 was taken from PDB (code: 2IBI), green being representative of the catalytic domain, ubiquitin being highlighted in orange.

The catalytic domain of USP2 was chosen, since the protein was previously researched to cleave various substrates, including K6/K11-linked and linear di-ubiquitin, therefore this enzyme was used as a control for the USP17 constructs in regard to purification, gel cleavage and fluorescence polarization assays.

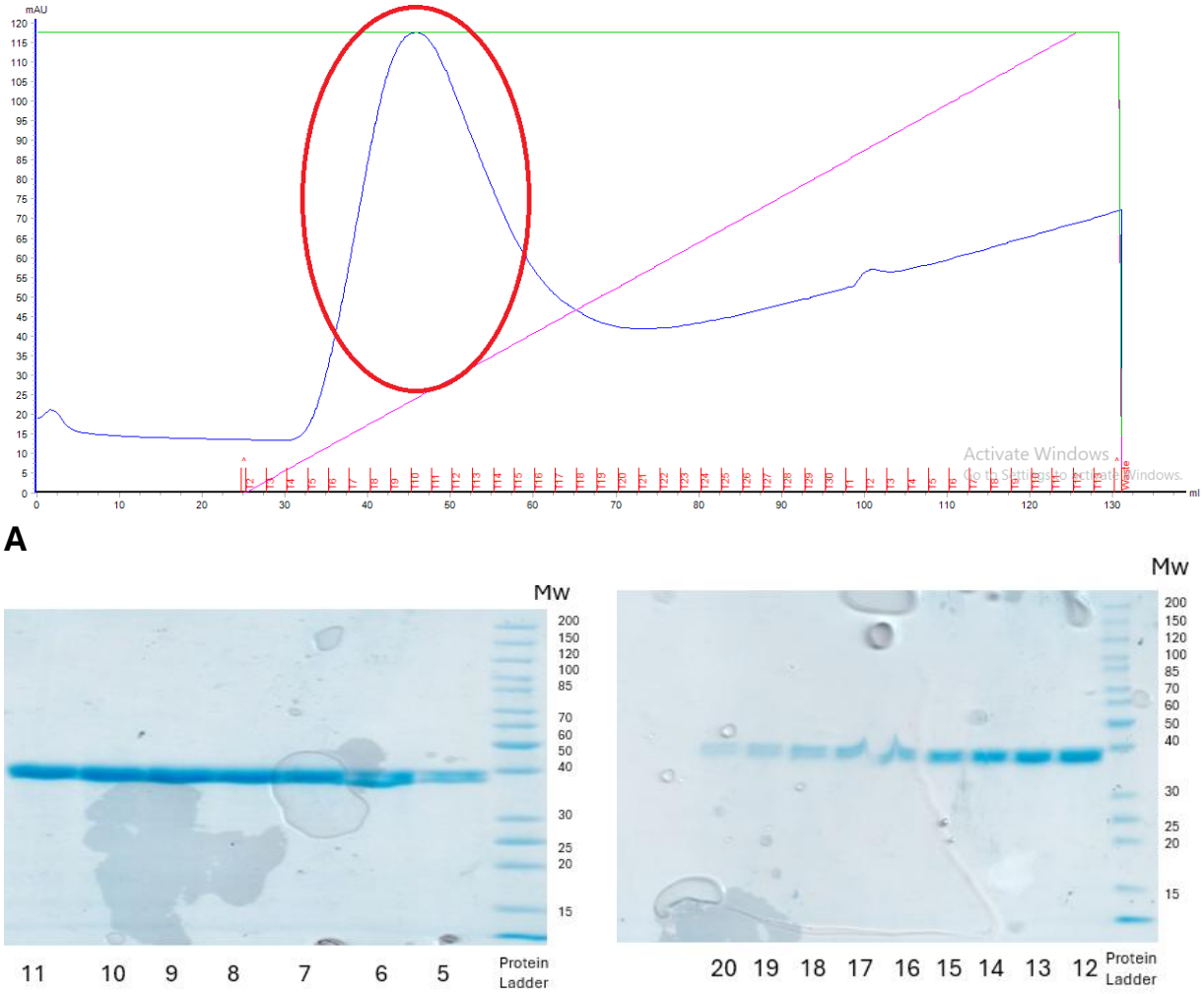


Figure 7- A) Chromatogram for the nickel purification of USP2 D1D2 using the Histrap column and being run on the Akta Start 2 protein purifier, with a linear increase of imidazole from 20mM to 500mM throughout the duration, fractions of interest, which were subsequently analysed using SDS-PAGE, included fractions 5-20, as the peak at 1.5 minutes was deemed most significant. B) Image of SDS-PAGE stained with Coomassie blue, the rightmost lane is occupied by the protein ladder, which was the 10-200kDa wide range protein molecular weight marker supplied by Thermo Scientific, with the correct molecular weight (kDa) being present on the right of the image. The other lanes show the fractions which were obtained from the purification, and which were present on the most significant peak, according to the chromatogram.

Protein purification of USP2 went according to the hypothesis, with a substantial peak being present within 10 collected fractions (fig. 7), fractions 5-20 were analysed using a 15% SDS-PAGE gel (fig. 7), with a consistent band present at ~39KDa, which is slightly short of the molecular weight of USP2-41 (41 KDa), no visible contaminations were observed as there was no significant bands at different molecular weights. Bands 8-13 had a more intense shade compared to the remaining bands; therefore these were selected for concentrating on a 10KDa Cytiva protein concentrator and buffer exchange using gel filtration buffer, the final concentration was measured to be 68.2μM with the aid of an IMPLEN NanoPhotometer N50-touch UV/Vis Spectrophotometer.

Mass= concentration x volume x molecular weight

Mass= 68.2μM x 1ml x 41000

Mass= 68.2 x 1 x 41000

Mass= 2796.2µg

Mass= 2.7962mg

6.2 Structural predictions of the K6 and K11 synthetic mimetics.

Due to an isopeptide bond being difficult to engineer for a protein in large volumes, the K6/K11-linked di-ubiquitins used throughout the research were mimetics, the purpose of the truncated residues was to mimic the orientation of the bond which would have occurred in the in-vivo versions of the proteins. Predicted structures of the proteins were generated to visualize the differences.

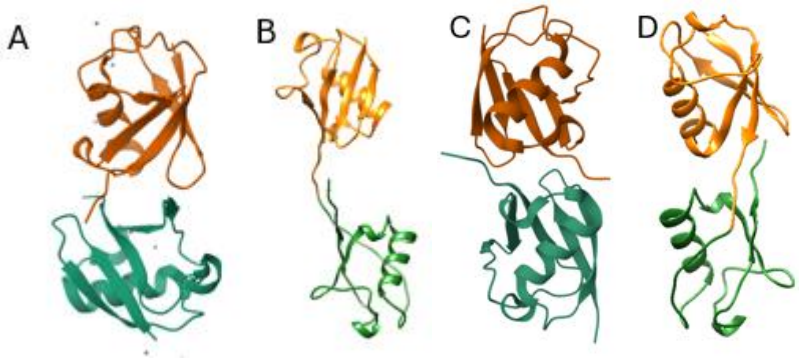


Figure 8- A) Official crystalline structure of K6-linked di-ubiquitin (PDB:2XK5). B) Predicted structure of the K6-linked di-ubiquitin mimetic generated in Alphafold, first 5 amino acids of ubiquitin 2 (green) truncated. C) Official crystalline structure of K11-linked di-ubiquitin (PDB: 3NOB) D) Predicted structure of the K11-linked di-ubiquitin mimetic generated in Alphafold, first 11 amino acids of ubiquitin 2 (green) truncated.

6.3 Gel cleavage assay using USP2CD to cleave a K6-linked di-ubiquitin mimetic.

All gel cleavage assay results were accomplished using 15% acrylamide gels mainly to allow increased visibility of the ubiquitin bands present at ~9kDa, any repeats were accomplished in the same day to eliminate any possibility of different environmental conditions such as temperature and airflow. Following the conclusion of each gel cleavage assay result, the colour of the images was inverted using ImageJ, with a sample area encompassing an unoccupied section of the gel and used as a blank against the bands, the exact sample area was used for all bands, the results of the densitometric data were subtracted from the blank and plotted into a scatter graph, showing the regression or lack thereof of the density. Initially the purified USP2CD was tested against a known substrate, 203µM K6-linked di-ubiquitin mimetic diluted to 20µM, provided by Krzysztof Raszpla, a stopwatch was kept as to maximise the accuracy for the reaction termination using 10µl 5x SDS loading dye.

A

B

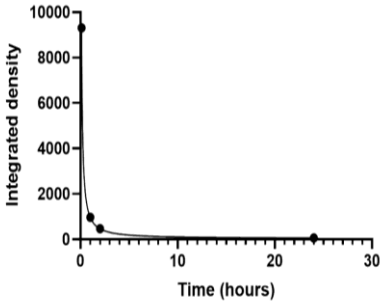


Figure 9- A) Image of SDS-PAGE stained with Coomassie blue, the leftmost lane is occupied by the protein ladder, which was the 10-200kDa wide range protein molecular weight marker supplied by Thermo Scientific, lane 2 contains 10µl 10µM of K6-linked di-ubiquitin and 10µl 5x SDS-PAGE loading dye, lane 3 contains an identical sample as lane 2 with the loading dye being added overnight, lane 4 contains 10µl 20µM K6 mimetic+ 10µl 3µM USP2CD + 10µl loading dye with the buffer being added immediately after introducing the enzyme, lanes 5-7 contain an identical sample as lane 4, however the dye was added at regular intervals being 1 hour, 2 hours, and 24 hours. Notable bands include one present at ~14kDa and ~39kDa. B) Densitometric data of the gel cleavage assay of 1.5µM USP2CD with 10µM K6-linked di-ubiquitin showing the results of the measured integrated density of the band relating to the K6 mimetic on the SDS-PAGE, the values being taken by ImageJ.

The gel cleavage assay (fig.9 A) shows USP2CD present at the same molecular weight as demonstrated in the expression and purification (fig. 6) of the deubiquitinase, with the bands retaining the same intensity, from the initial aliquoting of enzyme to overnight. Lanes 2 and 3 (fig.9 A) present K6 mimetic also keeping the same density from 0 minutes alone to 0 minutes upon being exposed to USP2CD, and overnight alone, however after an hour upon being introduced to USP2CD, the density of the band decreased by 9.62x, with the density decreasing further with time, additionally a band appears at the bottom, at the conclusion of development of the gel, since the band remains consistent with the dissipation of K6, this represents the product (ubiquitin) being formed.

The number of minutes 1.5µM USP2CD takes to hydrolyse 5µM of K6 mimetic (half of the total), can be presented with the densitometric graph (fig. 9 B). The EC50 was measured to be 0.04553 hours/ 2.7318 minutes for USP2CD to decrease the mass of K6 mimetic by a half.

6.4 Gel cleavage assay using USP2CD to cleave linear di-ubiquitin.

As linear di-ubiquitin is used throughout this project to test if other enzymatic proteins are able to cleave the peptide bond between the 2 ubiquitin molecules, in addition to the fact that USP2CD is yet again, known to cleave the complex, another gel cleavage assay was conducted to present how rapidly mono-ubiquitin is able to be produced, and where specifically on the gel ubiquitin will appear compared to the linear di-ubiquitin band. Linear di-ubiquitin was provided by Jiatong Zhang with a concentration of 231.6µM, diluted to 20µM.

A

B

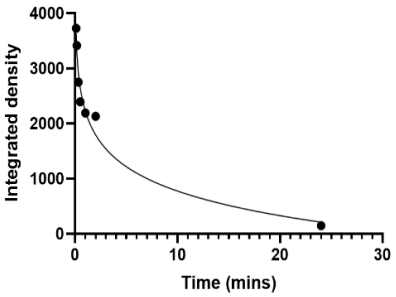


Figure 10- A) Image of an SDS-PAGE stained with Coomassie blue, the leftmost lane is occupied by the protein ladder, being the #26610 unstained protein marker by Thermo scientific, lane 2 was occupied by 10µl of 10µM of linear di-ubiquitin and 10µl 5x SDS loading dye, lanes 3-9 contained 1.5µM USP2CD + 10µM linear di-ubiquitin + 10µl 5x SDS loading dye. B) Densitometric data of the gel cleavage assay of 1.5µM of USP2CD and 10µM of linear di-ubiquitin with the measured integrated density taken from the band relating to linear di-ubiquitin present on the gel using an equal sample area, the readings were accomplished using ImageJ.

The gel cleavage assay (fig. 10 A) shows linear di-ubiquitin present at 17kDa, with its documented molecular weight being identical, and the USP2CD being present at 41kDa, this information being consistent throughout the gel despite different timepoints. Bands present at ~9kDa begin to appear after 10 minutes and intensifying over time, these bands are in relation to the product of the reaction, being ubiquitin.

The density of the di-ubiquitin decreases upon USP2CD being introduced, however although the concentrations of the enzyme and the substrate remained identical to the K6 mimetic gel cleavage assay (fig.9), the density decreases at a significantly slower rate, with the density reduction of linear di-ubiquitin from 0 minutes to 1 hour being 1.7x, the difference is even more visible when observing the estimated time to reduce the amount of substrate by half. The EC50 was calculated to be 5.167 hours, therefore upon comparing the EC50 of K6 mimetic (0.04553), USP2CD cleaves linear di-ubiquitin 113.5x slower than K6.

6.5 Fluorescence polarization assay using USP2CD to cleave a di-ubiquitin probe.

An initial fluorescence polarization assay was performed with the recently purified USP2CD protein against an isopeptide bond substrate mimetic (ISO-MIM) di-ubiquitin marked with fluorescein-5-maleimide, the procedure ran for a course of an hour, the FP readings being measured every 30 seconds to observe the change of the di-ubiquitin being cleaved at different speeds depending on the varying concentrations of USP2CD.

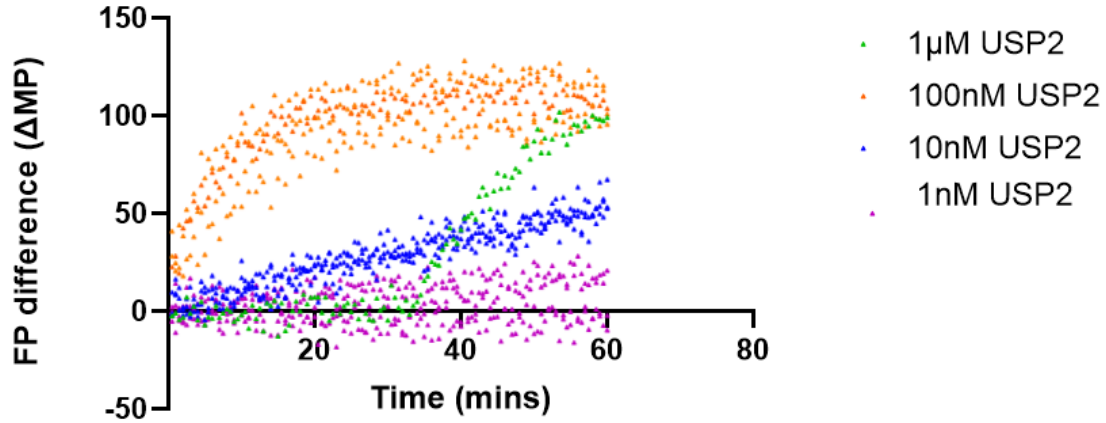


Figure 11- Scatter graph of the fluorescence polarization assay of the USP2CD and the di-ubiquitin fluorescent probe, the measurements were conducted using the EnVision Xcite multimode plate reader and conducted over a period of one hour, with readings being taken every 30 seconds, and a colour-coded figure legend showing the concentrations used during the assay.

The fluorescence polarization assay (fig. 11) had the FP difference values calculated using the formula:

FP difference= - (FP (sample)-FP (probe))

The assay was conducted using a tenfold 3.3μM USP2CD dilution, followed by the insertion of the 33nM di-ubiquitin probe and immediately inserted into the reader, with the graph displaying that although the highest concentration of enzyme does increase after 30 minutes, 100nM of USP2CD presents with a curved shape from the beginning of the assay and plateauing at approximately 100 ΔMP, with 10nM and 1nM increasing at a slower, more linear rate than the higher concentrations. The assay was conducted with triplicate repeats of each enzyme concentration, however as air bubbles were present within the wells of 1μM of USP2CD, with the two repeats showing random FP values, therefore only the repeat with no air bubbles was analysed.

6.6 Expression and purification of USP17D1D2-2GKG catalytic domain by affinity chromatography

Since the crystalline structure of USP17 has not been fully solved, with research currently ongoing, the following structures were predicted using the Alphafold™ server, with unclear regions being removed, with a prediction also being tested with a molecule of K6-linked di-ubiquitin as to test a possible site for the binding on the structure.

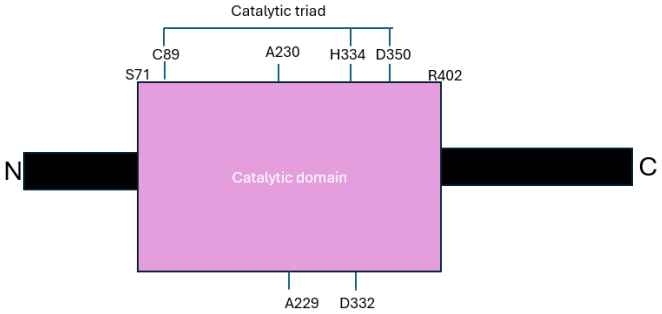


Figure 12- schematic depicting the domain structure of the full length USP17, with the domain boundaries and the catalytic triad being highlighted. Other notable residues include: A229 and A230- where the 2GKG tag was inserted, and D332 as the residue would be mutated.

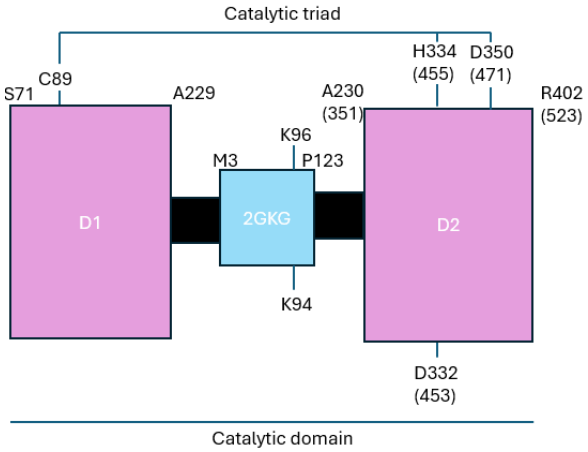


Figure 13- schematic depicting the USP17D1D22GKG domain structure, notable changes include the insertion of the 2GKG tag between A229 and A230, with the boundaries of the amino acid structure of 2GKG also being included. K94 and K96 were highlighted as the 2 residues were mutated to an alanine and serine respectively. Numbers in brackets represent the residue number, when the entire construct is considered as a single moiety.

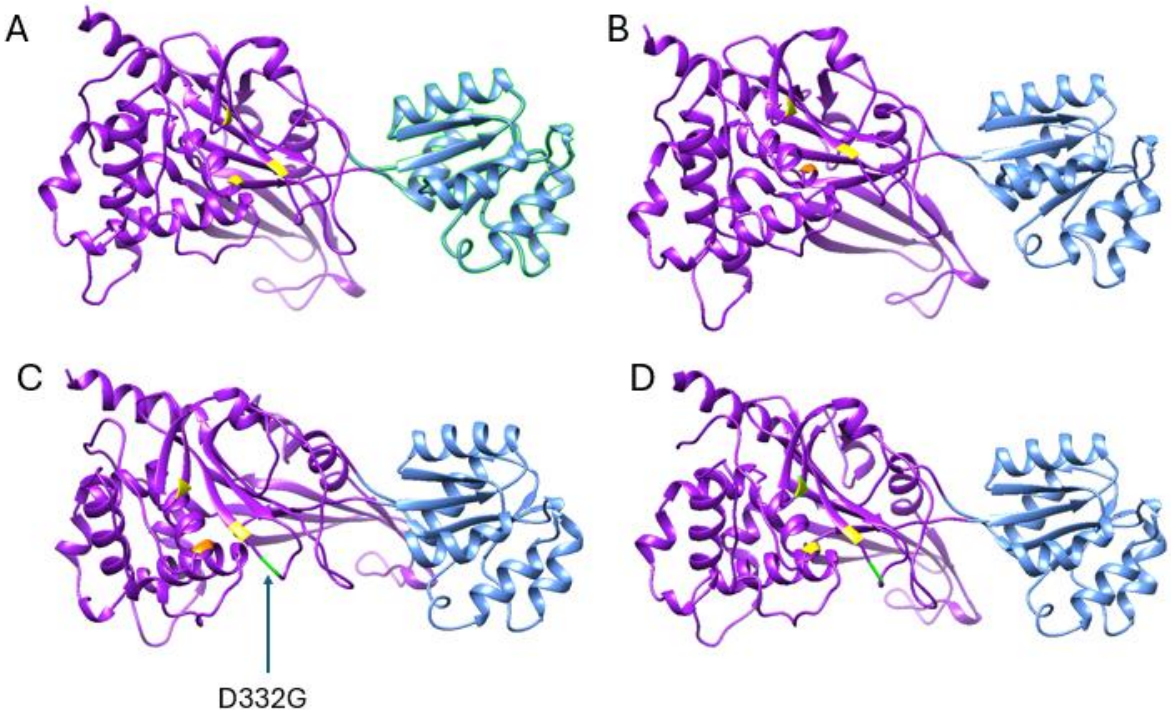


Figure 14- 4 USP17 constructs were used for this research, with the predicted structures being generated in Alphafold and edited in UCSF chimera. A) USP17D1D2-2GKG, B) USP17D1D2-C89S2GKG, C) USP17D1D2C89SD332G2GKG, D) USP17D1D2-D332G2GKG.Colour legend: Purple- catalytic domain of USP17 Blue- 2GKG, yellow- catalytic triad, orange- C89S, green- D332G.

Similarly to USP17D1D2-2GKG, USP17D1D2-D332G2GKG is an even more obscure molecule, as the aspartate to glycine mutation was so far documented only on the studies with USP11, attempting to enhance the peptide binding between linear di-ubiquitin, therefore this model was predicted using the Alphafold™ server, with the uncertain regions being removed.

The expression process for USP17D1D2-2GKG was identical to USP2CD, with the same E. coli strains, and the same antibiotics being used during overnight expression, the bacteria underwent transformation, and mini prep enhancement, and before any expression could begin, the protein was tested for the correct amino acid sequence, which resulted in a match, the expressed protein used an identical histrap excel column for the purification process.

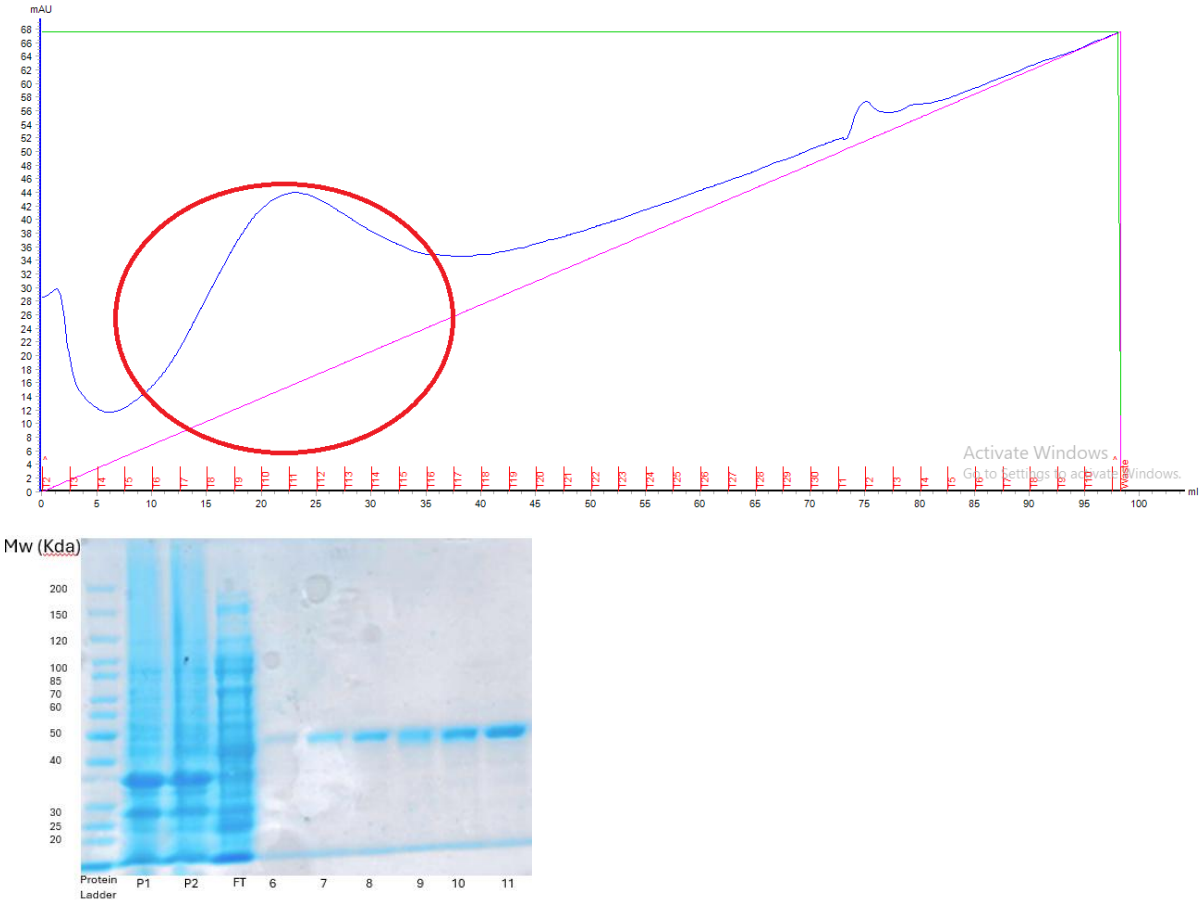


Figure 15- A) Chromatogram of the purification of USP17D1D2-2GKG using the histrap column on the Akta Start 2 protein purifier, with a linear increase of imidazole from 20mM to 500mM, fractions of interest included 6-11, as these were deemed to be related to the most significant peak, these were subsequently analysed on SDS-PAGE. B) Image of SDS-PAGE stained with Coomassie blue, the leftmost lane is occupied by the protein ladder, which was the 10-200kDa wide range protein molecular weight marker supplied by Thermo Scientific, with the correct KDa being present on the left of the image, lanes 2 and 3 (from left to right) are occupied by the pellets which were centrifuged at 12000rpm for 45 minutes post-sonication, lane 4 is occupied by the flow-through collected from the supernatant being introduced to the histrap column, lanes 5-10 were occupied by the corresponding fractions, which were related to the significant peak present on the chromatogram.

The protein purification of USP17 (fig.15) went similarly to USP2, with a significant peak being present at an identical fraction, with the most striking difference being that the peak of USP2 reached a maximum absorbance of 118mAU, whereas the peak of USP17 reached ~43mAU. The molecular weight of USP17D1D2-2GKG was measured to be 51.4634KDa, which aligned with the mW of the bands (fig.16), the band intensity on the resulting gel (fig.16), are comparable to the bands present on fig. 10, with the intensity increasing with each subsequent fraction, in relation to the absorbance of the peak (fig.16). All fractions present on fig.16 were aliquoted into two dialysis columns followed by overnight dialysis, overall concentration measured at 35.80µM, and overall volume measured at ~950µl therefore the entirety of the remaining solution was aliquoted to 19 tubes and flash frozen in liquid nitrogen. To calculate the overall yield, concentration= mass/volume was used.

Mass= concentration x volume x molecular weight

Mass= 35.8µM x 950µl x 51.4634KDa

Mass= 35.8 x 0.95 x 514634

Mass= 1750.3µg

Mass= 1.7503mg

6.7 Gel cleavage assay using USP17D1D2-2GKG to cleave a K6-linked di-ubiquitin mimetic

Upon retrieving USP17D1D2-2GKG via expression and purification, initially, the protein was tested with a K6-linked di-ubiquitin mimetic, despite literature stating an inefficient level of cleavage being performed by USP17D1D2-2GKG against K6, since the mimetic has potentially different properties than the substrate in vivo, the results could be different from the described outcome.

A B

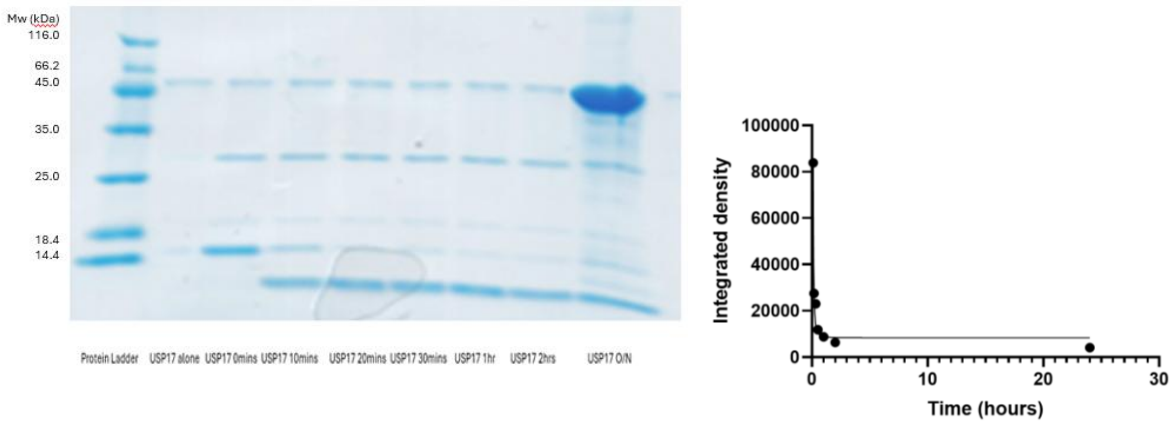


Figure 16- (A) Image of SDS-PAGE stained with Coomassie blue, the leftmost is occupied by the protein ladder, which was the #26610 unstained protein marker from Thermo Scientific, lane 2 was occupied by 10µl 1.5µM USP17D1D2-2GKG and 10µl SDS-PAGE loading dye, lane 3 contains 10µl 3.0µM USP17D1D2-2GKG + 10µl 20µM K6-linked di-ubiquitin +10µl SDS-PAGE loading dye, lanes 4-9 are composed of an identical sample as lane 3 with the SDS-PAGE loading dye being added at different timepoints (10 mins, 20 mins, 30 mins, 1 hour, 2 hours, and 24 hours). (B) Densitometric data of the gel cleavage assay of 1.5µM USP17D1D2-2GKG and 10µl K6 mimetic di-ubiquitin, with values of integrated density taken from the K6 mimetic band present on the SDS-PAGE gel using a sample size of equal area, the readings being accomplished using ImageJ.

The gel cleavage assay shows the presence of K6-linked di-ubiquitin mimetic present at 17kDa, the K6 mimetic sample also contains bands present at 19 and 30kDa, which in stark contrast, do not dissipate over time. Additionally USP17D1D2-2GKG is symbolised by bands present at 51kDa, which stay consistent with the exception of the overnight sample, presenting the bands in a more intense, saturated colour. The final band to note is of ubiquitin which appears at ~9kDa, intensifying in colour over the course of the assay.

The overall density of the K6 mimetic band depletes over a period of time, with the cleavage rate surpassing USP2CD, additionally upon analysing the densitometric data, the EC50 was estimated to be 0.002289 hours/ 0.13734 minutes/ 8.2404 seconds, 19.9x faster than USP2CD.

The regression line was found to plateau approximately at 8500 integrated density, whereas the lines in the previous graphs were found to reach or were reaching the x-axis.

6.8 Gel cleavage assay using USP17D1D2-2GKG to cleave a K11-linked di-ubiquitin mimetic.

Unlike K6, K11-linked di-ubiquitin is an active substrate of USP17, provided by Krzysztof Raszpla, a direct comparison in the speed of cleavage was desired to be observed, therefore an identical concentration of both USP17D1D2-2GKG and a K11-linked di-ubiquitin mimetic was used in a separate gel cleavage assay, with K11 having a concentration of 116µM.

A **B**

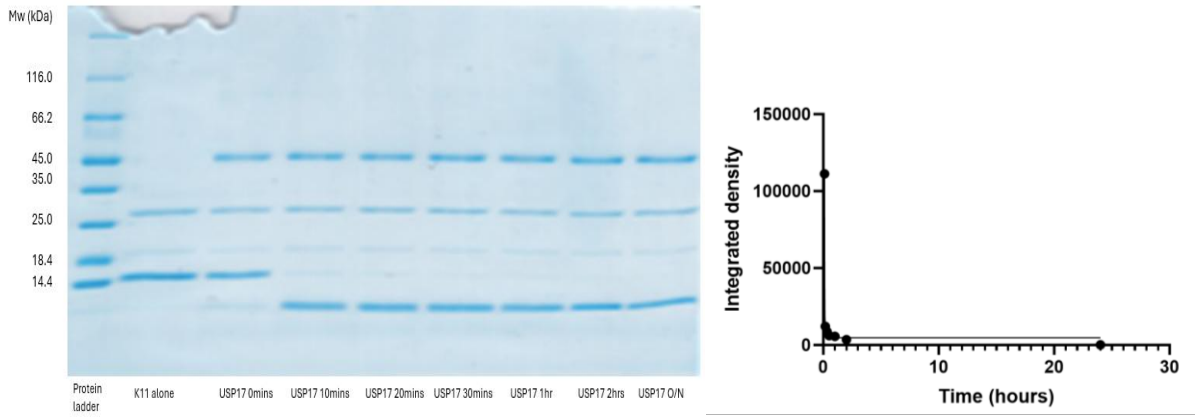


Figure 17-A) Image of SDS-PAGE stained with Coomassie blue, the leftmost is occupied by the protein ladder, which was the #26610 unstained protein marker from Thermo Scientific, lane 2 was occupied by 10µl 10µM K11 mimetic and 10µl SDS-PAGE loading dye, lane 3 contains 10µl 3.0µM USP17D1D2-2GKG + 10µl 20µM K11 mimetic +10µl SDS-PAGE loading dye, lanes 4-9 are composed of an identical sample as lane 3 with the SDS-PAGE loading dye being added at different timepoints (10 mins, 20 mins, 30 mins, 1 hour, 2 hours, and 24 hours). B) Densitometric data of the gel cleavage assay of 1.5µM of USP17D1D2-2GKG and 10µM of a K11-linked di-ubiquitin mimetic, with the values indicating the integrated density of the bands relating to the K11 mimetic on the SDS-PAGE gel using a sample size of equal area for each band, the density values were calculated using ImageJ.

The gel cleavage assay SDS-PAGE gel (fig. 17 A) shows bands signifying K11 di-ubiquitin at ~16kDa, almost aligning to the true molecular weight of the substrate (17kDa), with the band intensity decreasing rapidly over the duration of the assay, in relation to the band present at ~9kDa representing ubiquitin, which increases rapidly over time, most evident when observing the overnight sample, the bands relating to USP17D1D2-2GKG however remain consistent through every time point. Unexpected bands appeared at 19 and 30kDa, remaining stable throughout.

The densitometric data presents the decrease of density of the K11 band as the assay progressed, upon further analysis, the EC50 was calculated to be 0.0374 hours/ 2.244 minutes, marginally more rapid than USP2CD/K6 assay with 2.7318 minutes, yet significantly slower than USP17D1D2-2GKG/K6.

6.9 Gel cleavage assay using USP17D1D2-2GKG to cleave a GG-linked di-ubiquitin.

Following the success of the K6-linked di-ubiquitin mimetic with USP2CD and USP17D1D2-2GKG, GG-linked di-ubiquitin was intended to be tested next (provided by Jiatong Zhang, with a concentration of 221μM, since this di-ubiquitin is a standard lysine bond between the ubiquitin molecules with a diglycine tag, the substrate was tested against USP17D1D2-2GKG to observe any differences between the other lysine bound di-ubiquitin substrates.

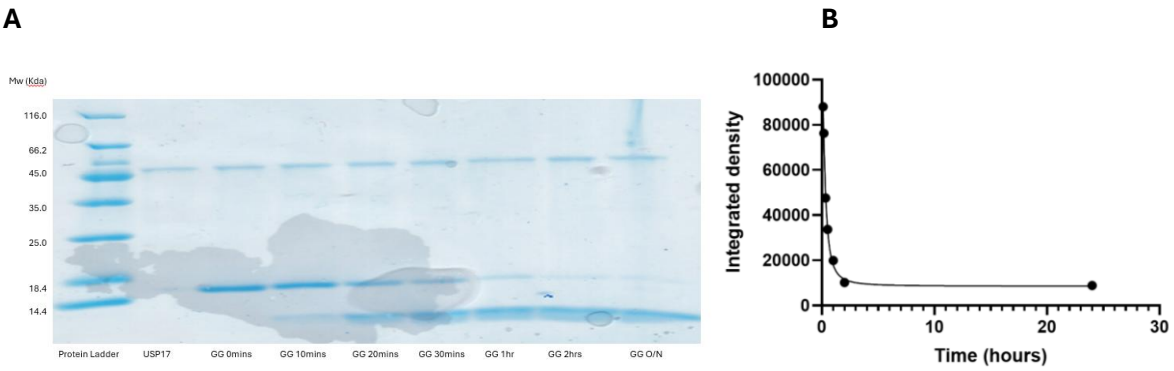


Figure 18- A) Image of SDS-PAGE stained with Coomassie blue, lane 1 was occupied by the protein ladder which was, lane 2 was occupied by 10μl 1.5μM USP17D1D2-2GKG and 10μl SDS-PAGE loading dye, lane 3 contains 10μl 3.0μM USP17D1D2-2GKG + 10μl 20μM GG-linked di-ubiquitin +10μl SDS-PAGE loading dye, lanes 4-9 are composed of an identical sample as lane 3 with the SDS-PAGE loading dye being added at different timepoints (10 mins, 20 mins, 30 mins, 1 hour, 2 hours, and 24 hours). B) Densitometric data of the 1.5μM USP17D1D2-2GKG and 10μM GG-linked di-ubiquitin gel cleavage assay, with the values of integrated density taken using an identical sample area of the gel, the values were initially deducted from the blank, the deducted values are present on the table, the density was calculated using ImageJ.

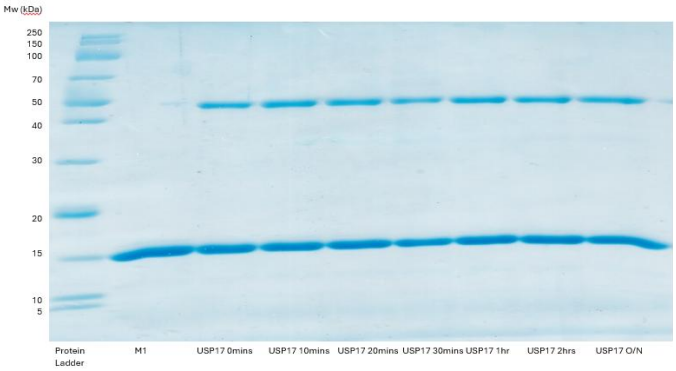
The gel cleavage assay between 1.5μM USP17D1D2-2GKG and 10μM GG-linked di-ubiquitin, present three bands of interest, one representing GG-linked di-ubiquitin present at 17kDa, aligning perfectly with its true molecular weight, the band initially has high density, however can be seen becoming less intense with each increasing time point, the band forming at 9kDa is inversely proportional to the intensity of GG, indicating the band being ubiquitin, due to the molecular weight being 8.6kDa, throughout the assay a band is present at 51kDa, which is representative of USP17D1D2-2GKG, additionally unlike GG, the bands remain identical regardless of time.

The densitometric data shows a steep regression curve, with the EC50 value being 0.282 hours/16.92 minutes, when observing other USP17 assays, the cleavage rate is 7.5x slower than K11, and 125x slower than the measured EC50 of the K6 mimetic, an observation which was also made was the fact that the regression line plateaus at 8500, a similarity shared with USP17/K6.

6.10 Gel cleavage assay using USP17D1D2-2GKG to cleave linear di-ubiquitin.

Since USP17D1D2-2GKG has been documented to not cleave the peptide bond in linear di-ubiquitin, a gel cleavage assay was conducted to ensure that the literature was factual.

A



B

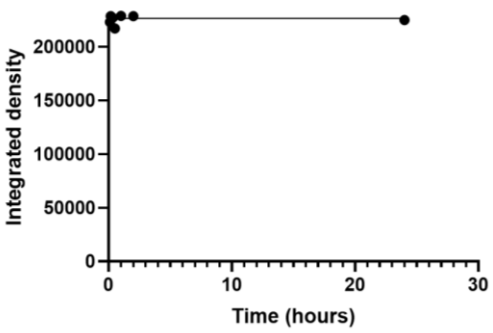


Figure 19-A) Image of SDS-PAGE stained with Coomassie blue lane 1 was occupied by 10µl protein ladder #26630 supplied by Thermo Scientific, lane 2 was occupied by 10µl 10µM linear di-ubiquitin and 10µl SDS-PAGE loading dye, lane 3 contains 10µl 3.0µM USP17D1D2-2GKG + 10µl 20µM linear di-ubiquitin +10µl SDS-PAGE loading dye, lanes 4-9 are composed of an identical sample as lane 3 with the SDS-PAGE loading dye being added at different timepoints (10 mins, 20 mins, 30 mins, 1 hour, 2 hours, and 24 hours). B) Densitometric data of the gel cleavage assay of USP17D1D2-2GKG and linear di-ubiquitin, the density was measured using an equally sized sample of the linear di-ubiquitin, and subtracted from the blank, with the density values already being deducted, with the density being measured using ImageJ.

Unlike previous gel cleavage figures, the SDS-PAGE gel of the gel cleavage assay between USP17D1D2-2GKG and linear di-ubiquitin (fig. 19 A), shows the linear di-ubiquitin present at ~16kDa, aligning to the molecular weight of 17kDa, the di-ubiquitin showing intense density for the duration of the assay, with no visible ubiquitin bands present at 9kDa, similarly USP17D1D2-2GKG remains consistent throughout the procedure as well.

The densitometric data (fig. 19 B) indicates that the EC50 is unattainable due to the fact that the regression line does not have a curve angle, due to the density value of the overnight sample being higher than the initial sample, while the sample whereby the reaction terminated at 30 minutes, shows the density decreasing slightly, when observing the SDS-PAGE gel (fig. 19 A), the band does appear smaller upon comparing with the bands at the same molecular weight.

6.11 Fluorescence polarization assay using USP17D1D2-2GKG to cleave di-ubiquitin.

Once the purified USP17D1D2-2GKG protein was tested using a gel cleavage against a known substrate, being the ISO-MIM di-ubiquitin, the USP17D1D2-2GKG was diluted into 4 concentrations, with a second sample of 1µM being denatured in 98.0°c, these samples were aliquoted to the same concentration of di-ubiquitin fluorescent probe marked with fluorescein-5-maleimide, with the protocol on the microplate reader being identical to the one of USP2CD, in order to create a direct comparison.

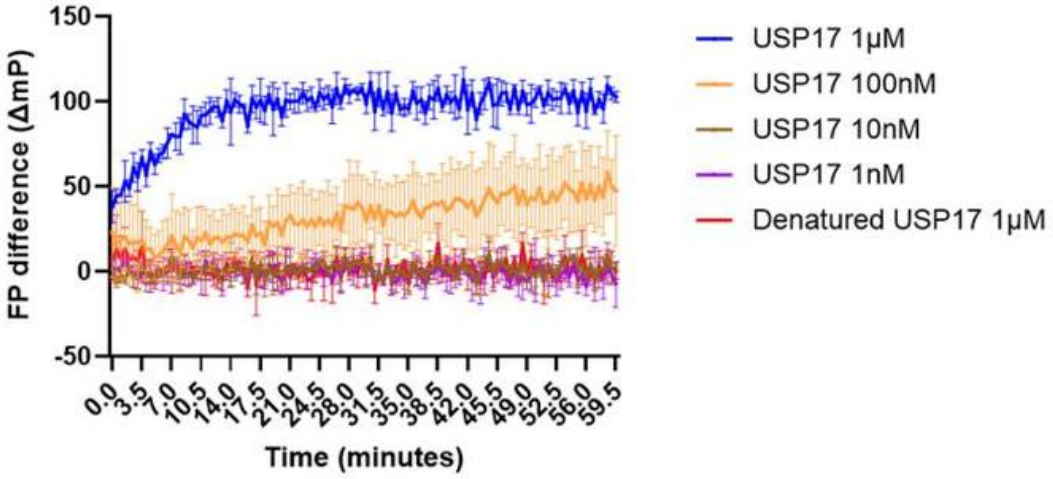
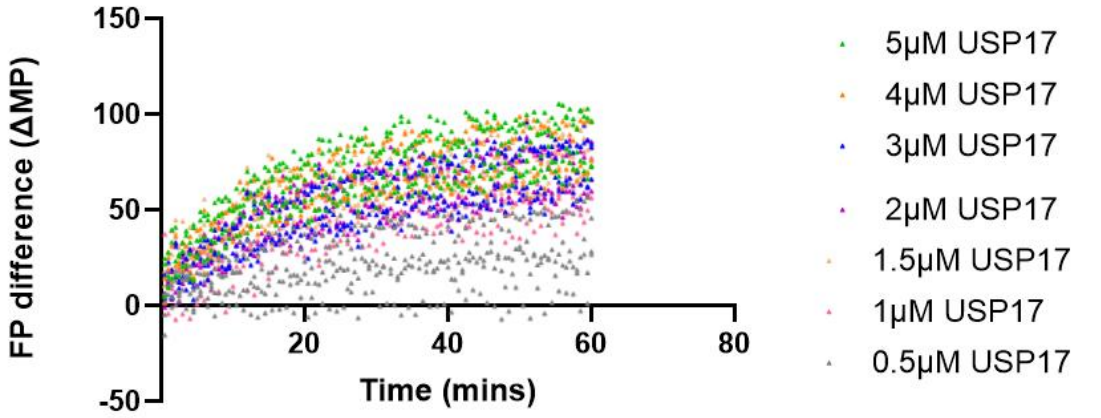


Figure 20- Scatter graph showing the FP difference of the di-ubiquitin fluorescent probe over a period of one hour with error bars, and the readings being conducted every 30 seconds, with different concentrations of USP17D1D2-2GKG being used, with the concentrations being colour coded by the figure present on the right of the graph. The measurements were conducted using the EnVision Xcite multimode plate reader.

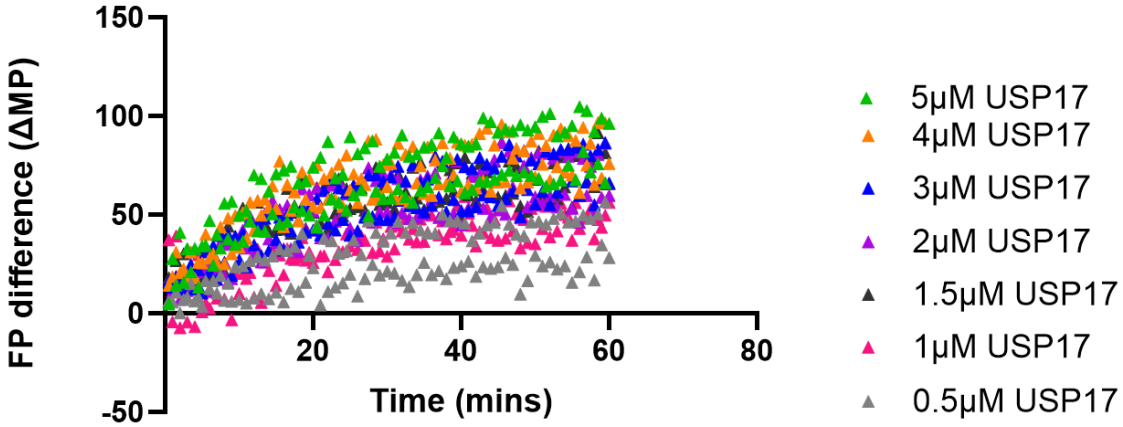
The fluorescence polarization assay between varying concentrations of USP17D1D2-2GKG and the fluorescent di-ubiquitin probe (fig. 20), the calculations to obtain the FP difference were identical to fig. 11, similarly, a tenfold dilution of USP17D1D2-2GKG was accomplished using 33nM of di-ubiquitin probe. Unlike fig. 11, the highest concentration of enzyme (1μM), presents with a dramatic increase in FP difference before plateauing at 14 minutes, 100nM of enzyme increases at a more linear range, with all other concentrations and the denatured enzyme sample remaining at 0 or increasing slightly, however temporarily. Three repeats were conducted for each enzyme concentration, with the mean being calculated and subtracted from the FP of the probe.

6.12 Fluorescence polarization competition assay using USP17D1D2-2GKG and a di-ubiquitin probe/ K6-linked di-ubiquitin mimetic.

Further to the assay only including the fluorescent di-ubiquitin probe, the task was to identify the specificity of USP17D1D2-2GKG, specifically against the K6-diubiquitin mimetic, therefore a competition fluorescence polarization assay was conducted to observe if an FP difference was present, with an increase suggesting that the di-ubiquitin probe outcompeted the K6-diubiquitin mimetic.



A



B

Figure 21- Scatter graph showing the results of the competition fluorescence polarization assay between USP17D1D2-2GKG and the fluorescent di-ubiquitin probe and the K6-linked di-ubiquitin mimetic over the course of 1 hour, with readings being conducted every 30 seconds, a colour-coded legend can be found on the right of both graphs. The readings were accomplished using the EnVision Xcite multimode plate reader.

USP17D1D2-2GKG was further tested with the K6-linked di-ubiquitin mimetic to compete for the catalytic domain with the fluorescent di-ubiquitin probe, the FP difference using the measurement of the probe and sample formula. Image A shows the overall results for the assay with all repeats included, image B shows the most idyllic repeat from each enzyme concentration being selected, due to image A showing results that are deemed anomalous, because of the presence of air bubbles. Image B indicates the curves increasing proportionally to the enzyme concentration, with all concentrations increasing the FP difference between the probe, however with some results showing that at certain time points, the 4μM USP17D1D2-2GKG had a higher FP difference compared to 5μM, yet for the remainder of the assay, 5μM had the most substantial increase.

6.13 Fluorescence polarization assay using USP17D1D2-2GKG to cleave a mono-ubiquitin probe

USP17D1D2-2GKG was also introduced to the mono-ubiquitin probe to test the binding potential to a singular molecule of ubiquitin with the only chemical modification being the addition of fluorescein-5-maleimide in order to quantify the general binding potential of the enzyme, with the concentration of USP17D1D2-2GKG being diluted by 10x to identify the difference between each diluted sample.

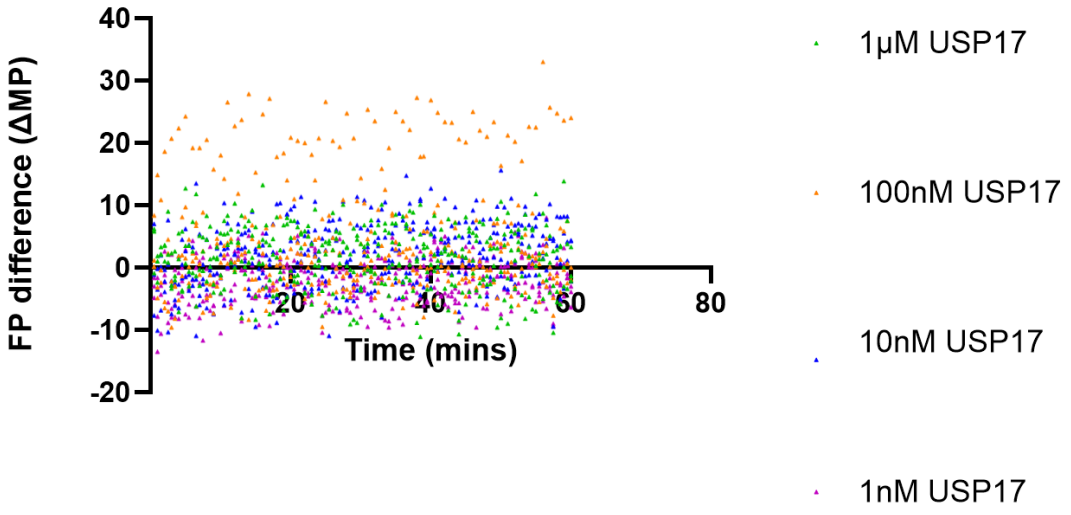
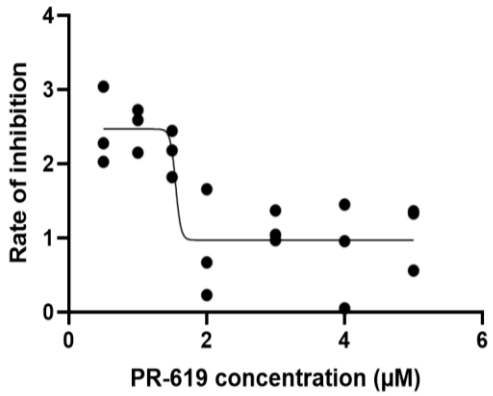


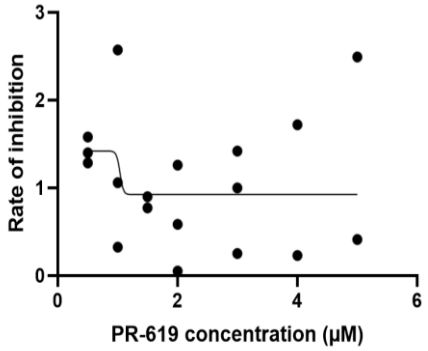
Figure 22- Scatter graph showing the results of the fluorescence polarization assay including various concentrations of USP17D1D2-2GKG and the mono-ubiquitin probe, with the colour coded figure legend present on the right of the graph, the assay was conducted on the EnVision Xcite multimode plate reader, over the course of one hour, with a reading every 30 seconds.

Following the introduction of a probe not yet studied on USP17D1D2-2GKG, the fluorescence polarization assay uses the calculated FP difference values using the appropriate formula, yet upon inspection of the obtained data, there is seemingly no correlation or activity present between USP17D1D2-2GKG and the 10nM mono-ubiquitin probe, despite no noted air bubbles present within the wells; which can be inferred when observing the graph, regardless of the concentration of enzyme, even with the 100nM USP17D1D2-2GKG, despite being present at 20ΔMP consistently, the highest other data point being 10nM USP17D1D2-2GKG at 51 minutes (15.6ΔMP), there is no stable increase in correlation with time.

6.14 IC50 studies on USP17D1D2-2GKG against the small molecule inhibitor, PR-619, using fluorescence polarization.

Due to the IC50 of the USP17D1D2-2GKG not being currently established, a study was conducted to investigate the inhibitive effects of PR-619 on the enzyme's potential to prevent the binding of a di-ubiquitin fluorescent probe to the catalytic domain of USP17D1D2-2GKG, the procedure was repeated three times to observe any consistencies/ invalidate any anomalies that may occur during the measurement.





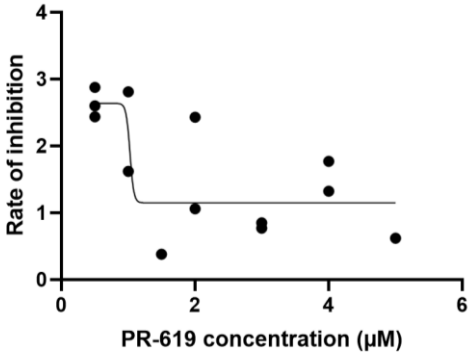


Figure 23-Three repeats of the IC50 studies on USP17D1D2-2GKG using the small molecule inhibitor, PR-619, each being numbered in the title, the corresponding tables of values each being present below the graphs, the measurements were conducted using an EnVision Xcite multimode plate reader for 15 minutes, with readings being accomplished every 30 seconds.

Once the data was received, initially the FP values of the triplicates for each enzyme concentration were converted into FP difference (ΔMP), the values taken from 0 minutes and 15 minutes were subtracted, resulting in the rate of inhibition, data omissions were decided as anomalous due to the presence of air bubbles within the wells. IC50 curve 1 is the only graph with zero omissions due to anomaly, in addition to having the largest span of 1.498, with IC50 curve 3 being slightly shorter with a span of 1.488, and the IC50 curve 2 having a significantly shorter span of 0.497. The calculated IC50 of each curve is as follows:

IC50 of curve 1: 1.549μM

IC50 of curve 2: 1.038μM

IC50 of curve 3: 1.026μM.

6.15 DNA sequencing of USP17D1D2-2GKG following mutagenesis

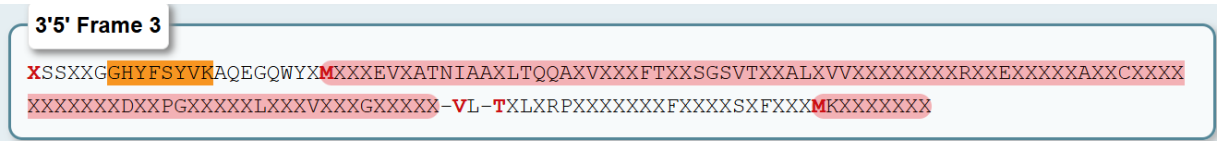
To ensure that the purified protein was mutated successfully, with surrounding residues also matching, USP17D1D2-D332G2GKG was sampled by Source Bioscience with appropriate forward and reverse primers.

Quality Value	> 35	20 - 35	< 20
LOR Value	> 600bp	301-600bp	< 300bp

Reaction	Template Name	Primer Name	N-Base Calls	Length of Read	Quality Value
1	USP17D332G2GKG	T7promoterF	36 %	153 bp	10
2	USP17D332G2GKG	T7terminateR	27 %	314 bp	15

Figure 24 – Table provided by Source Bioscience from the DNA sequencing procedure, with the figure legend above the table of results.

Sample 3 from the mutagenesis triplicates was selected as samples 2 and 1 consistently showed a substantially lower quality/ incorrect amino acid sequences. 2 reactions were conducted, both involving the T7 promoter and terminator for the forward and reverse reactions respectively. The DNA sequences for both were pasted into a Expasy (a translate tool).



A

B

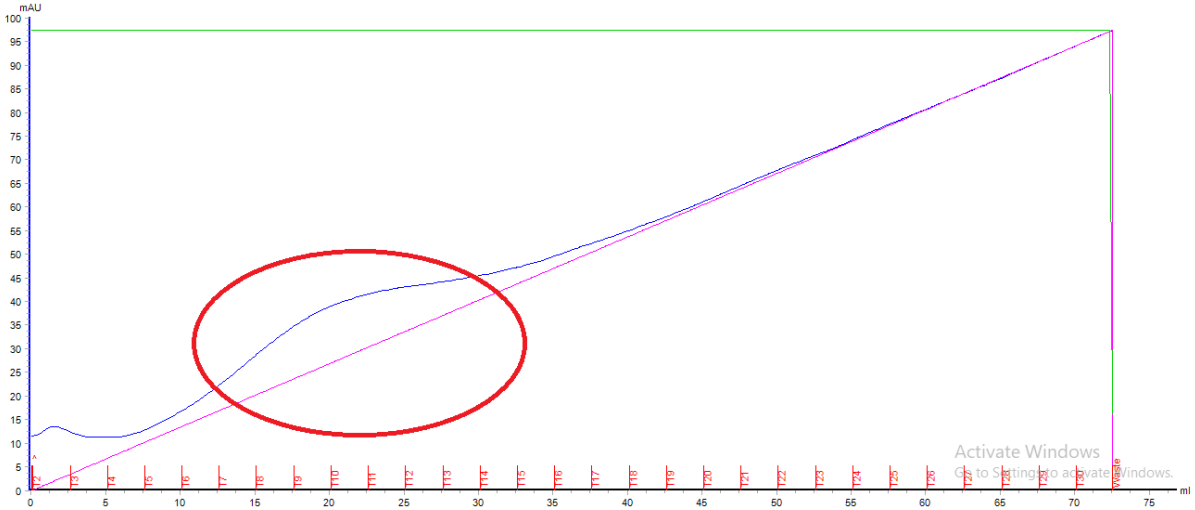
Figure 25 – A) Result from the 3' to 5' sequence read, translated using the Expasy tool. B) The full amino acid sequence of USP17 as shown by PDB. The matching sequences of interest were highlighted orange.

Although the sample was deemed as low quality by the system, inserting the forward and reverse sequences to Expasy, provided one of the frames with a sufficient residue sequence. Confirming two results: The USP17D1D2-2GKG initially used for mutagenesis is the correct protein, and the aspartic acid present in the 332 position has been successfully mutated to a glycine. Therefore, the analysis through the use of assays could commence for the USP17 constructs.

6.16 Expression and purification of USP17D1D2-D332G2GKG catalytic domain by affinity chromatography.

Initially, upon mini prepping the USP17D1D2-2GKG protein plasmid DNA, a sample was saved for the mutagenesis process, using annealing at 62°c and instead of leaving the DNA for a total of 3 hours, the DNA had an incubatory period of 21 hours, three repeats of the DNA were stored and their amino acid sequence was tested to ensure the correct mutation was completed before initiating expression.

A



B

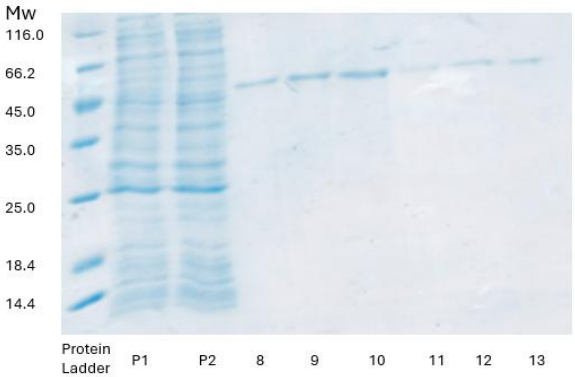
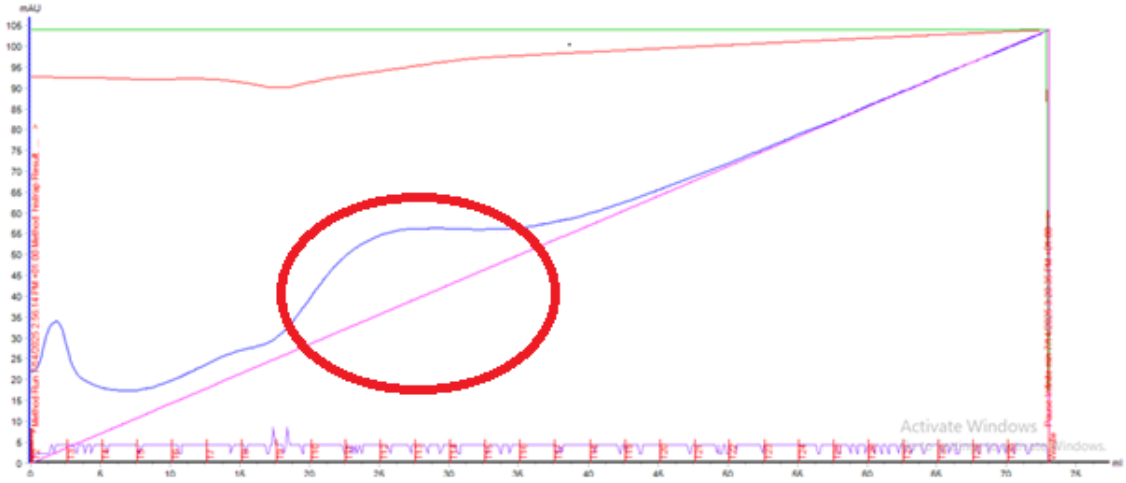


Figure 26- A) Chromatogram of the purification of USP17D1D2-D332G2GKG using the histrap column on the Akta Start 2 purifier, with a linear increase of imidazole from 20mM to 500mM, fractions of interest included 8-13, as these were deemed to be related to the most significant peak, the fractions were subsequently analysed on SDS-PAGE. B) Image of SDS-PAGE stained with Coomassie blue, the leftmost lane is occupied by the protein ladder, which was the unstained protein molecular weight marker 116-14.4 KDa supplied by Thermo Scientific, lanes 2 and 3 were occupied by both pellets retrieved from centrifugated post-sonication

supernatant, the remaining lanes were occupied by the fractions collected from the histrap column purification, the fractions were chosen as these were in relation to the most significant peak present on the chromatogram .

A



B

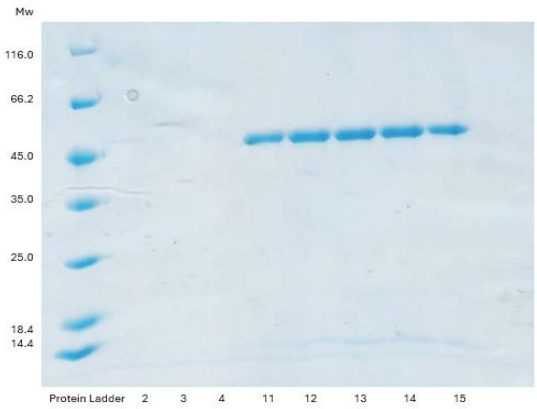


Figure 27- A) Chromatogram of the second purification of USP17D1D2-D332G2GKG using the histrap column on the Akta Start 2 purifier, with a linear increase of imidazole from 20mM to 500mM, fractions of interest included 8-13, as these were deemed to be related to the most significant peak, the fractions were subsequently analysed on SDS-PAGE. B) Image of SDS-PAGE stained with Coomassie blue, the leftmost lane is occupied by the protein ladder, which was the unstained protein molecular weight marker 116-14.4 KDa supplied by Thermo Scientific, lanes 2-4 were occupied by fractions 2-4 which were collected from the histrap column purifications, as seen on the minor peak on the chromatogram (fig.27 A), the remaining lanes were occupied by the fractions 11-15 which corresponded to the major peak on the chromatogram(fig.27 A).

The initial protein purification of USP17D1D2-D332G2GKG was successful, however it was noted that the band intensity was lower in comparison to the previous purifications, with peak absorbance reaching ~37mAU (fig.26 A), further inferred by the SDS-PAGE which showed bands of lower intensity by comparison (fig.27 A). Fractions 8-12 were selected for concentration, yet due to the omission of buffer exchange following the addition of fractions, meaning imidazole was deemed to be present within the fractions, to ensure that the protein was non-functioning, a fluorescence polarization assay was conducted using a standard di-ubiquitin probe, whereby the wildtype showed a curve being in direct relation to the concentration of enzyme, in direct

contrast, the batch of collected USP17D1D2-D332G2GKG showed a more random, nonsensical scatter graph, therefore a decision was made to discard the enzyme.

A repeated expression and purification of USP17D1D2-D332G2GKG was initiated, utilizing glycerol stocks which were assembled and stored at -80°c during the first attempt. Comparatively, the peak present on the chromatogram of the second attempt (fig. 27 A) is significantly less subtle and peaks at a higher absorbance (~55mAU). Since two peaks were present on the chromatogram, corresponding fractions for both peaks were analysed using SDS-PAGE (fig. 27 B), fractions 2-4 do not show any discernible bands present at the appropriate molecular weight (approximately 51.0KDa), therefore only fractions 11-15, containing bands at the appropriate Mw and a lack of visible contaminants, were chosen to undergo concentrating followed by buffer exchange, with the final concentration reaching 10.62µM and the protein being distributed evenly in 20 tubes each with a volume of 50µl. The initial concentration of the first purification reached 25.7µM, however this reading was disregarded as the large volume of imidazole would have given an inaccurate concentration of the protein alone, therefore yield was calculated only for the second purification.

Mass= concentration x volume x molecular weight

Mass= 10.62µM x 1ml x 51.4054kDa

Mass= 10.62 x 1 x 51405.4

Mass= 545921ng

Mass= 545.921µg

6.17 Gel cleavage assay using USP17D1D2-2GKG to cleave a K6-linked di-ubiquitin mimetic.

To address the lack of a confirmed 3D crystalline structure of all USP17 constructs, the binding for the K6/K11-linked di-ubiquitin was predicted for visualisation purposes.

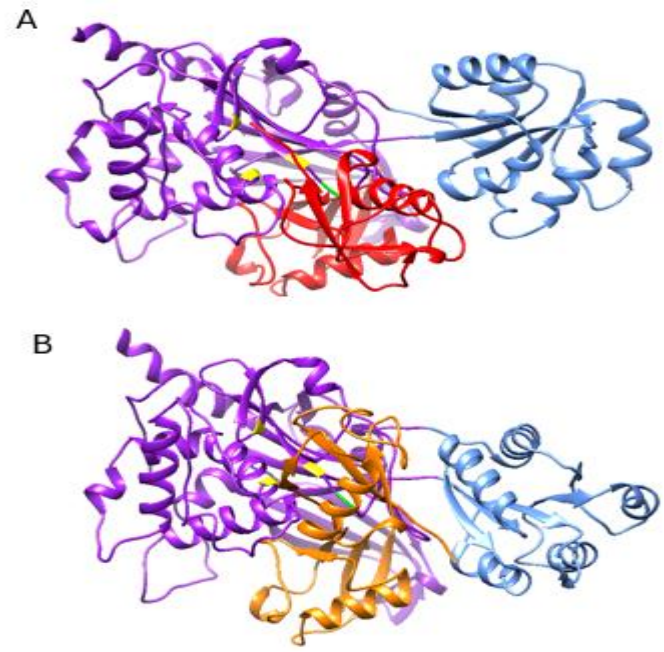
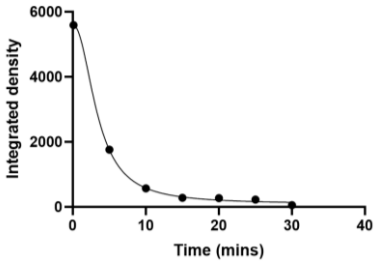


Figure 28-A) USP17D1D2-D332G2GKG bound to a K6-linked di-ubiquitin mimetic, with the colour legend: Purple- catalytic domain, yellow- catalytic triad, green- D332G mutation, blue- 2GKG insert, red- K6-linked di-ubiquitin mimetic. B) USP17D1D2-D332G2GKG bound to a K11-linked di-ubiquitin mimetic, with the same colour legend, except for K11, which is coloured orange.

Upon purification being successful, the mutation was hypothesized to not affect the binding of other di-ubiquitin complexes, therefore, to initiate the testing, USP17D1D2-D332G2GKG was retrieved and diluted to three varying concentrations: 1µM, 600nM, and 200nM, with the K6-linked di-ubiquitin mimetic being diluted to 20µM, all three gel cleavage assays were accomplished on the same day and were accomplished using three separate SDS-PAGE gels.

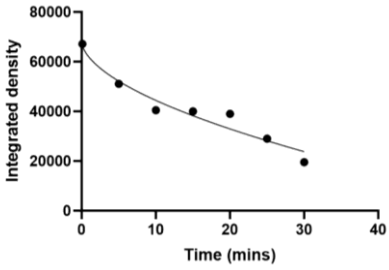
A

B



C

D



E F

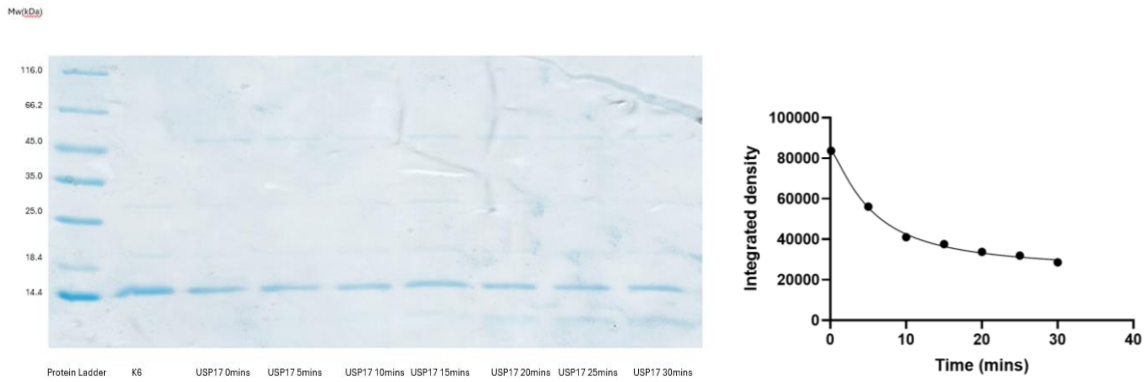


Figure 29- (A, C, E)Images of SDS-PAGE stained with Coomassie blue, lane 1 of all the gels is composed of the protein ladder, which was the #26610 unstained protein marker from Thermo Scientific, lanes 2 were composed of 10µl 10µM K6-linked di-ubiquitin and 10µl SDS-PAGE loading dye, lanes 3-9 of all gel images were composed of 10µl 20µM K6-linked di-ubiquitin + varying concentrations of 10µl USP17D1D2-D332G2GKG (1µM of USP17D1D2-D332G2GKG (A), 600nM of USP17D1D2-D332G2GKG (C), 200nM of USP17D1D2-D332G2GKG (E) + 10µl SDS-PAGE loading dye. (B, D, F) Densitometric data of the 500nM USP17D1D2-D332G2GKG(B), 300nM USP17D1D2-D332G2GKG (D), and 100nM USP17D1D2-D332G2GKG (F) gel cleavage assays, all assays used 10µM of the K6-linked di-ubiquitin mimetic, identical sample areas were used for individual gels, the sample area may differ across the gels however. Integrated density was measured using ImageJ.

The gel cleavage assays, although using differing concentrations of USP17D1D2-D332G2GKG, share visual similarities, such as the K6-linked di-ubiquitin mimetic bands all presenting at the same molecular weight, being 16kDa, which almost aligns with the molecular weight of 17kDa, another similarity shared among the gels are the additional bands associated with the K6 mimetic sample present at 19kDa and 30kDa (although less visible in image C), another visual similarity is the presence of the bands present at 51kDa, aligning with the molecular weight of USP17D1D2-D332G2GKG, and finally the presence of bands seen at 9kDa, representing the presence of the product formed during the reaction, ubiquitin, which has a molecular weight of 8.6kDa.

The difference between the gels are all related to their density, with USP17D1D2-D332G2GKG bands increasing in density in relation to their concentration, bands present in image A are most visible and decreasing steadily as the concentration lowers, as a result the intensity of the K6 mimetic bands decrease more intensely and faster when the concentration of USP17D1D2-D332G2GKG is higher, leading to ubiquitin being formed more slowly in image C as compared to image A.

The densitometric data of all gel cleavage assays shows scatter graphs with varying slope regression lines, image A signifying 500nM USP17D1D2-D332G2GKG, contains a regression line which plateaus near to the x-axis, matching with the information given on the SDS-PAGE gel (fig. 29 A), the band representing the K6 mimetic is almost dissipated once 30-minute passes, the EC50 value of image A was calculated to be 3.391 minutes. Image C presents a regression line that is unstable due to the decrease of values not following the decrease of density between 0 and 5 minutes, upon removing data points of 15 minutes, 20 minutes, and 25 minutes, resulting in the regression line becoming stable, the EC50 was calculated to be 19.76 minutes, a 5.83-

fold increase from 500nM. Image E presents the regression line with a steeper angle than image C, however a line which plateaus at ~25000 integrated density, yet due to the initial decrease being significant, the EC50 was calculated to be 5.388 minutes, 3.67-fold decrease in time as compared to 300nM.

6.18 Gel cleavage assay of USP17D1D2-D332G2GKG and linear di-ubiquitin.

According to the literature stating that USP11 will be able to cleave linear di-ubiquitin upon mutagenesis being conducted to alter the aspartic acid to a glycine, a USP17D1D2-D332G2GKG sample was tested against linear di-ubiquitin, with an expectation for cleavage to occur and mono-ubiquitin being formed, with the results of the reaction being visible on an SDS-PAGE gel.

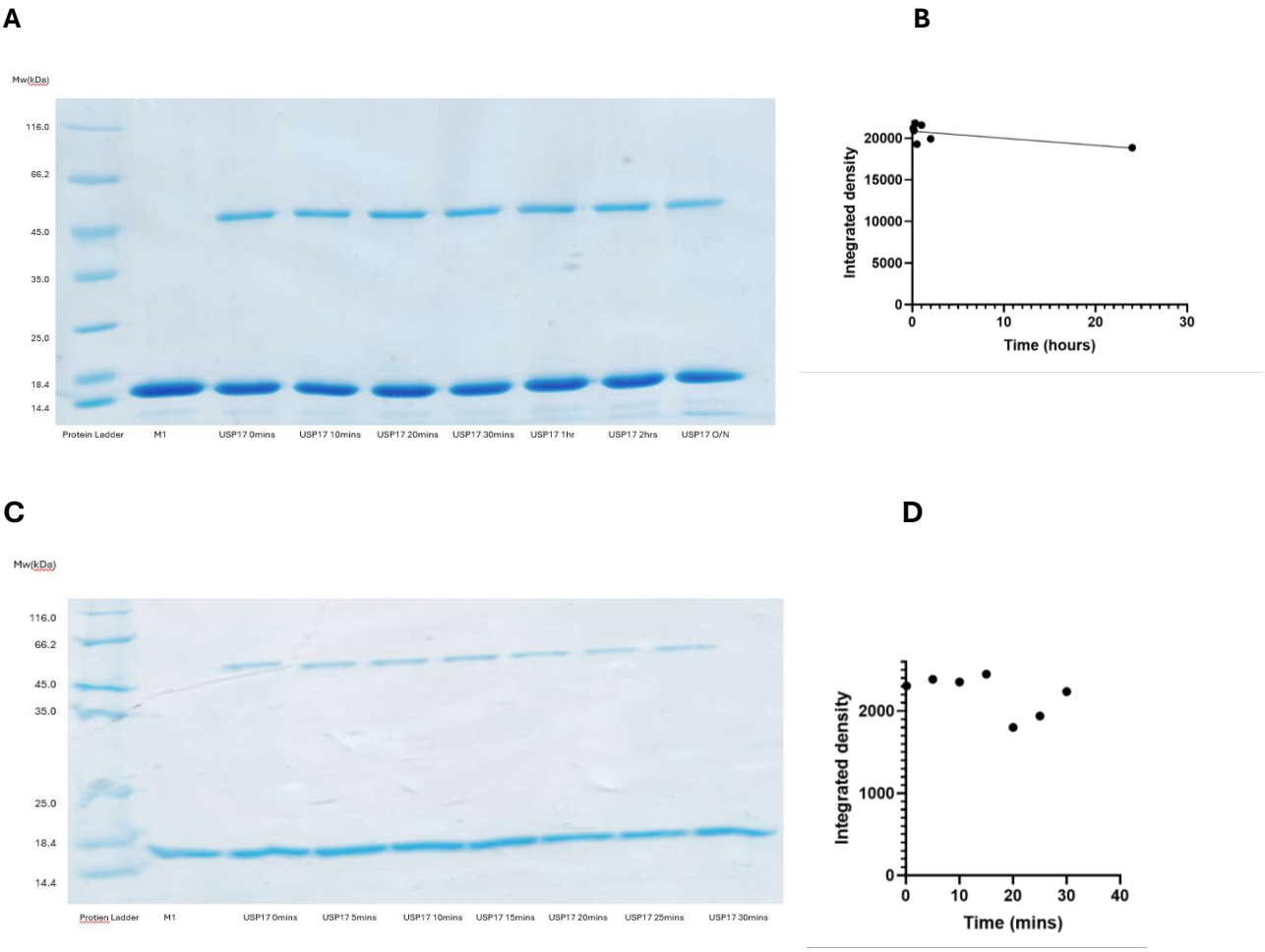


Figure 30- Images of SDS-PAGE gels stained with Coomassie blue, lane 1 for both images was occupied by the protein ladder, more specifically the #26610 unstained protein marker from Thermo Scientific, lane 2 for both images was occupied by 10µl 20µM linear di-ubiquitin + 10µl 5x SDS-PAGE loading dye, lanes 3-9 were occupied by 10µl 3µM (image A) and 1µM (image C) USP17D1D2-D332G2GKG + 10µl 20µM linear di-ubiquitin +10µl 5x SDS-PAGE loading dye. Densitometric data of the 1.5µM (B) and 500nM (D) USP17D1D2-D332G2GKG and 10µM linear di-ubiquitin gel cleavage assays, in addition to the respective tables of values, the integrated density values were calculated using an identical sample area and subtracted from the blank, also with an identical sample area, these values were received with imageJ.

Both gel cleave assays (fig. 30 A, C) show the linear di-ubiquitin present at 17kDa, hence matching the molecular weight of the protein complex, with the bands present at 51kDa being representative of USP17D1D2-D332G2GKG, similarly to the wildtype, no bands were shown to represent the formation of ubiquitin, indicated by the loss of density in the linear di-ubiquitin bands as time progressed, and the formation of the ubiquitin bands which would have presented at 9kDa, this is the case for image B, however upon inspection of the overnight sample present in image A, and comparing the band to the ones prior, the density does decrease slightly with seemingly an appearance of a band below linear di-ubiquitin, a repeat was conducted alongside USP2CD, ensuring the band appeared at the correct height.

The densitometric graphs (fig. 30 B, D) contain the data presented by the gel cleavage assays (fig. 30 A, C), although image C shows no correlation between the time and the density of linear di-ubiquitin upon being exposed to USP17D1D2-2GKG (further reinforced by the density values present in table D), image A shows a minor regression in relation from 0 minutes to overnight, the density regression being -85.06/hour, an EC50 value could not be estimated as the regression line was linear and not a curve, therefore if the assay were to theoretically continue indefinitely, the calculated time to reach 0 density on the linear di-ubiquitin band, is 221.66 hours/ 9.24 days.

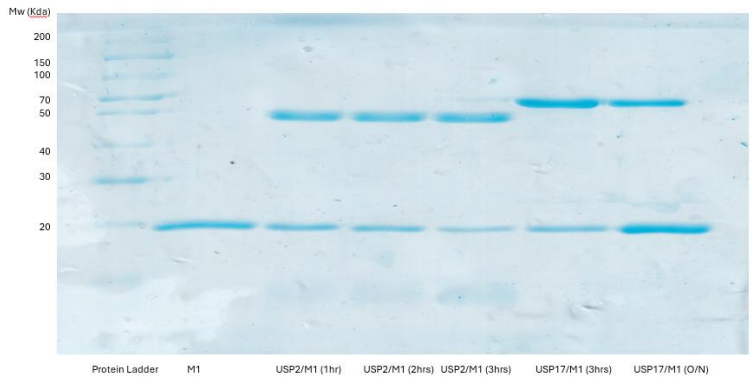


Figure 31- Image of an SDS-PAGE stained with Coomassie blue, lane 1 is composed of the protein ladder which is the 26630 broad range protein ladder, lane 2 is composed of 10μM of linear di-ubiquitin and 10μl SDS-PAGE loading dye, lanes 3-5 are composed of 3μM of USP2CD + 10μM linear di-ubiquitin whereby the reaction was terminated at an hourly rate up to 3 hours, lanes 6 and 7 are composed of 3μM USP17D1D2-D332G2GKG + 10μM linear di-ubiquitin with the reaction being cleaved at 3 hours and overnight respectively.

The repeated 1.5μM USP2CD/ USP17D1D2-D332G2GKG and 10μM linear di-ubiquitin assay (fig. 31), presents with linear di-ubiquitin bands present at 17kDa, remaining constant throughout both deubiquitinase reactions, lanes 3-5 contain bands at 39kDa representative of USP2CD, while lanes 6 and 7 contain bands at 51kDa signifying the presence of USP17D1D2-D332G2GKG, both enzyme bands appear to not decrease in density with the assay progressing. Unlike the enzymes, the density of linear di-ubiquitin decreases upon being exposed to USP2CD, further inferred by the appearance of bands at 9kDa, being ubiquitin, however for USP17D1D2-D332G2GKG, the density of linear di-ubiquitin increases in correlation to time, with no ubiquitin present.

6.19 Crystallisation Studies on USP17D1D2-2GKG and the USP17D1D2-2GKG/GG-linked di-ubiquitin complex.

Three crystallisation studies were conducted, one for the USP17D1D2-2GKG alone, one for the mutant of the subsequent deubiquitinase (USP17D1D2-D332G2GKG), and one for the complex of USP17D1D2-2GKG and GG-linked di-ubiquitin. All solutions were aliquoted to Morpheus screens with a reservoir of gel filtration buffer, the screens were placed at a constant temperature of 20°c for a duration of 2 months. Upon microscopic evaluation being done for all proteins, despite a single well on the USP17D1D2-2GKG GG complex showing a promise of precipitation being visible, once placed under ultraviolet light, the mass showed little protein in its composition, it was deemed that the identity of the mass was a contaminant due to a sealing error, with zero significant crystal structures present on the remaining screens, no results were obtained for the trials.

7. Discussion

7.1 Rationale of construct use.

The first construct purified was the USP2 catalytic domain (USP2CD), as this deubiquitinating enzyme, from known literature at least, is capable of cleaving all the ubiquitin moieties used in the catalytic research of USP17. Only the catalytic domain was purified, as no crystallization studies were planned for USP2, with the only interest of use being a direct comparison against the USP17 constructs. 4 different USP17 constructs were utilised throughout the research (fig.14), $USP17^{71\text{-}402}$-$2GKG^{3\text{-}123}$ was used to analyse the catalytic domain as the 2GKG crystallization tag does not prohibit the enzyme from any catalytic activity and would be expected to behave identically to the full length USP17. As the crystal structure of USP17 is yet to be solved, the 2GKG tag was inserted to increase the protein's solubility in gel filtration buffer, truncation of the non-catalytic residues was deemed necessary for an even higher likelihood of crystal formation as smaller proteins are more likely to crystallize[80].

Mutating the cysteine residue at position 89 into a serine was essential to ensuring that if one of the catalytic triad residues was mutated, USP17 was unable to cleave any ubiquitin moieties it would typically be able to. Therefore, USP17D1D2-C89S2GKG and USP17D1D2-D332GC89S2GKG upon being introduced to a liquid matrix containing the K6-linked di-ubiquitin mimetic, both proteins were unable to produce any ubiquitin, which is in stark contrast to USP17 constructs that did not contain the C89S mutation. The mutagenesis of USP17D1D2-2GKG to USP17D1D2-D332G2GKG was heavily inspired by the USP11D1D2D886G mutation, documenting the significant improvement of binding affinity against linear di-ubiquitin compared to the wildtype- therefore the concluding remarks were due to steric hindrance, the aspartate repelled the substrate, prohibiting binding[83].

The rationale behind using the mimetic was to reduce the difficulty of manufacturing a substrate in large quantities for research purposes, therefore, to mimic the isopeptide bond which is present in the in-vivo variant of the K6 and K11-linked di-ubiquitin, the first 6 and 11 residues were truncated respectively. Truncation is considered as a reliable modification to ensure the bond occurs at the correct orientation and position, without the possibility of binding to an incorrect residue. Although this was achieved, the predicted structures (fig. 8), suggest that the entire tertiary structure of the di-ubiquitin is slightly altered, therefore the stability and catalytic functionality was put into question. However as the data analysis spanned from November 2024- August 2025 using the ubiquitin mimetics from the same batch, no instability

or change in function was noted, therefore truncation is ideal as it allows the moieties to behave identically to the in-vivo counterparts, without noticeable compromise.

7.2 Purification of USP2/USP17 and mutagenesis of USP17

The multiple purifications performed were successful in purifying the appropriate proteins with minimal impurities, all purifications used the same excel histrap column, which does not require ethylenediaminetetraacetic acid (EDTA) and nickel to be reintroduced into the column, allowing for more convenient usage of the column, the column has also demonstrated success in the appropriate removal of more complex proteins such as USP17D1D2-2GKG and its mutated variant, USP17D1D2-D332G2GKG. Histrap excel columns are also most commonly used in samples of eukaryotic supernatant containing proteins rich in histidine, the nickel ions present in the column are not susceptible to contaminants binding to the nickel ions due to their increased resistance as compared to an HP histrap column.

Further analysis of the chromatograms solidifies the validation of the correct protein being present, as the complexes which were targeted for elution all exceed 50kDa, meaning as a larger complex, the protein is expected to elute within the first ten fractions (fig.7)(fig.15)(fig.27).

The peak which was associated with USP2 (fig.7 A) is significantly more prominent than the peak associated with USP17D1D2-2GKG, USP2 peaking at an absorbance of 118mAU, meanwhile USP17D1D2-2GKG peaking at 44mAU, which suggests a 2.68x lower yield, interestingly the actual yield of protein for USP17D1D2-2GKG (being 33.05mg) was 2.05x lower than that of USP2 (being 68.0mg), suggesting either the selection of fractions for concentrating was better for USP17, or some USP2 protein was lost during the concentrating process. The peak associated with USP17D1D2-D332G2GKG (fig.27) was even less prominent than that of the wildtype (fig.15), peaking at approximately 37mAU, suggesting that potentially some protein was denatured or lost during the mutagenesis process, which can occur if the annealing temperature is not desirable for the protein to mutate correctly (the temperature used was 62.0° c), furthermore the yield was significantly lower than that of the non-mutant, at 15.8mg.

Another factor which can impact the overall yield of the protein is the chosen method of clearing the imidazole from the sample, which is a necessary procedure as imidazole impacts the efficacy of the protein upon being flash frozen. Two methods wsafern to clear the imidazole, buffer exchange (used for USP2 and USP17D1D2-D332G2GKG) and dialysis (used for USP17D1D2-2GKG), both methods utilised an identical gel cleavage buffer recipe, with dialysis being deemed a more appropriate method for the protein, due to buffer exchange, while more time-efficient at retrieving the protein, ultimately product can be lost due to permeation, which can be evident based on the yield of USP17D1D2-D332G2GKG and USP2, dialysis, while time consuming, allows for a more safe approach to clearance, potentially causing the higher yield in USP17D1D2-2GKG.

Although histrap affinity protein purification was used to have a pure sample of both USP17 variants, there are several other procedures that can separate the enzyme from the gathered supernatant, for instance ammonium sulphate precipitation is a commonly used low-cost method for preparing samples before further chromatography, which uses multiple salt concentrations to salt out proteins based on their solubility, causing separation[84].

7.3 Gel Cleavage Assays

K6: USP17 showed a 19.9x increase in cleavage than USP2.

M1: 113.5x slower than K6 when cleaved by USP2, no activity for USP17 and the mutant.

K11: Selectivity comparable to USP2, much slower than K6.

The USP2 gel cleavage assay against the K6-linked di-ubiquitin mimetic(fig.8) was used as a control to ensure that the appropriate method of purification, concentration, and storage was being chosen and repeated for USP17, the choice of substrate, being K6-linked di-ubiquitin, was due to USP2 having a strong affinity towards K6, K11, K48 linked di-ubiquitin, therefore an assay which validated the protein was correct was imperative. This strong affinity can be confirmed on the image (fig.8 A) as the K6 mimetic was cleaved within one hour and did not degrade overnight, confirming that the substrate can be used for further testing on USP17.

Upon using the K6 mimetic and K11 in a gel environment, the samples contained contaminants present at 19 and 30kDa, most likely originating from the E. coli lysate used for expression and purification, since the bands remained not cleaved by the enzymes, infers that these contaminants are not multichain ubiquitin compounds. One of the potential theories for the identification of the protein with 19kDa, could be leucine-responsive regulatory protein, a 164bp regulatory protein, with the binding site only containing 15bp[85], the overall amino acid sequence shows 2 histidine amino acids at the 68 and 117 positions, therefore explaining the contamination when exposed to histags present within the column. On the contrary, proteins produced by E. coli with a molecular weight of 30kDa include the peptidyl-prolyl cis/trans isomerase, slyD[86], out of 48 amino acids present on the C-terminal tail, 15 of the amino acids were identified to be a histidine[87], therefore the likelihood of nickel binding and eluting at an identical time as the K6 mimetic is high.

After purification, USP17 was tested against 4 common di-ubiquitin constructs to identify the substrates which the enzyme can cleave successfully and within a 24-hour period, the findings confirm that K11 is a cleavable target by USP17, meanwhile there has been speculation surrounding the efficiency of cleaving K6, however as fig. 16 suggests, the di-ubiquitin showed clear signs of cleavage within only 10 minutes of the enzyme being introduced, which could be misinterpreted as USP17 being in excess due to the high concentration, yet fig. 28 further suggests that not only did the USP17D1D2-D332G2GKG variant cleave the K6, but the concentration was significantly reduced (up to 15x lower than the concentration present at fig.16), with 300nM of both proteins showing sufficient cleaving of the K6 mimetic at 5 minutes upon introduction.

Some improvements that could have been implemented include the gel cleavage assay of USP17D1D2-2GKG and K6, as the rightmost lane shows signs of the tube being improperly closed while the assay was ongoing, this misuse would have resulted in the contamination which can be seen on the SDS-PAGE gel (fig. 16), what can also be deduced is the fact that the contents were exposed to a different environment than the sealed tubes, primarily the temperature shift is less impactful on an unopen sample, therefore to mitigate this from occurring is to ensure the tube produces an audible click upon closing. Due to the intensity of the bands not reflecting the true result, a second blank was taken directly above the measured sample, as the initial blank would not be intense enough and give an incorrect reading.

Some gel images were also not taken with the best quality, as when water presses onto a transparent folder, darker circles will appear, potentially obstructing a clear view of the whole gel, this can clearly be observed at fig. 18, where the compressed water causes 4 bands of GG-linked di-ubiquitin to be submerged, additionally air bubbles pose the same issue. To rectify this

issue, multiple blanks were taken, since 4 bands of interest were within the patch of water, the blank value was strikingly higher than for the remaining 3 bands, however a simpler solution is to redo the image.

Furthermore, the estimated EC50 resulting from the densitometric data (fig. 16), was 19.9x faster than USP2CD, a possible explanation for this occurrence is that more time points were used for USP17D1D2-2GKG (10 minutes, 20 minutes, and 30 minutes) as compared to USP2, the greater number of timepoints would allow the graph to have a significantly steeper curve, therefore a repeat should have been replicated on USP2 and K6 mimetic to investigate whether the EC50 would have been altered by the increase in samples. Yet, when observing the USP17D1D2-2GKG assay with the K11-linked di-ubiquitin mimetic, both assays used identical timepoints, however the EC50 remained within the same range as USP2CD/ K6, a possible reason for the extreme EC50 value is due to the regression line being unstable because of density values which were deemed anomalous, possibly the 20-minute band density, which does not decrease enough according to the regression line.

A yet another improvement that can be implemented is using the enzyme at a lower concentration immediately, 1.5μM although presents with clear and rapid results, the curve on the associated densitometric data is too steep, hence the EC50 could also be perceived as too rapid, and as a result the EC50 is greater than expected. Gel cleavage assays which use the enzyme at a lower concentration, such as 100-500nM present as more realistic, evident in fig. 29 where a range of concentrations were used to provide multiple densitometric graphs with varying EC50 values, however a parameter which is to be considered depending on the choice of enzyme, is potency, as the substrates were already known to be cleaved by the enzyme, such as USP17D1D2-2GKG and the K11-linked di-ubiquitin mimetic, yet if an enzyme is chosen and the activity on a desired substrate is unknown, a higher concentration could be beneficial to determine the potency.

7.4 Structural Determination

Crystallography is the study of determining the 3D crystalline structure of a protein of choice, this can be achieved using X-Ray crystallography whereby a high concentration purified sample is crystallized- these results can be viewed using an X-Ray beam, whereby electrons which strike a copper anode, causing multiple wavelengths to be released- causing the formation of the beam[88]. One major disadvantage in using X-Ray crystallography is that the sample needs to match the requirements necessary to be screened, for instance the crystal lattice which is formed needs to allow the binding site to not be obstructed by other crystal lattices[89].

Another widely used method of structural determination is through the use of cryogenic electron microscopy (cryo-EM)- which uses a transmission electron microscopy to view samples at -183°C which allows vitreous ice to be formed[90]. Throughout the time of development, cryo-EM technology was able to improve in resolution imaging and overall performance with enhancements being made for the freezing technique and sample preparation[91].

Cryo-EM has multiple advantages over X-Ray crystallography such as samples not needing to be in crystal form to be analysed, capturing the protein in multiple conformations, and its main asset is to analyse proteins which have a higher molecular weight, which in the case of USP17D1D2-2GKG having a molecular weight of 51kDa- cryo-EM would be more beneficial[92]. However, cryo-EM used to be not suitable for smaller protein complexes, this would include the

K6 mimetic which has a molecular weight of <50kDa, a homogenous sample preparation is a necessity as protein flexibility will distort imaging. Finally 2Ä is the typical resolution requirement for pharmaceutical research, although initially cryo-EM had a limit of 3Ä, recently manufactured cryo-EM technology is capable of reaching as little as 1.2Ä, therefore the consistently evolving technology is able to mitigate the initial limitations[93].

7.5 Fluorescence Polarization

As demonstrated by fig. 21, USP17D1D2-2GKG was tested against the ISO-MIM di-ubiquitin, a moiety with an engineered isopeptide bond and a fluorescein maleimide fluorescent tag. If the enzyme were to inefficiently or only partially bind to the ISO-MIM, there would be little to no improvement regardless of the concentration of the USP, however as can be seen, the activity of the enzyme against the ISO-MIM increases dramatically, while also following a linear pattern when the concentration is lowered by 10x. Moreover, the optimal concentration of USP17 was discovered to be at 350nM, as the FP difference continues to increase at a linear, steady rate despite reaching the one-hour limit.

USP17D1D2-2GKG was also tested against the mono-ubiquitin probe, a singular ubiquitin moiety with the fluorescein maleimide bound to the allosteric site. As USP17D1D2-2GKG, according to the results shown in fig. 21, was able to efficiently cleave the ISO-MIM at various concentrations, therefore the protocol was repeated using the mono-ubiquitin probe. However, looking at the scatter graph (fig. 22), the results were mostly random, as there was no discernible correlation between the concentration of enzyme used and the fluorescence polarization. A considerable disadvantage of fluorescein maleimide is bulkiness, when bound to a substrate, especially a smaller moiety such as ubiquitin, the difference between the tumbling speed of the bound and liberated forms of fluorescein is less distinguishable, potentially causing the random results present in fig. 22.

PR-619 is a small molecule inhibitor, with a diverse inhibiting profile, therefore inhibiting the catalytic triad present on USP17, as presented in fig. 23, the representative graph for the calculated IC50 value would be graph 1, due to the sheer fact that none of the values from the table were deleted and all were incorporated into the curve, giving an overall clearer impression, therefore the calculated IC50 value for USP17D1D2-2GKG, specifically in the presence of PR-619, is 1.549μM, to improve on the IC50 curve and fluorescence polarization assays overall, would be to ensure the aliquoting of substances in the wells is concise and accurate, to mitigate the likelihood of air bubbles becoming a hinderance to the resulting data analysis.

8. Conclusion

From the speculated hypothesis and the data generated, USP17D1D2-2GKG was deemed as a protein with selective cleaving of multiple di-ubiquitin substrates, specifically GG, K6, and K11 with the gel cleavage assay results confirming the enzyme is able to produce monoubiquitin as rapidly as 5 minutes, with linear di-ubiquitin not being able to be cleaved. This issue of the cleaving not being successful was hypothesized to be remedied with mutagenesis of the aspartic acid present at the 332 position into a glycine, as the binding cannot occur in the wildtype due to potential steric hindrance, causing USP17 to inefficiently bind with the ubiquitin, hence its incapability of being cleaved, as although aspartate has low steric hindrance individually, glycine experiences none, potentially improving the binding domain to accommodate linear di-ubiquitin.

Following gel cleavage assays, between both protein constructs, confirmed that the steric hindrance of the catalytic domain present on USP17D1D2-2GKG was too high to sufficiently bind to linear di-ubiquitin, however lowering that steric strain with the introduction of a smaller amino acid with glycine, caused a minor improvement in the cleaving possibilities of linear di-ubiquitin.

USP17 was wrongly described as a protein which inefficiently binds with K6, as the gel cleavage assays and fluorescence polarization assays have shown, USP17 is able to cleave the K6-linked bond with ease and haste, as monoubiquitin was able to appear within five minutes of being introduced, and with the fluorescence polarization results, 1μM of enzyme cleaved the fluorescent tag present on the K6 almost immediately upon introduction, therefore out of all the substrates tested in this study, K6 would be the second most efficient substrate to bind to USP17, only behind K11. It is to be noted however, that the K6 used in the procedures was a mimetic, the purpose being to increase stability and to limit the risk of denaturation, therefore any interactions that could occur in vivo are not considered, such as the protein behaving differently as was demonstrated, therefore literature that referenceconsistently butSP17 is inefficient at cleaving K6-linked di-ubiquitin could potentially be correct if a mimetic was not used, however since the structure of the protein remains intact bar a few PTMs, the conclusion would be that USP17 not only cleaves K6 consistently, but also is selective.

Regarding the fluorescence polarization assays, USP2CD was used for one assay due to its nature of being a control to ensure the protocol was being accomplished correctly, with the scatter graph (fig. 11) proving that expression, purification, and storage of the protein was successful, along with proving that the fluorescent di-ubiquitin probe was functioning as intended , and therefore the protocols can be repeated on USP17D1D2-2GKG and USP17D1D2-D332G2GKG.

As the crystallization failed to produce any meaningful crystal samples for analysis, only testing on the Morpheus screen was deemed inefficient, therefore for any future work that can be conducted, the ligand-friendly screen is a respectable alternative, as the slight alkaline conditions that the screen contains allows the protein to be more accommodated and active, increasing the likelihood of crystal formation. JCSG-plus screen can also be a consideration as the various conditions that are present within one screen, hypothetically would allow USP17 to remain active over an extended period, causing crystallization to occur. Moreover, the protocols conducted on the Morpheus screens should have been repeated at least two more times to reduce the possibility of errors being made with the gel filtration process, the storage of the screens and proteins, or crystallisation being fully accomplished within a singular day rather than the protein or complexes being stored at -80°c overnight.

Regarding any future work that can be accomplished to further the results gathered from this project, crystallization studies should be prioritized, due to the lack of results presented by the screens used.

9. Acknowledgments

I would like to personally thank the biodiscovery institute, for providing the necessary equipment and laboratorial space which was necessary to the completion of this project. I would also like to extend my gratitude towards both of the supervisors, Ingrid Dreveny and Jonas Emsley, for supporting me physically and mentally throughout the whole process. I would also like to thank Reyes Ruiz Campoy, Mostafa Mohamed, and Krzysztof Raszpla, along with all the

other Ph.D. students/ researchers that were a part of the C84 laboratory, as all of them have eased me into the degree and the overall structure of the lab, along with any problems I may have faced. Additionally I would like to give special thanks to Krzysztof Raszpla and Jiatong Zhang for providing the K6 mimetic, GG, linear, and K11 di-ubiquitin compounds. I would also like to thank Lee Hibbett for managing the lab excellently along with giving the initial induction to the lab equipment along with inducting me to the more complex laboratory equipment further along my degree.

10. Bibliography

[1]Damgaard, R.B. The ubiquitin system: from cell signalling to disease biology and new therapeutic opportunities. *Cell Death Differ* **28**, 423–426 (2021).

[2] Stewart MD, Ritterhoff T, Klevit RE, Brzovic PS. E2 enzymes: more than just middle men. Cell Res. 2016 Apr;26(4):423-40. doi: 10.1038/cr.2016.35. Epub 2016 Mar 22. PMID: 27002219; PMCID: PMC4822130.

[3] Yang Q, Zhao J, Chen D, Wang Y. E3 ubiquitin ligases: styles, structures and functions. Mol Biomed. 2021 Jul 30;2(1):23. doi: 10.1186/s43556-021-00043-2. PMID: 35006464; PMCID: PMC8607428.

[4] Medvar B, Raghuram V, Pisitkun T, Sarkar A, Knepper MA. Comprehensive database of human E3 ubiquitin ligases: application to aquaporin-2 regulation. Physiol Genomics. 2016 Jul 1;48(7):502-12. doi: 10.1152/physiolgenomics.00031.2016. Epub 2016 May 13. Erratum in: Physiol Genomics. 2025 Aug 1;57(8):472. doi: 10.1152/physiolgenomics.00031.2016_COR. PMID: 27199454; PMCID: PMC4967219.

[5] Li Y, Okatsu K, Fukai S, Sato Y. Structural basis for specific recognition of K6-linked polyubiquitin chains by the TAB2 NZF domain. Biophys J. 2021 Aug 17;120(16):3355-3362. doi: 10.1016/j.bpj.2021.06.037. Epub 2021 Jul 7. PMID: 34242591; PMCID: PMC8391078.

[6] Bhat SA, Vasi Z, Adhikari R, Gudur A, Ali A, Jiang L, Ferguson R, Liang D, Kuchay S. Ubiquitin proteasome system in immune regulation and therapeutics. Curr Opin Pharmacol. 2022 Dec;67:102310. doi: 10.1016/j.coph.2022.102310. Epub 2022 Oct 23. PMID: 36288660; PMCID: PMC10163937.

[7] Wickliffe KE, Williamson A, Meyer HJ, Kelly A, Rape M. K11-linked ubiquitin chains as novel regulators of cell division. Trends Cell Biol. 2011 Nov;21(11):656-63. doi: 10.1016/j.tcb.2011.08.008. Epub 2011 Oct 4. PMID: 21978762; PMCID: PMC3205209.

[8] Ishii T, Funakoshi M, Kobayashi H, Sekiguchi T. Yeast Irc22 Is a Novel Dsk2-Interacting Protein that Is Involved in Salt Tolerance. Cells. 2014 Mar 27;3(2):180-98. doi: 10.3390/cells3020180. PMID: 24709957; PMCID: PMC4092868.

[9] Fishbain, S., Prakash, S., Herrig, A. *et al.* Rad23 escapes degradation because it lacks a proteasome initiation region. *Nat Commun* **2**, 192 (2011). https://doi.org/10.1038/ncomms1194.

[10] Matsumoto ML, Wickliffe KE, Dong KC, Yu C, Bosanac I, Bustos D, Phu L, Kirkpatrick DS, Hymowitz SG, Rape M, Kelley RF, Dixit VM. K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. Mol Cell. 2010 Aug 13;39(3):477-84. doi: 10.1016/j.molcel.2010.07.001. Epub 2010 Jul 22. PMID: 20655260.

[11] Faggiano S., Alfano C., Pastore A., The missing links to link ubiquitin: Methods for the enzymatic production of polyubiquitin chains, Analytical Biochemistry, Volume 492, 2016, Pages 82-90, ISSN 0003-2697, https://doi.org/10.1016/j.ab.2015.09.013. (https://www.sciencedirect.com/science/article/pii/S0003269715004431)

[12] Kirisako T, Kamei K, Murata S, Kato M, Fukumoto H, Kanie M, et al. A ubiquitin ligase complex assembles linear polyubiquitin chains. EMBO J. 2006;25:4877–87.

[13] Tokunaga F, Sakata S, Saeki Y, Satomi Y, Kirisako T, Kamei K, et al. Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation. Nat Cell Biol. 2009;11:123–32.

[14] Oeckinghaus A, Ghosh S . The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol* 2009; **1**: a000034.

[15] Fulzele A, Bennett EJ. Ubiquitin diGLY Proteomics as an Approach to Identify and Quantify the Ubiquitin-Modified Proteome. Methods Mol Biol. 2018;1844:363-384. doi: 10.1007/978-1-4939-8706-1_23. PMID: 30242721; PMCID: PMC6791129.

[16] Goldknopf IL, Busch H (1977) Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. Proceedings of the National Academy of Sciences 74 (3):864–868.

[17] T.T. Huang, S.M. Nijman, K.D. Mirchandani, P.J. Galardy, M.A. Cohn, W. Haas, *et al.*, Regulation of monoubiquitinated PCNA by DUB autocleavage, Nat. Cell Biol., 8 (2006), pp. 339-347

[18] Alexei F. et al., Proteasome inhibitors: from research tools to drug candidates, Kisselev, , Chemistry & Biology, Volume 8, Issue 8, 739 – 758.

[19] Schauer NJ, Magin RS, Liu X, Doherty LM, Buhrlage SJ. Advances in Discovering Deubiquitinating Enzyme (DUB) Inhibitors. J Med Chem. 2020 Mar 26;63(6):2731-2750. doi: 10.1021/acs.jmedchem.9b01138. Epub 2019 Nov 20. PMID: 31682427.

[20] Bonacci T, Emanuele MJ. Dissenting degradation: Deubiquitinases in cell cycle and cancer. Semin Cancer Biol. 2020 Dec;67(Pt 2):145-158. doi: 10.1016/j.semcancer.2020.03.008. Epub 2020 Mar 20. PMID: 32201366; PMCID: PMC7502435.

[21] Zhang F, Zhao Y, Sun Y. USP2 is an SKP2 deubiquitylase that stabilizes both SKP2 and its substrates. J Biol Chem. 2021 Oct;297(4):101109. doi: 10.1016/j.jbc.2021.101109. Epub 2021 Aug 21. PMID: 34425107; PMCID: PMC8446802.

[22] Newton K, Matsumoto ML, Wertz IE, Kirkpatrick DS, Lill JR, Tan J, Dugger D, Gordon N, Sidhu SS, Fellouse FA, Komuves L, French DM, Ferrando RE, Lam C, Compaan D, Yu C, Bosanac I, Hymowitz SG, Kelley RF, Dixit VM. Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. Cell. 2008 Aug 22;134(4):668-78. doi: 10.1016/j.cell.2008.07.039. PMID: 18724939.

[23] Bozza WP, Liang Q, Gong P, Zhuang Z. Transient kinetic analysis of USP2-catalyzed deubiquitination reveals a conformational rearrangement in the K48-linked diubiquitin substrate. Biochemistry. 2012 Dec 18;51(50):10075-86. doi: 10.1021/bi3009104. Epub 2012 Dec 4. PMID: 23211065; PMCID: PMC3617563.

[24] Emmerich CH, Cohen P. Optimising methods for the preservation, capture and identification of ubiquitin chains and ubiquitylated proteins by immunoblotting. Biochem Biophys Res Commun. 2015 Oct 9;466(1):1-14. doi: 10.1016/j.bbrc.2015.08.109. Epub 2015 Aug 29. PMID: 26325464; PMCID: PMC4709362.

[25] Sacco JJ, Coulson JM, Clague MJ, Urbé S. Emerging roles of deubiquitinases in cancer-associated pathways. IUBMB Life. 2010 Feb;62(2):140-57. doi: 10.1002/iub.300. PMID: 20073038; PMCID: PMC7165618.

[26] Kuang Z, Liu X, Zhang N, Dong J, Sun C, Yin M, Wang Y, Liu L, Xiao D, Zhou X, Feng Y, Song D, Deng H. USP2 promotes tumor immune evasion via deubiquitination and stabilization of PD-L1. Cell Death Differ. 2023 Oct;30(10):2249-2264. doi: 10.1038/s41418-023-01219-9. Epub 2023 Sep 5. PMID: 37670038; PMCID: PMC10589324.

[27] de la Vega M, Kelvin AA, Dunican DJ, McFarlane C, Burrows JF, Jaworski J, Stevenson NJ, Dib K, Rappoport JZ, Scott CJ, Long A, Johnston JA. The deubiquitinating enzyme USP17 is essential for GTPase subcellular localization and cell motility. Nat Commun. 2011 Mar 29;2:259. doi: 10.1038/ncomms1243. PMID: 21448158; PMCID: PMC3072070.

[28] Ducker C, Shaw PE. USP17-mediated de-ubiquitination and cancer: Clients cluster around the cell cycle. Int J Biochem Cell Biol. 2021 Jan;130:105886. doi: 10.1016/j.biocel.2020.105886. Epub 2020 Nov 20. PMID: 33227393.

[29] de la Vega M, Kelvin AA, Dunican DJ, McFarlane C, Burrows JF, Jaworski J, Stevenson NJ, Dib K, Rappoport JZ, Scott CJ, Long A, Johnston JA: The deubiquitinating enzyme USP17 is essential for GTPase subcellular localization and cell motility. Nat Commun 2011, 2:259.

[30] Nicole M. Hjortland, Andrew D. Mesecar, 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,Volume 612, 2016, Pages 35-45, ISSN 0003-9861, https://doi.org/10.1016/j.abb.2016.10.008.

[31] Hjortland NM, Mesecar AD. Steady-state kinetic studies reveal that the anti-cancer target Ubiquitin-Specific Protease 17 (USP17) is a highly efficient deubiquitinating enzyme. Arch

Biochem Biophys. 2016 Dec 15;612:35-45. doi: 10.1016/j.abb.2016.10.008. Epub 2016 Oct 15. PMID: 27756680; PMCID: PMC5868344.

[32] Chen X, Wang C, Liao K, Zhou S, Cao L, Chen J, Xu C, Lin Y. USP17 Suppresses Tumorigenesis and Tumor Growth through Deubiquitinating AEP. Int J Biol Sci. 2019 Jan 29;15(4):738-748. doi: 10.7150/ijbs.30106. PMID: 30906206; PMCID: PMC6429017.

[33] McFarlane C, McFarlane S, Paul I, Arthur K, Scheaff M, Kerr K, Stevenson M, Fennell DA, Johnston JA. The deubiquitinating enzyme USP17 is associated with non-small cell lung cancer (NSCLC) recurrence and metastasis. Oncotarget. 2013 Oct;4(10):1836-43. doi: 10.18632/oncotarget.1282. PMID: 24123619; PMCID: PMC3858568.

[34] Song C, Liu W, Li J. USP17 is upregulated in osteosarcoma and promotes cell proliferation, metastasis, and epithelial-mesenchymal transition through stabilizing SMAD4. Tumour Biol. 2017 Jul;39(7):1010428317717138. doi: 10.1177/1010428317717138. PMID: 28670958.

[35] Ramakrishna S, Suresh B, Bae SM, Ahn WS, Lim KH, Baek KH. Hyaluronan binding motifs of USP17 and SDS3 exhibit anti-tumor activity. PLoS One. 2012;7(5):e37772. doi: 10.1371/journal.pone.0037772. Epub 2012 May 25. PMID: 22662218; PMCID: PMC3360616..

[36] Davis MI, Pragani R, Fox JT, Shen M, Parmar K, Gaudiano EF, Liu L, Tanega C, McGee L, Hall MD, McKnight C, Shinn P, Nelson H, Chattopadhyay D, D'Andrea AD, Auld DS, DeLucas LJ, Li Z, Boxer MB, Simeonov A. Small Molecule Inhibition of the Ubiquitin-specific Protease USP2 Accelerates cyclin D1 Degradation and Leads to Cell Cycle Arrest in Colorectal Cancer and Mantle Cell Lymphoma Models. J Biol Chem. 2016 Nov 18;291(47):24628-24640. doi: 10.1074/jbc.M116.738567. Epub 2016 Sep 28. PMID: 27681596; PMCID: PMC5114414.

[37] Tian Y, Liu K, Wu D, Wu L, Xu Q, Wei W, Li Z, Du Q, Bian J. The discovery of potent USP2/USP8 dual-target inhibitors for the treatment of breast cancer via structure guided optimization of ML364. Eur J Med Chem. 2024 Mar 15;268:116275. doi: 10.1016/j.ejmech.2024.116275. Epub 2024 Feb 24. PMID: 38452725.

[38] Dang F, Bai L, Dong J, Hu X, Wang J, Paulo JA, Xiong Y, Liang X, Sun Y, Chen Y, Guo M, Wang X, Huang Z, Inuzuka H, Chen L, Chu C, Liu J, Zhang T, Rezaeian AH, Liu J, Kaniskan HÜ, Zhong B, Zhang J, Letko M, Jin J, Lan K, Wei W. USP2 inhibition prevents infection with ACE2-dependent coronaviruses in vitro and is protective against SARS-CoV-2 in mice. Sci Transl Med. 2023 Dec 6;15(725):eadh7668. doi: 10.1126/scitranslmed.adh7668. Epub 2023 Dec 6. PMID: 38055802; PMCID: PMC10787358.

[39] Liu T, Yu J, Deng M, Yin Y, Zhang H, Luo K, Qin B, Li Y, Wu C, Ren T, Han Y, Yin P, Kim J, Lee S, Lin J, Zhang L, Zhang J, Nowsheen S, Wang L, Boughey J, Goetz MP, Yuan J, Lou Z. CDK4/6-dependent activation of DUB3 regulates cancer metastasis through SNAIL1. Nat Commun. 2017 Jan 9;8:13923. doi: 10.1038/ncomms13923. PMID: 28067227; PMCID: PMC5228031.

[40] Kaufhold S, Bonavida B. Central role of Snail1 in the regulation of EMT and resistance in cancer: a target for therapeutic intervention. J Exp Clin Cancer Res. 2014 Aug 2;33(1):62. doi: 10.1186/s13046-014-0062-0. PMID: 25084828; PMCID: PMC4237825.

[41] Peiró S, Escrivà M, Puig I, Barberà MJ, Dave N, Herranz N, Larriba MJ, Takkunen M, Francí C, Muñoz A, Virtanen I, Baulida J, García de Herreros A. Snail1 transcriptional repressor binds to its own promoter and controls its expression. Nucleic Acids Res. 2006 Apr 14;34(7):2077-84. doi: 10.1093/nar/gkl141. Erratum in: Nucleic Acids Res. 2026 Mar 19;54(6):gkag313. doi: 10.1093/nar/gkag313. PMID: 16617148; PMCID: PMC1440880.

[42] Chen Y, Gao F, He Y, Liu M, Quan Y, Zhang P. DUB3 is a MAGEA3 deubiquitinase and a potential therapeutic target in hepatocellular carcinoma. iScience. 2024 Feb 8;27(3):109181. doi: 10.1016/j.isci.2024.109181. PMID: 38414853; PMCID: PMC10897913.

[43] Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol Cell. 2005 Aug 19;19(4):523-34. doi: 10.1016/j.molcel.2005.06.027. PMID: 16109376.

[44] Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, Young RA. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013 Apr 11;153(2):320-34. doi: 10.1016/j.cell.2013.03.036. PMID: 23582323; PMCID: PMC3760967.

[45] Ross NT, Katt WP, Hamilton AD. Synthetic mimetics of protein secondary structure domains. Philos Trans A Math Phys Eng Sci. 2010 Mar 13;368(1914):989-1008. doi: 10.1098/rsta.2009.0210. PMID: 20123744; PMCID: PMC3263801.

[46] Costa F, Riziotis I, Andreeva A, Kalwan D, de Jong J, Hinchliffe P, Parmeggiani F, Race PR, Burston SG, Bateman A, Barringer R. A global survey of intramolecular isopeptide bonds. Protein Sci. 2025 Dec;34(12):e70342. doi: 10.1002/pro.70342. PMID: 41236442; PMCID: PMC12617251.

[47] Andersen KA, Raines RT. Creating site-specific isopeptide linkages between proteins with the traceless Staudinger ligation. Methods Mol Biol. 2015;1248:55-65. doi: 10.1007/978-1-4939-2020-4_4. PMID: 25616325; PMCID: PMC4307611.)

[48]Ghosh D, Kohli AG, Moser F, Endy D, Belcher AM. Refactored M13 bacteriophage as a platform for tumor cell imaging and drug delivery. ACS Synth Biol. 2012 Dec 21;1(12):576-582. doi: 10.1021/sb300052u. Epub 2012 Sep 24. PMID: 23656279; PMCID: PMC3905571.

[49] Rosano GL, Ceccarelli EA. Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol. 2014 Apr 17;5:172. doi: 10.3389/fmicb.2014.00172. PMID: 24860555; PMCID: PMC4029002.

[50]Francis DM, Page R. Strategies to optimize protein expression in E. coli. Curr Protoc Protein Sci. 2010 Aug;Chapter 5(1):5.24.1-5.24.29. doi: 10.1002/0471140864.ps0524s61. PMID: 20814932; PMCID: PMC7162232.

[51]Piorino F, Styczynski MP. Complex Dependence of *Escherichia coli*-based Cell-Free Expression on Sonication Energy During Lysis. ACS Synth Biol. 2023 Oct 20;12(10):3131-3136. doi: 10.1021/acssynbio.3c00312. Epub 2023 Sep 19. PMID: 37725792; PMCID: PMC10594866.

[52] Ley C, Holtmann D, Mangold KM, Schrader J. Immobilization of histidine-tagged proteins on electrodes. Colloids Surf B Biointerfaces. 2011 Dec 1;88(2):539-51. doi: 10.1016/j.colsurfb.2011.07.044. Epub 2011 Jul 26. PMID: 21840689.

[53]Yildiz AN, Kasapbasi EE, Yurtsever Z, Yurtsever M. DFT Study on the His-Tag Binding Affinity of Metal Ions in Modeled Hexacationic Metal Complexes. ACS Omega. 2024 Nov 8;9(46):45920-45925. doi: 10.1021/acsomega.4c05422. PMID: 39583711; PMCID: PMC11579778.

[54]Du M, Hou Z, Liu L, Xuan Y, Chen X, Fan L, Li Z, Xu B. [1]Progress, applications, challenges and prospects of protein purification technology. Front Bioeng Biotechnol. 2022 Dec 6;10:1028691. doi: 10.3389/fbioe.2022.1028691. PMID: 36561042; PMCID: PMC9763899.

[55]Bornhorst JA, Falke JJ. Purification of proteins using polyhistidine affinity tags. Methods Enzymol. 2000;326:245-54. doi: 10.1016/s0076-6879(00)26058-8. PMID: 11036646; PMCID: PMC2909483.

[56] Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970 Aug 15;227(5259):680-5. doi: 10.1038/227680a0. PMID: 5432063.

[57] Lee A, Tang SK, Mace CR, Whitesides GM. Denaturation of proteins by SDS and tetraalkylammonium dodecyl sulfates. Langmuir. 2011 Sep 20;27(18):11560-74. doi: 10.1021/la201832d. Epub 2011 Aug 23. PMID: 21834533; PMCID: PMC3172379.

[58] Na N, Ouyang J. TEMED-enhanced photoluminescent imaging of human serum proteins by quantum dots after PAGE. Methods Mol Biol. 2012;869:511-20. doi: 10.1007/978-1-61779-821-4_45. PMID: 22585516.

[59] Liu X, Sun Q, Zhang Y, Feng Y, Su X. Rapid RAFT Polymerization of Acrylamide with High Conversion. Molecules. 2023 Mar 13;28(6):2588. doi: 10.3390/molecules28062588. PMID: 36985559; PMCID: PMC10057598.

[60] Packirisamy V, Pandurangan P. Polyacrylamide gel electrophoresis: a versatile tool for the separation of nanoclusters. Biotechniques. 2023 Jan;74(1):51-62. doi: 10.2144/btn-2022-0086. Epub 2022 Dec 14. PMID: 36517970; PMCID: PMC9887536.

[61]Braun O, Coquery C, Kieffer J, Blondel F, Favero C, Besset C, Mesnager J, Voelker F, Delorme C, Matioszek D. Spotlight on the Life Cycle of Acrylamide-Based Polymers Supporting Reductions in Environmental Footprint: Review and Recent Advances. Molecules. 2021 Dec 22;27(1):42. doi: 10.3390/molecules27010042. PMID: 35011281; PMCID: PMC8746853.

[62] Wang D, Wang C, Wang L, Chen Y. A comprehensive review in improving delivery of small-molecule chemotherapeutic agents overcoming the blood-brain/brain tumor barriers for glioblastoma treatment. Drug Deliv. 2019 Dec;26(1):551-565. doi: 10.1080/10717544.2019.1616235. PMID: 31928355; PMCID: PMC6534214.

[63] Ma YS, Xin R, Yang XL, Shi Y, Zhang DD, Wang HM, Wang PY, Liu JB, Chu KJ, Fu D. Paving the way for small-molecule drug discovery. Am J Transl Res. 2021 Mar 15;13(3):853-870. PMID: 33841626; PMCID: PMC8014367.

[64] Zhao Z, Ukidve A, Kim J, Mitragotri S. Targeting Strategies for Tissue-Specific Drug Delivery. Cell. 2020 Apr 2;181(1):151-167. doi: 10.1016/j.cell.2020.02.001. PMID: 32243788.

[65] Cowell IG, Ling EM, Swan RL, Brooks MLW, Austin CA. The Deubiquitinating Enzyme Inhibitor PR-619 is a Potent DNA Topoisomerase II Poison. Mol Pharmacol. 2019 Nov;96(5):562-572. doi: 10.1124/mol.119.117390. Epub 2019 Sep 12. PMID: 31515282; PMCID: PMC6776009.

[66] Wang L, Li M, Sha B, Hu X, Sun Y, Zhu M, Xu Y, Li P, Wang Y, Guo Y, Li J, Shi J, Li P, Hu T, Chen P. Inhibition of deubiquitination by PR-619 induces apoptosis and autophagy via ubi-protein aggregation-activated ER stress in oesophageal squamous cell carcinoma. Cell Prolif. 2021 Jan;54(1):e12919. doi: 10.1111/cpr.12919. Epub 2020 Oct 31. PMID: 33129231; PMCID: PMC7791184.

[67] Seiberlich V, Goldbaum O, Zhukareva V, Richter-Landsberg C. The small molecule inhibitor PR-619 of deubiquitinating enzymes affects the microtubule network and causes protein aggregate formation in neural cells: implications for neurodegenerative diseases. Biochim Biophys Acta. 2012 Nov;1823(11):2057-68. doi: 10.1016/j.bbamcr.2012.04.011. Epub 2012 Apr 28. PMID: 22565157.

[68] Hsu FS, Lin WC, Kuo KL, Chiu YL, Hsu CH, Liao SM, Dong JR, Liu SH, Chang SC, Yang SP, Chen YT, Chang RJ, Huang KH. PR-619, a General Inhibitor of Deubiquitylating Enzymes, Diminishes Cisplatin Resistance in Urothelial Carcinoma Cells through the Suppression of c-Myc: An In Vitro and In Vivo Study. Int J Mol Sci. 2021 Oct 28;22(21):11706. doi: 10.3390/ijms222111706. PMID: 34769137; PMCID: PMC8584183.

[69] Kuo KL, Liu SH, Lin WC, Chow PM, Chang YW, Yang SP, Shi CS, Hsu CH, Liao SM, Chang HC, Huang KH. The Deubiquitinating Enzyme Inhibitor PR-619 Enhances the Cytotoxicity of Cisplatin via the Suppression of Anti-Apoptotic Bcl-2 Protein: In Vitro and In Vivo Study. Cells. 2019 Oct 17;8(10):1268. doi: 10.3390/cells8101268. PMID: 31627336; PMCID: PMC6830310.

[70]Lin WC, Chiu YL, Kuo KL, Chow PM, Hsu CH, Liao SM, Dong JR, Chang SC, Liu SH, Liu TJ, Hsu FS, Wang KC, Lin YC, Chang CC, Huang KY. Anti-tumor effects of deubiquitinating enzyme inhibitor PR-619 in human chondrosarcoma through reduced cell proliferation and endoplasmic reticulum stress-related apoptosis. Am J Cancer Res. 2023 Jul 15;13(7):3055-3066. PMID: 37559983; PMCID: PMC10408468.

[71] Chen L, Chen X, Yang X, He C, Wang M, Xi P, Gao J. Advances of super-resolution fluorescence polarization microscopy and its applications in life sciences. Comput Struct Biotechnol J. 2020 Jun 26;18:2209-2216. doi: 10.1016/j.csbj.2020.06.038. PMID: 32952935; PMCID: PMC7476067.

[72] Marshall J, Johnsen S. Fluorescence as a means of colour signal enhancement. Philos Trans R Soc Lond B Biol Sci. 2017 Jul 5;372(1724):20160335. doi: 10.1098/rstb.2016.0335. PMID: 28533452; PMCID: PMC5444056.

[73] Maziarz M, Garcia-Marcos M. Fluorescence polarization assays to measure interactions between Gα subunits of heterotrimeric G proteins and regulatory motifs. Methods Cell Biol. 2017;142:133-143. doi: 10.1016/bs.mcb.2017.07.007. Epub 2017 Sep 11. PMID: 28964332; PMCID: PMC5654624.[70]Wittig I, Schägger H. Proteomics. 2005;5:4338–4346. doi: 10.1002/pmic.200500081.

[74] Cohen AE, Moerner WE. Controlling Brownian motion of single protein molecules and single fluorophores in aqueous buffer. Opt Express. 2008 May 12;16(10):6941-56. doi: 10.1364/oe.16.006941. PMID: 18545398; PMCID: PMC2435051.

[75]Brown TA. Genomes. 2nd edition. Oxford: Wiley-Liss; 2002. Chapter 14, Mutation, Repair and Recombination. Available from: https://www.ncbi.nlm.nih.gov/books/NBK21114

[76]Durland J, Ahmadian-Moghadam H. Genetics, Mutagenesis. Treasure Island (FL): StatPearls Publishing; 2026 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560519/

[77]Khehra N, Padda IS, Zubair M. Polymerase Chain Reaction (PCR) Treasure Island (FL): StatPearls Publishing; 2026 Jan-.

[78]Garibyan L, Avashia N. Polymerase chain reaction. J Invest Dermatol. 2013 Mar;133(3):1-4. doi: 10.1038/jid.2013.1. PMID: 23399825; PMCID: PMC4102308.

[79]Bijak V, Szczygiel M, Lenkiewicz J, Gucwa M, Cooper DR, Murzyn K, Minor W. The current role and evolution of X-ray crystallography in drug discovery and development. Expert Opin Drug Discov. 2023 Jul-Dec;18(11):1221-1230. doi: 10.1080/17460441.2023.2246881. Epub 2023 Aug 17. PMID: 37592849; PMCID: PMC10620067.

[80] Maurer SK, Mayer MP, Ward SJ, Boudjema S, Halawa M, Zhang J, Caulton SG, Emsley J, Dreveny I. Ubiquitin-specific protease 11 structure in complex with an engineered substrate mimetic reveals a molecular feature for deubiquitination selectivity. J Biol Chem. 2023 Nov;299(11):105300. doi: 10.1016/j.jbc.2023.105300. Epub 2023 Sep 28. PMID: 37777157; PMCID: PMC10637973.

[81] Fraser JS, Merlie JP Jr, Echols N, Weisfield SR, Mignot T, Wemmer DE, Zusman DR, Alber T. An atypical receiver domain controls the dynamic polar localization of the Myxococcus xanthus social motility protein FrzS. Mol Microbiol. 2007 Jul;65(2):319-32. doi: 10.1111/j.1365-2958.2007.05785.x. Epub 2007 Jun 15. PMID: 17573816; PMCID: PMC1974792.

[82] https://sourcebioscience.com/genomics/rapid-sequencing/sanger-sequencing/plasmid-dna/

[83] Sigrun K. Maurer, Matthias P. Mayer, Stephanie J. Ward, Sana Boudjema, Mohamed Halawa, Jiatong Zhang, Simon G. Caulton, Jonas Emsley, Ingrid Dreveny, Ubiquitin-specific protease 11 structure in complex with an engineered substrate mimetic reveals a molecular feature for deubiquitination selectivity, Journal of Biological Chemistry, Volume 299, Issue 11, 2023, 105300, ISSN 0021-9258, https://doi.org/10.1016/j.jbc.2023.105300. (https://www.sciencedirect.com/science/article/pii/S0021925823023281)

[84] Wingfield P. Protein precipitation using ammonium sulfate. Curr Protoc Protein Sci. 2001 May;Appendix 3:Appendix 3F. doi: 10.1002/0471140864.psa03fs13. PMID: 18429073; PMCID: PMC4817497.

[85] Calvo JM, Matthews RG. The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. Microbiol Rev. 1994 Sep;58(3):466-90. doi: 10.1128/mr.58.3.466-490.1994. Erratum in: Microbiol Rev 1995 Jun;59(2):323. PMID: 7968922; PMCID: PMC372976.

[86] Scholz C, Eckert B, Hagn F, Schaarschmidt P, Balbach J, Schmid FX. SlyD proteins from different species exhibit high prolyl isomerase and chaperone activities. Biochemistry. 2006 Jan 10;45(1):20-33. doi: 10.1021/bi051922n. PMID: 16388577.

[87] Hottenrott S, Schumann T, Plückthun A, Fischer G, Rahfeld JU. The Escherichia coli SlyD is a metal ion-regulated peptidyl-prolyl cis/trans-isomerase. J Biol Chem. 1997 Jun 20;272(25):15697-701. doi: 10.1074/jbc.272.25.15697. PMID: 9188461.

[88]Smyth MS, Martin JH. x ray crystallography. Mol Pathol. 2000 Feb;53(1):8-14. doi: 10.1136/mp.53.1.8. PMID: 10884915; PMCID: PMC1186895.

[89]Davis AM, St-Gallay SA, Kleywegt GJ. Limitations and lessons in the use of X-ray structural information in drug design. Drug Discov Today. 2008 Oct;13(19-20):831-41. doi: 10.1016/j.drudis.2008.06.006. Epub 2008 Aug 27. PMID: 18617015; PMCID: PMC7185550.

[90]Milne JL, Borgnia MJ, Bartesaghi A, Tran EE, Earl LA, Schauder DM, Lengyel J, Pierson J, Patwardhan A, Subramaniam S. Cryo-electron microscopy--a primer for the non-microscopist. FEBS J. 2013 Jan;280(1):28-45. doi: 10.1111/febs.12078. Epub 2012 Dec 17. PMID: 23181775; PMCID: PMC3537914.

[91]Chiu W, Schmid MF, Pintilie GD, Lawson CL. Evolution of standardization and dissemination of cryo-EM structures and data jointly by the community, PDB, and EMDB. J Biol Chem. 2021 Jan-Jun;296:100560. doi: 10.1016/j.jbc.2021.100560. Epub 2021 Mar 18. PMID: 33744287; PMCID: PMC8050867.

[92] García-Nafría J, Tate CG. Structure determination of GPCRs: cryo-EM compared with X-ray crystallography. Biochem Soc Trans. 2021 Nov 1;49(5):2345-2355. doi: 10.1042/BST20210431. PMID: 34581758; PMCID: PMC8589417.

[93] Van Drie JH, Tong L. Cryo-EM as a powerful tool for drug discovery. Bioorg Med Chem Lett. 2020 Nov 15;30(22):127524. doi: 10.1016/j.bmcl.2020.127524. Epub 2020 Sep 2. PMID: 32890683; PMCID: PMC7467112.