

**INVESTIGATION OF CHEMICAL UBIQUITINATION  
OF PCNA AND MECHANISM, INHIBITION OF USP1/UAF1 &  
THE MOLECULAR RECOGNITION  
OF RNA BY THE PSEUDOURIDINE SYNTHASE RLUA**

by

Junjun Chen

A dissertation submitted to the Faculty of the University of Delaware in  
partial fulfillment of the requirements for the degree of Doctor of Philosophy in  
Chemistry and Biochemistry

Winter 2011

Copyright 2011 Junjun Chen  
All Rights Reserved

**INVESTIGATION OF CHEMICAL UBIQUITINATION  
OF PCNA AND MECHANISM, INHIBITION OF USP1/UAF1 &  
THE MOLECULAR RECOGNITION  
OF RNA BY THE PSEUDOURIDINE SYNTHASE RLUA**

by

Junjun Chen

Approved: \_\_\_\_\_  
Klaus H. Theopold, Ph.D.  
Chair of the Department of Chemistry and Biochemistry

Approved: \_\_\_\_\_  
George H. Watson, Ph.D.  
Dean of the College of Arts and Science

Approved: \_\_\_\_\_  
Charles G. Riordan, Ph.D.  
Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Zhihao Zhuang, Ph.D.  
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Eugene G. Mueller, Ph.D.  
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Joseph M. Fox, Ph.D.  
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Kristi L. Kiick, Ph.D.  
Member of dissertation committee

## ACKNOWLEDGMENTS

I would never have been able to finish my dissertation without the guidance of my two advisors, Dr. Zhihao Zhuang and Dr. Eugene Mueller. I'm so grateful for the time I have spent under both of your tutelage. Dr. Zhuang, thank you very much for taking me as your graduate student when I was not sure whether I could continue on with graduate school in Delaware. I benefited a lot from the discussion with you on my research and working with you in the lab. Your enthusiasm and persistency in science has invaluable influence on my development as a researcher. Also, thank you for always reminding me to develop good lab habits, which helped me a lot in increasing my work efficiency. I'm very lucky to get the first project "chemical ubiquitination of PCNA" from your lab. It gave me a chance to combine the knowledge I learned from Mueller group with what you taught me. Dr. Mueller, thank you for your excellent guidance, caring, patience and providing me with a good atmosphere for doing research during my first three years in graduate school. Your encouragement made me believe things will always work out in the end. Your immense knowledge and motivation for science have a great impact on my development as a scientist. I feel very lucky to work in your lab.

Next, I want to thank people both from my present and past group. Yongxing Ai, William Bozza, Mark Villamil, Jialiang Wang, Qin Liang, Yu Peng and Chris Weinacht from Zhuang group; Chappie Wright, Chris Hamilton, Caroline Fanslau, Tara Baney, Edd Miracco and Maggie McDonald from Mueller group. I learnt a lot research expertise from all of you and also had so much fun with you guys

every day. Particularly, I need to thank two people. Yonxing, thank you very much for your contribution in the chemical ubiquitination project and HTS project. Chappie, I appreciated all your patience and help in teaching me those basic biochemistry lab techniques when I first joined Mueller lab.

I also want to thank my committee members, Dr. Fox and Dr. Kiick. Dr. Fox, thank you for your generous permission to allow me to work in your lab to do the organic synthesis. Dr. Kiick, I truly appreciate your time and effort on discussing my project. I want to thank both of you for your encouragement, good advice and support over the years!

I also wish to thank Dr. Thorpe who provided helpful thinking in our chemical ubiquitination project!

I would like to say “thank you” to John Dykins, who patiently taught me how to use the ESI instrument step by step.

Most importantly, I would like to thank my parents and my elder brother. They are always supporting me and encouraging me with their best wishes. Finally, I wish to thank my loving husband, Shangjin Sun, who always there cheering me up and stood by me through the good times and bad.

## TABLE OF CONTENTS

LIST OF TABLES.....	x
LIST OF FIGURES .....	xii
ABSTRACT .....	xxvi

### Chapter

1.	The Molecular Recognition of RNA by the Pseudouridine Synthase	
	RluA .....	1
1.1	Introduction .....	1
1.2	Results and Discussion .....	10
1.2.1	Preliminary kinetic characterization of substrates ASLs with wild-type RluA .....	10
1.2.2	Assessing the fraction of ASL substrates in an incompetent conformation by chromatography analysis .....	14
1.2.3	Full kinetic characterization of ASLs via HPLC analysis .....	15
1.2.4	Investigation of the reactivity between RluA and two altered [ <sup>5</sup> F]ASLs .....	21
1.3	Materials and methods.....	29
1.3.1	General .....	29
1.3.2	Overexpression and Purification of RluA .....	30
1.3.3	Deprotection and Purification of 2'-ACE RNA Oligonucleotides.....	31
1.3.4	Preliminary kinetic characterization for wild-type RluA .....	32
1.3.5	Kinetic assays .....	34
1.3.6	Band shift analysis.....	34
1.3.7	Intact and S1 digestion HPLC analysis .....	35
	References .....	36
2.	Chemically ubiquitinated PCNA as a probe for eukaryotic translesion DNA synthesis .....	40
2.1	Introduction .....	40
2.1.1	Ubiquitin is a versatile signaling molecule .....	40
2.1.2	Ubiquitination and SUMOylation of PCNA .....	42
2.1.3	Translesion synthesis (TLS) .....	44
2.1.4	Chemical ubiquitination .....	46

2.1.5	New approaches for PCNA ubiquitination is imperative .....	53
2.2	Results .....	53
2.2.1	Chemical ubiquitination of PCNA .....	53
2.2.2	Disulfide linkage characterization .....	55
2.2.3	Confirm position-specific monoubiquitination by tryptic digestion and MS/MS .....	56
2.2.4	Assess the chemical stability of the disulfide linkage .....	59
2.2.5	Chemical ubiquitination on PCNA position 127, 44 and 107 .....	60
2.2.6	Chemical SUMOylation on PCNA position 164 and 127 .....	66
2.3	Discussion .....	72
2.4	Materials and Methods .....	79
2.4.1	Preparation of ubiquitin-cysteamine (Ub-SH) .....	79
2.4.2	Preparation of SUMO-cysteamine .....	80
2.4.3	Preparation of PCNA mutants .....	80
2.4.4	Chemical ubiquitination of PCNA .....	81
2.4.5	Determine the molecular weight of ubiquitinated PCNA by ESI-MS .....	82
2.4.6	Tryptic digestion of the chemically ubiquitinated PCNA .....	83
2.4.7	Complete cleavage of the disulfide linkage with high concentration of DTT .....	83
2.4.8	Sensitivity of chemically ubiquitinated PCNA to reducing reagents .....	84
2.4.9	Chemical SUMOylation of PCNA .....	84
2.4.10	Determine the content of free thiol in unmodified and ubiquitinated PCNA .....	85
	References .....	86
3.	A novel regulatory mechanism of ubiquitin specific protein 1 by reactive oxygen species and the requirement of the UAF1 subunit .....	91
3.1	Introduction .....	91
3.1.1	Deubiquitinase .....	91
3.1.2	Crosstalk between translesion synthesis and Fanconi anemia pathway .....	93
3.1.3	Molecular mechanism of USP1 regulation by UAF1 .....	95
3.2	Results .....	96
3.2.1	Kinetic constants for USP1/UAF1 and USP1 .....	96
3.2.2	Reversibility of inactivation by H <sub>2</sub> O <sub>2</sub> .....	97
3.2.2.1	Inactivation of USP1 and USP1/UAF1 by H <sub>2</sub> O <sub>2</sub> .....	97
3.2.2.2	Inactivation of UCHs by H <sub>2</sub> O <sub>2</sub> .....	102
3.2.3	Time- and concentration-dependent inhibition of USP1/UAF1 by H <sub>2</sub> O <sub>2</sub> .....	103

3.2.4	Inactivation of USP1/UAF1 and USP1 by iodoacetamide.....	105
3.2.5	pH dependence of inactivation of USP1/UAF1 by iodoacetamide.....	106
3.2.6	Probe the sulfenic acid species in USP1/UAF1 upon oxidation by DAz1-Biotin .....	108
3.2.7	Phosphorylation of USP1/UAF1 .....	110
3.2.8	<i>In vitro</i> deubiquitination of monoubiquitinated human PCNA (hPCNA-Ub) by USP1/UAF1 .....	111
3.2.9	Mass spectrometry analysis of USP1/UAF1 following in- gel digestion.....	113
3.3	Discussion.....	118
3.3.1	UAF1 acts as an activator of USP1 .....	118
3.3.2	USP1/UAF1 complex is more sensitive to H <sub>2</sub> O <sub>2</sub> .....	119
3.3.3	Modification of USP1/UAF1 by iodoacetamide .....	124
3.3.4	Phosphorylation of USP1 .....	129
3.3.5	<i>In vitro</i> deubiquitination of native monoubiquitinated human PCNA-Ub .....	130
3.4	Materials and Methods .....	131
3.4.1	Determine the steady state kinetic parameters of USP1 and USP1/UAF1.....	131
3.4.2	Time- and concentration-dependent inhibition of USP1/UAF1 and USP1 by H <sub>2</sub> O <sub>2</sub> .....	131
3.4.3	Reversibility of inactivation by H <sub>2</sub> O <sub>2</sub> of USP1/UAF1 or USP1.....	132
3.4.3.1	Inactivation of USP1/UAF1 or USP1 by H <sub>2</sub> O <sub>2</sub> .....	132
3.4.3.2	Reactivation by DTT or glutathione.....	132
3.4.4	Reversibility of inactivation by H <sub>2</sub> O <sub>2</sub> of USP7, UCH-L3 and UCH-L1 in DTT .....	133
3.4.4.1	Inactivation of DUBs by H <sub>2</sub> O <sub>2</sub> .....	133
3.4.4.2	Reactivation of treated DUBs by DTT .....	133
3.4.5	Inactivation of USP1/UAF1 and USP1 by iodoacetamide under different pHs.....	134
3.4.6	pH dependence of the inactivation rate of USP1/UAF1 by iodoacetamide.....	134
3.4.7	Modifications of USP1/UAF1 by DAz1-Biotin probe .....	135
3.4.8	Detect phosphorylation of USP1/UAF1 .....	136
3.4.9	<i>In vitro</i> deubiquitination assay of monoubiquitinated human PCNA by USP1/UAF1 .....	137
3.4.10	Sample preparation for in-gel digestion .....	137
References	.....	138

4.	Discovery of small-molecule inhibitors against the human USP1/UAF1 deubiquitinating enzyme complex .....	143
4.1	Introduction .....	143
4.2	Results .....	148
4.2.1	qHTS against USP1/UAF1 using Ub-Rho as a substrate .....	148
4.2.2	Develop an orthogonal diubiquitin cleavage assay .....	148
4.2.3	Validate the hits using the di-Ub cleavage assay.....	150
4.2.4	Specificity of the USP1/UAF1 inhibitors against human USPs .....	152
4.2.5	Specificity of the USP1/UAF1 inhibitors against the UCH-family DUBs and other cysteine proteases .....	155
4.2.6	Pimozide and GW7647 are reversible inhibitors .....	156
4.2.7	Pimozide and GW7647 inhibit USP1/UAF1 by a noncompetitive mechanism .....	158
4.2.8	Inhibition mechanism determination of USP1/UAF1 by methylergonovine and beta-Lapachone using Ub-AMC as the substrate.....	160
4.2.9	Validate the hits using the di-Ub cleavage assay for the second set of compounds.....	162
4.3	Discussion.....	165
4.4	Materials and Methods .....	170
4.4.1	Kinetic parameters determination for USP1/UAF1 using K63 di-Ub as a substrate.....	170
4.4.2	Assay inhibition of USP1/UAF1 using K63 di-Ub as a substrate .....	171
4.4.3	Assess the specificity of USP1/UAF1 inhibitors.....	171
4.4.4	Assay inhibition of USP1 and UCH using Ub-rhodamine .....	172
4.4.5	Assay inhibition of SENP1 using SUMO1-AMC .....	172
4.4.6	$K_i$ determination for top USP1/UAF1 inhibitors.....	173
4.4.7	Reversibility of inhibition.....	173
4.4.8	$K_i$ determination for USP1/UAF1 inhibitor methylergonovine and beta-Lapachone using Ub-AMC as the substrate.....	174
4.4.9	Assess the specificity of USP1/UAF1 inhibitors methylergonovine and beta-Lapachone .....	175
	References .....	176
	Appendix .....	182

## LIST OF TABLES

Table 1.1	Kinetic parameters for RluA with tRNA and variant ASLs .....	18
Table 3.1	Kinetic parameters for USP1 and USP1/UAF1 with Ub-AMC .....	97
Table 3.2	Inactivation rate constants ( $M^{-1} \text{ min}^{-1}$ ) measured in $H_2O$ or $D_2O$ under different pLs (H or D). ND indicated not determined. ....	108
Table 3.3	Mass spectrometry results of various USP1 samples subjected to different modification. ....	118
Table 3.4	Inactivation rates by $H_2O_2$ of various proteins containing active site cysteines.....	120
Table 4.1	The $IC_{50}$ ( $\mu M$ ) value of the top forty two compounds in inhibiting USP1/UAF1 by Ub-Rho or diubiquitin. NI, no significant inhibition was observed at the highest inhibitor concentration of 114 $\mu M$ . FI, full inhibition was observed at the lowest inhibitor concentration of 0.08 $\mu M$ . ....	150
Table 4.2	The $IC_{50}$ ( $\mu M$ ) value of the top five compounds in inhibiting human USPs determined using K63-linked diubiquitin substrate. NI, no significant inhibition was observed at the highest inhibitor concentration of 114 $\mu M$ . ....	154
Table 4.3	The $IC_{50}$ ( $\mu M$ ) value of the top five compounds in inhibiting human DUBs and other cysteine proteases.....	155
Table 4.4	The $IC_{50}$ ( $\mu M$ ) value of methylergonovine and beta-Lapachone in inhibiting various DUBs using Ub-AMC as the substrate. ....	161
Table 4.5	The $IC_{50}$ ( $\mu M$ ) value of twenty three compounds in inhibiting USP1/UAF1 by diubiquitin. NI, no significant inhibition was observed at the highest inhibitor concentration of 114 $\mu M$ .....	163

Table 4.6	The IC <sub>50</sub> (μM) value of the top four compounds in inhibiting human USPs determined using K63-linked diubiquitin substrate. NI, no significant inhibition was observed at the highest inhibitor concentration of 114 μM. ....	165
-----------	---	-----

## LIST OF FIGURES

Figure 1.1:	The conversion from uridine to pseudouridine. ....	2
Figure 1.2:	Mechanisms proposed for $\Psi$ synthases. A, the Michael mechanism; B, the acylal mechanism. ....	4
Figure 1.3:	Schemes showing the different behavior of RluA, TruA and TruB towards [F <sup>5</sup> U]RNA.....	6
Figure 1.4:	Cloverleaf schematic of tRNA <sup>Phe</sup> showing uridine residues that are isomerized to $\Psi$ by RluA and TruB. ASL is <u>anticondon stem-loop</u> and TSL is <u>T-armed stem-loop</u> .....	7
Figure 1.5:	Superposition of the free (green) and RluA-bound (blue) ASLs.....	8
Figure 1.6:	Various interactions observed from the cocrystal structure of RluA (blue) bound with ASL (brown) containing 5-fluorouridine (PDB ID: 2I82). The water-mediated hydrogen bonding between the side chain of Arg-62 and N6 of A38 is not shown because the water molecules are not included in the structure file. Dashed lines denote hydrogen bonds. ....	9
Figure 1.7:	The relative positions of A37 and A38 (yellow) with RluA (blue). The pink sphere denotes for Trp-119. ....	9
Figure 1.8:	HPLC analysis of the nuclease S1/alkaline phosphatase digestion of mixture of (A) 3 $\mu$ M ASL incubated in the absence (above) and presence (below) of 1 $\mu$ M RluA (B) 3 $\mu$ M U33C ASL incubated in the absence (above) or presence (below) of 1 $\mu$ M RluA, and (C) 3 $\mu$ M ASL incubated with 1 $\mu$ M RluA that had been heat denatured. All incubations were 1 h at 37 $^{\circ}$ C. ....	11
Figure 1.9:	RluA activity towards ASL and seven variant ASLs. (A) 3 $\mu$ M ASL incubated with 50 nM RluA (white bar) and 1 $\mu$ M RluA (black bar) for 1 h. (B) 3 $\mu$ M ASL variants (U33C, U33G, A36C) incubated with 1 $\mu$ M RluA overnight.....	12

Figure 1.10: The structure comparison between U and G. Panel B is the overlap of panel A and panel C. ....	13
Figure 1.11: The hydrogen bonds made by U33 A36 and C33 A36. ....	14
Figure 1.12: Wild-type ASL on POROS HQ strong anion exchange column before (A) and after (B) heating and cooling. ....	15
Figure 1.13: Representative HPLC traces of (A) three time points of 3 $\mu$ M WT ASL incubated with 10 nM RluA (B) three time points of 1 $\mu$ M A36C ASL incubated with 50 nM RluA. ....	16
Figure 1.14: Briggs-Haldane plots of the kinetic data for RluA with variant ASLs as substrates. The RluA concentration was 10 nM except for A36C ASL and U33C ASL, for which 50 nM RluA was used. Each point is the average of at least two independent determinations. Error bars smaller than the data point are not shown. ....	18
Figure 1.15: (A) The water mediated hydrogen bonds formed between Arg-62, Pro-36 of RluA and A38 of ASL. (B) The hypothetical water mediated hydrogen bonds formed between Arg-62, Pro-36 of RluA and C38 of ASL. The pictures are generated by Pymol software <sup>33</sup> using PDB file: 2I82. The depicted water molecules are not included in the pdb file, so they were added manually during the preparation of the graphic based on their position as presented in the report of the cocrystal. <sup>25</sup> ....	20
Figure 1.16: SDS-PAGE showing a time course of adduct formation between RluA and variant [F <sup>5</sup> U]ASLs. ....	21
Figure 1.17: HPLC traces of the intact variant [F <sup>5</sup> U]ASLs and the [F <sup>5</sup> U]ASLs modified after incubation with different amount of RluA. ....	24
Figure 1.18: The sequence comparison of [F <sup>5</sup> U]ASL, A36C [F <sup>5</sup> U]ASL, U33C [F <sup>5</sup> U]ASL and [F <sup>5</sup> U]TSL. ....	25
Figure 1.19: HPLC analysis of the products of F <sup>5</sup> U. (A) buffer components (upper), [F <sup>5</sup> U]ASL incubated in the absence (middle) and presence (bottom) of RluA. (B) Blow-ups of the boxed areas in panel A. Inosine (I) appears as the result of the presence of adenosine deaminase in the S1 nuclease. ....	26

Figure 1.20: HPLC analysis of the products of F <sup>5</sup> U. (A) buffer components (upper), A36C [F <sup>5</sup> U]ASL incubated in the absence (middle) and presence (bottom) of RluA. (B) Blow-ups of the boxed areas in panel A. Inosine (I) appears as the result of the presence of adenosine deaminase in the S1 nuclease. ....	26
Figure 1.21: HPLC analysis of the products of F <sup>5</sup> U. (A) buffer components (upper), U33C [F <sup>5</sup> U]ASL incubated in the absence (middle) and presence (bottom) of RluA. (B) Blow-ups of the boxed areas in panel A. Inosine (I) appears as the result of the presence of adenosine deaminase in the S1 nuclease. ....	27
Figure 1.22: Overlaid HPLC traces of nuclease S1/alkaline phosphatase digestion of [F <sup>5</sup> U]TSL incubated with TruB (blue) and U33C [F <sup>5</sup> U]ASL incubated with RluA (red). The blue star denotes the buffer peak that always elutes before the digested [F <sup>5</sup> U]TSL products, and the red star denotes the buffer peak that always elutes after the digested U33C [F <sup>5</sup> U]ASL products. ....	28
Figure 2.1: Cellular pathways associated with different forms of ubiquitin modification. K48 or K29 linked polyubiquitin chain targets protein for proteasome degradation. Monoubiquitination or polyubiquitination formed through other lysine linkages have nonproteolytic functions, such as endocytosis and DNA repair.....	41
Figure 2.2: PCNA ubiquitination and SUMOylation pathways. PCNA (orange ring) can be modified by SUMO (purple sphere), monoubiquitin or ubiquitin chains (green spheres) at different sites. SUMOylation on PCNA can occur either on K164 or K127. Monoubiquitination was observed on K164. Polyubiquitination on PCNA K164 is critical for an error-free DNA damage tolerance pathway.....	42
Figure 2.3: Two models of translesion synthesis. A, postreplicative gap-filling model. B, polymerase switch model. Note: this figure was generated by William Bozza and Dr. Yongxing Ai.....	45
Figure 2.4: Three different strategies to generate homogenous diubiquitins. (a) Genetic engineering and synthetic chemistry was used to create K6 and K29 di-Ub. (b) Native chemical ligation to synthesize K48 di-Ub. (c) Native chemical ligation to synthesize all seven homogenous di-Ub. Figure a and c were originally published in references. <sup>36, 40</sup> Permission is granted by publishers. ....	49

Figure 2.5:	Native chemical ligation. Unprotected peptide segments are reacted by means of reversible thiol/thioester exchange to give thioester-linked initial reaction products. Uniquely, the thioester-linked intermediate involving an N-terminal Cys residue (boxed) is able to undergo nucleophilic rearrangement by a highly favored intramolecular mechanism; this step is irreversible (under the conditions used) and gives a polypeptide product that is linked by a native amide (i.e. peptide) bond. Only a single reaction product is obtained, even in the presence of additional Cys residues in either segment. The product polypeptide is subsequently folded to give the desired synthetic protein molecule. ....	51
Figure 2.6:	(a) Retrosynthetic analysis of uH2B synthesis. uH2B was generated by a three-piece ligation with several polypeptides. (b) Synthetic scheme for the ubiquitination of H2B. This figure was originally published in the reference. <sup>41</sup> Permission is granted by the publisher. ....	52
Figure 2.7:	Chemical ubiquitination of PCNA. (a) The disulfide-mediated ligation between ubiquitin-cysteamine, generated by cleaving ubiquitin off chitin resin with cysteamine, and K164C PCNA. The chemically ubiquitinated PCNA mimics the native ubiquitinated PCNA. (b) Denaturing SDS-PAGE gel showing the progression of the chemical ligation reaction and the final product following an anion exchange chromatography. (c) The multiple charge states of the ubiquitinated PCNA determined by electrospray ionization mass spectrometry. The inset shows the deconvoluted peaks of K164C PCNA-Ub (37,440 Da). ....	55
Figure 2.8:	Denaturing SDS-PAGE analysis of the chemically ubiquitinated PCNA (K164C PCNA-Ub) treated with 5 mM DTT at different time points (up to one hour) or 100 mM DTT for 5 min. ....	55
Figure 2.9:	A control chemical ubiquitination reaction using a cysteine-free PCNA mutant (C22S/C30S/C62S/C81S PCNA) that retains Lys164 PCNA (designated as K164 PCNA). ....	56
Figure 2.10:	Deconvoluted electrospray ionization mass spectra of the tryptic peptides of K164C PCNA-Ub. ....	57

Figure 2.11: The MS/MS spectra of the ion at m/z of 1157.60 that corresponds to the precursor peptide from K164C PCNA-Ub trypsin digestion that contains the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys164 (DLSQLSDSINIMITCETIK) with a molecular weight of 2313.07 Da (m/z 1157.60 for the doubly charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The MS/MS spectrum with the measured mass and the fragments labeled. (c) The zoomed-in MS/MS spectrum between m/z 200 and 1160. ....	59
Figure 2.12: Denaturing SDS-PAGE analysis of the chemically ubiquitinated PCNA (K164C PCNA-Ub) treated with (a) 0.25, 0.5 mM DTT (left) and (b) 0.25, 0.5 mM glutathione (right) respectively for 5 minutes at 37 °C. Top, SDS-PAGE gel of the treated K164C PCNA-Ub (comparable amount of protein was loaded); Bottom, the normalized percentage of intact K164C PCNA-Ub after the treatment. The amount of K164C PCNA-Ub loaded onto the SDS-PAGE gel was normalized for the different incubation reactions to give a comparable intensity. ....	60
Figure 2.13: Denaturing SDS-PAGE analysis of the chemically ubiquitinated PCNA at residue 164, 127, 44 and 107 respectively. ....	61
Figure 2.14: The multiple charge states of the ubiquitinated PCNA subunit (K127C PCNA-Ub, R44C PCNA-Ub and K107C PCNA-Ub) determined by electrospray ionization mass spectrometry. The inset shows the deconvoluted peak of a molecular weight of 37,440 Da for K127C PCNA-Ub, 37,410 Da for R44C PCNA-Ub and 37,440 Da for K107C PCNA-Ub. ....	62
Figure 2.15: The MS/MS spectra of the precursor peptide from K127C PCNA-Ub trypsin digestion that contains the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys127 (LMDIDADFLCIEELQYDSTLSLPSSEFSK) with a molecular weight of 3498.64 Da (m/z 1750.76 for the doubly charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The full MS/MS spectrum of the ion at m/z 1,750.76. (c)-(e) The zoomed-in regions of the MS/MS spectra with the measured mass and the fragments labeled. ....	63

Figure 2.16: The MS/MS spectra of the precursor peptide from R44C PCNA-Ub trypsin digestion that contains the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys44 (EDGIIAQAVDDSC <u>V</u> LLVSLEIGVEAFQEYR) with a molecular weight of 3470.70 Da (m/z 1736.80 for the doubly charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The full MS/MS spectrum of the ion at m/z 1,736.80. (c) & (d) The zoomed-in regions of the MS/MS spectra with the measured mass and the fragments labeled. ....	64
Figure 2.17: The MS/MS spectra of the precursor peptide from K107C PCNA-Ub trypsin digestion that contains the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys107 (SGNNTDTLTLIADNTPDSIILLFEDT <u>C</u> K) with a molecular weight of 3213.57 Da (m/z 1072.52 for the triply charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The full MS/MS spectrum of the ion at m/z 1,072.52. (c) & (d) The zoomed-in regions of the MS/MS spectra with the measured mass and the fragments labeled. ....	65
Figure 2.18: Denaturing SDS-PAGE gel analysis of the chemically SUMOylated PCNA at residue 164 (K164C PCNA-SUMO) and 127 (K127C PCNA-SUMO). ....	66
Figure 2.19: The multiple charge states of (a) SUMO-cysteamine (SUMO-SH), (b) K164C PCNA-SUMO, (c) K127C PCNA-SUMO determined by electrospray ionization mass spectrometry. The inset in (a) shows the deconvoluted peak with a molecular weight of 11,190.0 Da for SUMO-SH, which indicates the N-terminal methionine of the recombinant SUMO is removed by methionine aminopeptidase. The theoretical molecular weight of SUMO-SH (amino acid 2-98, without N-ter Met) is 11,189.5 Da. The measured molecular weight of K164C PCNA-SUMO (40,010 Da) and K127C PCNA-SUMO (40,010 Da) agrees with the calculated molecular weight (40,014.3 Da) for PCNA-SUMO (without the N-ter Met in SUMO). ....	69

Figure 2.20: The MS/MS spectra of the precursor peptide from K164C PCNA-SUMO trypsin digestion that contains the SUMO C-terminal pentapeptide (EQIGG) conjugated to the PCNA peptide flanking Cys164 (DLSQLSDSINIMITCETIK) with a molecular weight of 2683.65 Da (m/z 1342.66 for the doubly charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The full MS/MS spectrum with the measured mass and the fragments labeled. (c) The zoomed-in MS/MS spectrum between m/z 200 and 1340. ....	71
Figure 2.21: The MS/MS spectra of the precursor peptide from K127C PCNA-SUMO trypsin digestion that contains the ubiquitin C-terminal pentapeptide (EQIGG) conjugated to the PCNA peptide flanking Cys127 (LMDIDADFLCIEELQYDSTLSLPSEFSK) with a molecular weight of 3869.14 Da (m/z 1290.89 for the triply charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The full MS/MS spectrum with the measured mass and the fragments labeled. (c) & (d) The zoomed-in MS/MS spectra between m/z 200-780 and 1725-2205 respectively. ....	72
Figure 2.22: New methods for monoubiquitination of PCNA. (a) Chemical ubiquitination. (b) Splitting PCNA method by Washington group.....	73
Figure 2.23: Diagram of the two polypeptides used to generate monoubiquitinated PCNA by Washington group. <sup>49</sup> The N fragment contained amino acid residues 1 to 163 of PCNA and was N-terminally FLAG-tagged. The C fragment contained the entire ubiquitin sequence (residues 1–76) fused via a short linker to residues 165 to 258 of PCNA and was N-terminally His <sub>6</sub> -tagged. The short linker consisted of two glycine residues because this is nearly isosteric with the side chain of Lys 164 and the isopeptide bond to the C-terminus of ubiquitin.....	74

Figure 2.24:	(Left) The positions of chemical ubiquitination mapped to the homotrimeric PCNA, shown as front and side views. Lys164 (red), Lys127 (green), Lys107 (magenta), Arg44 (blue) and the C-terminal residue Glu258 (yellow) are shown in one PCNA subunit. (Right) (a) The extent of polymerase exchange reported by the decrease in DNA synthesis for ubiquitinated and SUMOylated PCNA. The different forms of modified PCNA include chemically ubiquitinated PCNA at residue 164 (K164C PCNA-Ub), enzymatically ubiquitinated PCNA (WT PCNA-Ub), chemically SUMOylated PCNA at residue 164 (K164C PCNA-SUMO) and chemically SUMOylated PCNA at residue 127 (K127C PCNA-SUMO). (b) Comparing the function of PCNA ubiquitinated at different positions in the polymerase exchange assay. The positions of chemical ubiquitination include residues 164, 127, 107 and 44. PCNA-Ub fusion refers to a chimeric protein with ubiquitin fused to the C terminus of PCNA. The ubiquitinated PCNAs and the unmodified wild-type and mutant PCNAs are compared. Note: Bar graph a&b were generated by Dr. Yongxing Ai.....	77
Figure 2.25:	Cartoon illustration of possible interactions among PCNA, ubiquitin and Pol $\eta$ when ubiquitin is conjugated to various positions (164, 127, 44 and 107) on PCNA. Pol $\eta$ is composed of catalytical domain, UBZ (ubiquitin binding) domain and PIP (PCNA interacting peptide). Pol $\eta$ interacts with the front side (the right side in the cartoon) of PCNA. The long linkage formed by the C-terminal peptide of ubiquitin and the lysine side chain of PCNA allows ubiquitin to adopt various conformations that can lead to a functional ternary complex with Pol $\eta$ . .....	78
Figure 3.1:	Oxidation states of protein cysteines that are implicated in biological function. Thiol can be oxidized into sulfenic, sulfinic and sulfonic acid.....	99
Figure 3.2:	Selective reaction of dimedone with a sulfenic acid affords a new thioether bond.....	99

- Figure 3.3: A flow chart shows the experimental design of the inactivation of DUBs by H<sub>2</sub>O<sub>2</sub>. The DUB was inactivated when it was oxidized into sulfinic acid and sulfonic acid because the oxidation is not reversible. Sulfinic acid form can be reduced back to -SH in the presence of reducing agent, like DTT. To verify the existence of sulfinic acid form, dimedone was used to trap the sulfinic acid intermediate. This modification is irreversible and thus inactivating the DUB. However, if the DUB was first treated with H<sub>2</sub>O<sub>2</sub> then followed by catalase instead of dimedone, the sulfinic acid intermediate will be reduced back to -SH in the presence of DTT. This will lead to the reactivation of the DUB. .... 100
- Figure 3.4: Inactivation and reactivation of various DUBs. DUBs inactivated by 2 μM – 1 mM H<sub>2</sub>O<sub>2</sub> treatment (black bar) is diluted into either DTT or glutathione and monitored for recovery of original activity. Dimedone was used to trap sulfinic acid form after H<sub>2</sub>O<sub>2</sub> treatment (white bar) or catalase was used to degrade H<sub>2</sub>O<sub>2</sub> following H<sub>2</sub>O<sub>2</sub> treatment (grey bar). The plotted data represents mean value ± s.d. (a) USP1/UAF1 in DTT (b) USP1/UAF1 in glutathione (c) USP1 in DTT (d) USP1 in glutathione (e) USP7 in DTT. .... 101
- Figure 3.5: Inactivation and reactivation of (a) UCH-L3 and (b) UCH-L1. UCHs inactivated by 50 μM – 1 mM H<sub>2</sub>O<sub>2</sub> treatment (black bar) is diluted into DTT and monitored for recovery of original activity. Dimedone was used to trap sulfinic acid form after H<sub>2</sub>O<sub>2</sub> treatment (white bar) or catalase was used to degrade H<sub>2</sub>O<sub>2</sub> following H<sub>2</sub>O<sub>2</sub> treatment (grey bar). The plotted data represents mean value ± s.d. .... 102

- Figure 3.6: Time- and concentration-dependent inhibition of USP1/UAF1 and USP1 by H<sub>2</sub>O<sub>2</sub>. (a) Time-dependent loss in USP1/UAF1 activity is observed after pre-incubation with 5 (■), 10 (□), 20 (▲), 40 (Δ) μM H<sub>2</sub>O<sub>2</sub>, with observed inactivation rates of 0.69, 1.19, 2.60 and 5.61 × 10<sup>-3</sup> s<sup>-1</sup>, respectively. Fits are obtained using Equation 1. (b) Observed pseudo first-order inactivation rates for USP1/UAF1 vary linearly with H<sub>2</sub>O<sub>2</sub> concentration and are fit using a second order rate constant of 140.33 M<sup>-1</sup> s<sup>-1</sup>. (c) Time-dependent loss in USP1 activity is observed after pre-incubation with 100 (■), 150 (□), 200 (▲), 300 (Δ) μM H<sub>2</sub>O<sub>2</sub>, with observed inactivation rates of 0.75, 1.25, 1.53 and 2.22 × 10<sup>-3</sup> s<sup>-1</sup>, respectively. Fits are obtained using Equation 1. (d) Observed pseudo first-order inactivation rates for USP1 vary linearly with H<sub>2</sub>O<sub>2</sub> concentration and are fit using a second order rate constant of 7.41 M<sup>-1</sup> s<sup>-1</sup>. ..... 104
- Figure 3.7: Inactivation of USP1 (white bar) and USP1/UAF1 (black bar) by iodoacetamide under different pHs. Both species were inactivated by 4 mM iodoacetamide for 2 min under room temperature. The remaining activity was monitored by diluting the treated DUBs into an assay buffer containing Ub-AMC and DTT. The plotted data represents mean value ± s.d. .... 105
- Figure 3.8: Inactivation of USP1/UAF1 by iodoacetamide under pH 7.6. Time-dependent inactivation of USP1/UAF1 activity was observed after pre-incubation with (A) 2 (●), 4 (■), 8 (▲) mM iodoacetamide. The pseudo first-order inactivation rate constants are plotted against iodoacetamide concentration as shown in the inset. The inactivation rate constant (k<sub>2</sub>) obtained is 1.1 M<sup>-1</sup> s<sup>-1</sup> (66 M<sup>-1</sup> min<sup>-1</sup>). ..... 106
- Figure 3.9: Plot of log(k<sub>2</sub>) measured for the inactivation of USP1/UAF1 with iodoacetamide vs reaction solution pH fitted to equation 2 to define the pK<sub>a</sub>. The points marked with (●) signs were determined in H<sub>2</sub>O and with (▲) were determined in 98% deuterium oxide. .... 107

- Figure 3.10: Trapping proteins that undergo sulfenic acid modification. (a) Structure of DAz-1 which is a derivative of dimedone. Biotin is conjugated to DAz1. (b) Strategy of utilizing dimedone derived probes to modify sulfenic acid in USP1. The modified USP1 can be detected by anti-Biotin which is HRP-streptavidin in western blot..... 109
- Figure 3.11: Western blot of DAz1-Biotin labeled USP1/UAF1. USP1/UAF1 was treated with or without H<sub>2</sub>O<sub>2</sub> (20 or 40 μM). It was supposed to see increased USP1 band intensity with the increased H<sub>2</sub>O<sub>2</sub> concentration. P-6 column was used to remove reducing reagent TCEP here. The reaction treated without P-6 column was used as a negative control in which no difference was observed between with or without H<sub>2</sub>O<sub>2</sub>..... 110
- Figure 3.12: Phosphorylation of the USP1 subunit. A polyacrylamide gel containing USP1, UAF1 and USP1/UAF1 with or without calf intestinal alkaline phosphatase (CIP) were stained with Pro-Q Diamond phosphoprotein stain..... 111
- Figure 3.13: Time-dependent deubiquitination of native hPCNA-Ub by USP1/UAF1. 200 ng native hPCNA-Ub were incubated with 20 nM (36 ng) or 50 nM (90 ng) USP1/UAF1 for 10, 30, and 60 mins. hPCNA-Ub and hPCNA were detected by anti-PCNA antibody. .... 111
- Figure 3.14: The correlation between the Western blotting signal and known amounts of hPCNA. (a) 15 – 75 ng hPCNA was loaded on SDS-PAGE and detected by Western blotting (with anti-PCNA antibody). (b) A linear response was observed between 15 and 75 ng hPCNA by plotting the known amount of hPCNA against the Western blotting signal..... 112
- Figure 3.15: (a) SDS-PAGE of N84K USP1 (left) and N84K USP1/UAF1 complex (right) treated without or with H<sub>2</sub>O<sub>2</sub> followed by dimedone. (b) The trypsin digested fragment containing active site cysteine 90 (in red) can be modified by dimedone or iodoacetamide. Both cysteines (C90 and C102) in this fragment can be modified and the corresponding m/z value was listed. .... 116

- Figure 3.16: The MS/MS spectra of the precursor peptide from N84K USP1 trypsin digestion that contains the active site C90 ((K)NLGNTCYLNSILQVLYFCPGFK(S)) with a molecular weight of 2621.29 Da. This spectra was generated by Leila Choe..... 116
- Figure 3.17: (a) SDS-PAGE of N84K/C102S USP1/UAF1 complex treated without, with H<sub>2</sub>O<sub>2</sub> followed by dimedone, or with iodoacetanilide. (b) The trypsin digested fragment containing active site cysteine 90 (in red) can be modified by dimedone, iodoacetamide and iodoacetanilide. Only C90 in this fragment can be modified and the corresponding m/z value was listed..... 117
- Figure 3.18: (a) Aligned crystal structure of USP7 bound (red, PDB: 1NBF) with ubiquitin (pink) and unbound USP7 (blue, PDB: 1NB8). Catalytic triad area is boxed. (b) Zoomed-in catalytic triad of bound and unbound USP7. The distance from the N<sup>δ1</sup> of His 464 to the sulfur atom of Cys 223 is 10.17 Å in unbound form while it is 3.60 Å in the bound form..... 121
- Figure 3.19: Comparison of sensitivity to H<sub>2</sub>O<sub>2</sub> of various DUBs. Each DUB's remaining activity was plotted against different concentrations of H<sub>2</sub>O<sub>2</sub>. ..... 122
- Figure 3.20: (a) Aligned crystal structure of UCH-L3 (red, PDB: 1UCH) with UCH-L1 (blue, PDB: 2ETL). Catalytic triad area is boxed. The loop in green from UCH-L1 distinguishes itself from UCH-L3 in which no such loop was found. (b) Zoomed-in catalytic triad of UCH-L3 and UCH-L1. The distance from the N<sup>δ1</sup> of His 169 to the sulfur atom of Cys 95 is 3.76 Å in UCH-L3 while it is 7.69 Å in UCH-L1 from the N atom of His 161 to the S atom of Cys 90..... 123
- Figure 3.21: Two proposed mechanisms for cysteine proteases using Ub-AMC as the substrate. (a) Thioate-imidazolium ion pair mechanism. The boxed area indicates the uncharged pair of Cys-SH/His-Im and the ion pair of Cys-S<sup>-</sup>/His-ImH<sup>+</sup>. (b) General base mechanism. Figure b was generated by Mark Villamil. Figure a was modified by Junjun Chen based on Figure b. .... 127
- Figure 4.1: (a) SDS-PAGE showing the progression of the diubiquitin cleavage at different substrate concentration by USP1/UAF1 at different time points. (b) Michaelis-Menten plot of the kinetic data for USP1/UAF1 with K63 diubiquitin as substrate. The USP1/UAF1 concentration was 100 nM. .... 149

Figure 4.2:	Inhibition of USP1/UAF1 by pimozide (a) and GW7647 (b). Dose-dependent inhibition of USP1/UAF1 activity (left) and SDS-PAGE analysis (right) of the cleavage of K63-linked diubiquitin in the presence of different concentrations of inhibitors are shown. ....	151
Figure 4.3:	Pimozide (a) and GW7647 (b) tested as inhibitors against human USP7, USP2, USP5 and USP8 using K63 diubiquitin as substrate. ....	152
Figure 4.4:	Reversibility of inhibition by USP1/UAF1 inhibitors. (a) SDS-PAGE showing the diubiquitin cleavage by USP1/UAF1 that was pre-incubated with different inhibitors and then rapidly diluted into assay buffer. (b) The remaining activity of USP1/UAF1 treated with five inhibitors. The bar graph represents mean of three independent experiments $\pm$ s.d. Pimozide, GW7647, flupenthixol and trifluoperazine are reversible inhibitors. Rottlerin is a largely irreversible inhibitor. Beta-Lapachone was tested as an irreversible inhibitor control here. ....	157
Figure 4.5:	SDS-PAGE showing the progression of the diubiquitin cleavage of USP1/UAF1 in the presence of 0, 0.5 $\mu$ M, 1 $\mu$ M and 2 $\mu$ M (a) Pimozide and (b) GW7647. ....	159
Figure 4.6:	Determine the inhibition mechanism of USP1/UAF1 by Pimozide and GW7647. (a) Lineweaver-Burk plots of USP1/UAF1 catalyzed hydrolysis of diubiquitin in the presence of 0 (■), 0.5 (●), 1 (▲), 2 (◆) $\mu$ M pimozide. (b) Lineweaver-Burk plots of USP1/UAF1 catalyzed hydrolysis of diubiquitin in the presence of 0 (■), 0.5 (●), 1 (▲), 2 (◆) $\mu$ M GW7647. Each data point represents mean of two independent experiments $\pm$ s.d. The $K_i$ values for pimozide ( $0.50 \pm 0.02 \mu\text{M}$ ) and GW7647 ( $0.75 \pm 0.04 \mu\text{M}$ ) were determined by plotting the slopes obtained from the Lineweaver-Burk plot against the inhibitor concentrations (inset). ....	159

Figure 4.7: Lineweaver-Burk plot of the inhibition of USP1/UAF1 with (a) methylergonovine maleate. Inhibition was by 0 (□), 100 (■), 200 (Δ), 400 (▲), 800 (○) nM methylergonovine maleate against varying concentrations of Ub-AMC. (b) Inhibition was by 0 (□), 1 (■), 2 (Δ), 4 (▲) μM beta-Lapachone against varying concentrations of Ub-AMC. The  $K_i$  for methylergonovine maleate and beta-Lapachone were determined from a plot of the slopes obtained from the Lineweaver-Burk plot at each concentration of inhibitors (inset). ..... 161

## ABSTRACT

Chapter 1: Pseudouridine synthases ( $\Psi$  synthases) catalyze the isomerization of uridine to pseudouridine ( $\Psi$ ) and are present in all domains of life. The  $\Psi$  synthases fall into six families based on sequence similarity, and crystal structure of members of each family show that they share the same core fold and have one universally conserved aspartic acid residue in motif II. The cocrystal structures of the *E. coli*  $\Psi$  synthase RluA bound to an RNA oligomer corresponding to the anticodon stem-loop (ASL) of *E. coli* tRNA<sup>Phe</sup> has been determined. This structure is only the second of a  $\Psi$  synthase bound to an RNA substrate. In the RluA•RNA cocrystal, the RNA has undergone significant conformational changes from its unbound form. The isomerized uridine U32, is everted in the active site of RluA. This conformation is stabilized by the reverse-Hoogsteen base pair formed by U33 and A36 and the hydrogen bond formed between A36 and Pro 36 on RluA. To test the importance of particular interactions in the RluA•RNA cocrystal structure, kinetic studies were undertaken. RNA containing 5-fluorouridine, F<sup>5</sup>U, has been used as a mechanistic probe to distinguish the two proposed  $\Psi$  synthases' mechanisms. RluA forms heat-sensitive and apparently covalent adduct with [F<sup>5</sup>U]RNA. Upon heating, hydrated F<sup>5</sup>U products are also formed by the collapse of the adduct of RluA. Thus, the reactivity between RluA and two altered [F<sup>5</sup>U]ASLs (A36C and U33C) was investigated.

Chapter 2 to Chapter 4: The non-proteolytic function of ubiquitination has attracted increasing attention in recent years, including protein trafficking, immune

response, transcription regulation and DNA damage response. Monoubiquitination of proliferating cell nuclear antigen (PCNA) plays an important role in eukaryotic translesion synthesis (TLS), a mechanism utilized by cells to synthesize past DNA lesion. One obstacle in studying eukaryotic TLS resides in the difficulty of preparing sufficient amount of ubiquitinated PCNA for in-depth biochemical and biophysical investigation. In Chapter 2, we developed a chemical approach that combines the power of intein chemistry and the facile disulfide exchange chemistry for efficient protein ubiquitination and SUMOylation. The chemically ubiquitinated PCNA is functionally equivalent to the native ubiquitinated PCNA in effecting polymerase switch between the replicative and the specialized DNA polymerases. We also demonstrated the strict requirement of PCNA ubiquitination for polymerase switch. Moreover, we probed the effect of the site of ubiquitination by preparing chemically ubiquitinated PCNAs that differ only in the position of modification. Our study revealed a surprising degree of flexibility of ubiquitin modification.

Deubiquitinating enzymes (DUBs) cleave the ubiquitin moiety from mono- and poly-ubiquitinated proteins. Close to 100 DUBs have been identified in the human proteome. Abnormal cellular expression of DUBs or the loss of function due to mutation in certain DUB genes have been linked to various human diseases, including cancer and neurodegenerative diseases. Ubiquitin-specific proteases (USPs) constitute the largest DUB family. There is growing evidence suggesting that the activity of DUBs, in particular USPs, is stringently regulated through their interaction with many other protein partners. A recent global proteomic analysis of human DUBs identified 774 interacting proteins for the 75 DUBs studied. Remarkably, 34 human USPs were found to be associated with WD40-repeat proteins that adopt a  $\beta$ -propeller structure

comprising up to eight blades. Given its widespread occurrence, the interaction between WD40-repeat proteins and USPs likely represents a fundamentally important way of regulating USP activity. Based on the reported kinetic data for USP1/UAF1 complex and USP1 alone, we hypothesize that the interaction between UAF1 and USP1 may reorganize the catalytic triad into a more productive conformation. To probe the active site conformation, in Chapter 3, cysteine-reactive small organic molecules ( $\text{H}_2\text{O}_2$  or iodoacetamide) were used to determine the reactivity of the catalytic sulfhydryl group in USP1/UAF1 complex or USP1 alone.

DUBs are promising targets for pharmacological intervention. The advantage of inhibiting DUB lies in the specificity of therapeutic intervention that can lead to better efficacy and eliminate nonspecific side effects seen in proteasome inhibitors. In Chapter 4, we identified small-molecule inhibitors against the USP1/UAF1 complex through high throughput screening. Two highly selective inhibitors, pimozone and GW7647, inhibit USP1/UAF1 noncompetitively with a  $K_i$  of 0.50 and 0.75  $\mu\text{M}$ , respectively. We demonstrated that both compounds are reversible inhibitors that bind the USP1/UAF1 complex at a site different from the active site. Because USP1/UAF1 is involved in DNA damage response through deubiquitinating PCNA and FANCD2 in translesion synthesis and Fanconi anemia pathways, we tested USP1/UAF1 inhibitors as a sensitizer of cancer cells to DNA damaging agent, cisplatin. Our results indicated that the USP1/UAF1 inhibitors act synergistically with cisplatin in inhibiting the proliferation of cancer cells.

## Chapter 1

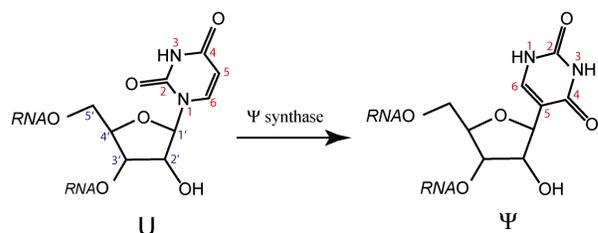
### THE MOLECULAR RECOGNITION OF RNA BY THE PSEUDOURIDINE SYNTHASE RLUA

#### 1.1 Introduction

Over 100 post-transcriptional modifications of RNA have been reported.<sup>1</sup> The conversion of uridine (U) to pseudouridine ( $\Psi$ ) occurs in all three domains of life (archaea, bacteria and eukaryotes) and is the most common post-transcriptional modification of RNA.<sup>2</sup>  $\Psi$  is an isomer of uridine containing a C–C instead of a C–N glycosidic bond (Figure 1.1). The C–C bond in  $\Psi$  has more rotational freedom than the C–N bond in uridine.  $\Psi$  can form one more hydrogen bond than uridine by using the protonated N1 in addition to the protonated N3 shared with U.<sup>3</sup>

$\Psi$  is the most prevalent modified nucleoside found in many cellular RNAs that are structurally and functionally important.  $\Psi$  has been reported to exist in ribosomal RNA (rRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and transfer RNA (tRNA).<sup>4</sup> The most well known  $\Psi$  in tRNA is the universal  $\Psi$ 55, which contributes to the stabilization of the local structural motifs through strong base pairing with other nucleosides in tRNA.<sup>4</sup> Pseudouridylation sites in RNA are highly conserved. For example, it is clustered at or near the peptidyltransferase center and the decoding center in rRNA.<sup>4</sup> The pseudouridine synthases ( $\Psi$  synthases) catalyze the isomerization of U to  $\Psi$ , which involves the breakage of the glycosidic bond that connects the ribose to the pyrimidine base, followed by a rotation of the uracil ring and formation of a C5-glycosyl bond (Figure

1.1). In addition to the formation of a new hydrogen-bond-donating N-H group, pseudouridylation may play an important role in increasing the conformational stability of RNA because  $\Psi$  can form a more stable base stacking arrangement than U.<sup>5</sup>



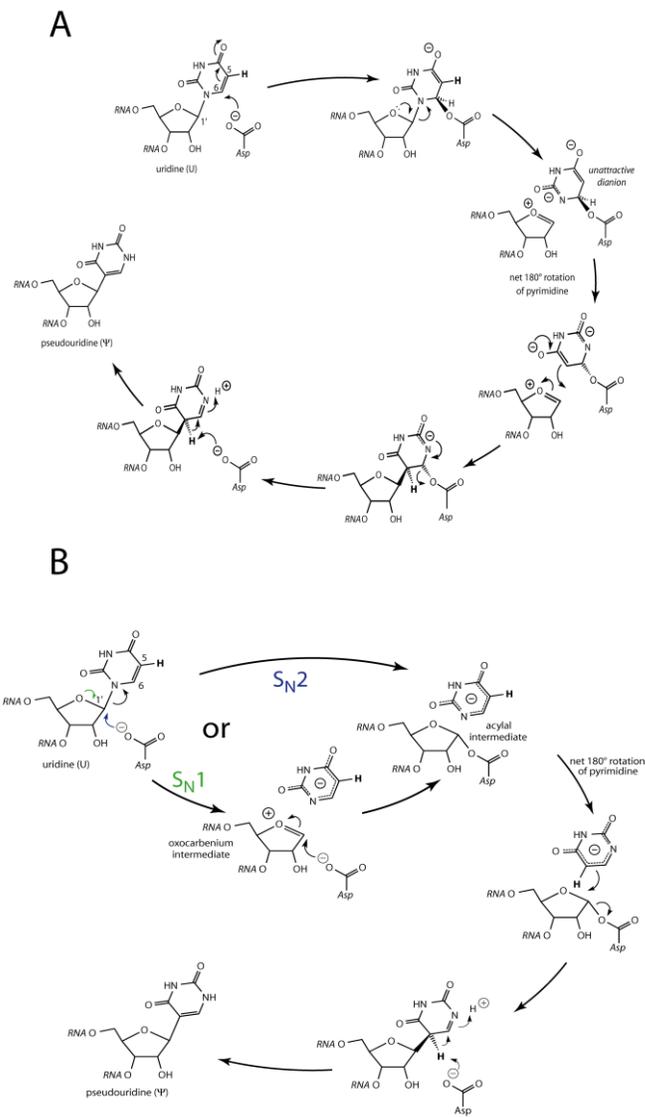
**Figure 1.1: The conversion from uridine to pseudouridine.**

All known  $\Psi$  synthases fall into six families that share no statistically significant global sequence similarity.<sup>6</sup> Each family is named for the representative enzyme that was cloned first: TruA, TruB, TruD, RluA, RsuA and Pus10. The enzyme names for TruA, TruB, TruD, RluA, RsuA from *Escherichia coli* are derived from their RNA substrate modified: TruA, TruB, and TruD are tRNA pseudouridine synthases A, B, and D; RluA stands for ribosomal large subunit pseudouridine synthase A; and RsuA is ribosomal small subunit pseudouridine synthase A. Pus10 is a human  $\Psi$  synthase. The structure determinations of  $\Psi$  synthases revealed that the six families share a core  $\beta$ -sheet fold and a conserved active-site cleft, indicating that all  $\Psi$  synthases diverged from a common ancestral protein.<sup>7-9</sup> An active site Asp is the only absolutely conserved residue and is essential for activity in all families.<sup>6</sup> Besides the core  $\beta$ -sheet, the  $\Psi$  synthases generally contain other secondary structural elements of helices and loops around the active site cleft, as well as a variety N- and C- terminal independently folded domains.

$\Psi$  synthases have been shown to have important biological functions in some cellular RNAs. Mutations of the human  $\Psi$  synthase dyskerin can result in the X-linked form of the skin and bone marrow disease dyskeratosis congenita.<sup>10</sup> Posttranscriptional modifications of U2 snRNA are essential for proper assembly of U2 snRNA into functional ribonuclear protein complexes (snRNPs) that are required for pre-mRNA splicing.<sup>11</sup> *Drosophila* gene *minifly* encodes an ubiquitous nucleolar protein that is a homolog of human dyskerin and is critical in rRNA processing and pseudouridylation. The mutation of *minifly* triggers pleiotropic defects, such as extreme reduction of body size and developmental delay.<sup>12, 13</sup>

Two different mechanisms have been proposed for  $\Psi$  synthase action.<sup>14</sup> The “Michael mechanism” (Figure 1.2A) begins with the essential Asp adding to the uracil ring to form a Michael adduct. This step is followed by the cleavage of the glycosidic bond. The detached pyrimidine ring rotates, reconnects to the ribose ring, and is deprotonated by Asp. The “acylal mechanism” (Figure 1.2B) starts either *via* an  $S_N1$  or  $S_N2$  reaction. Asp attacks C1' of the ribose to form the acylal intermediate. The detached pyrimidine ring undergoes rotation, reattachment, and deprotonation. The essential Asp functions as both a nucleophile and a general base in both proposed mechanisms.

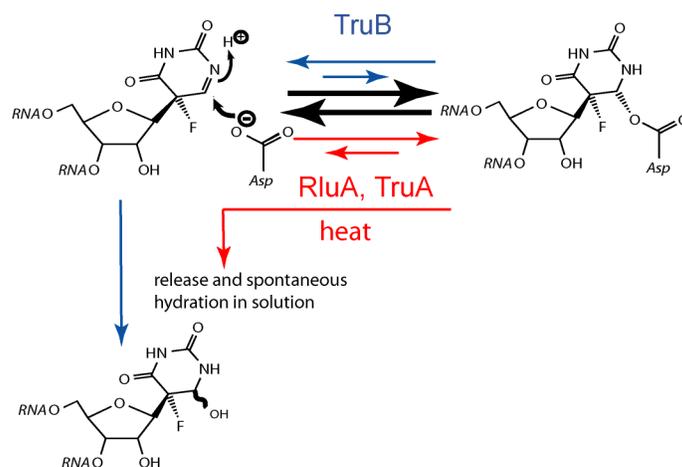
RNA containing 5-fluorouridine, F<sup>5</sup>U, has been used as a mechanistic probe to distinguish the two proposed mechanisms. All  $\Psi$  synthases were reported to



**Figure 1.2: Mechanisms proposed for  $\Psi$  synthases. A, the Michael mechanism; B, the acyl mechanism.**

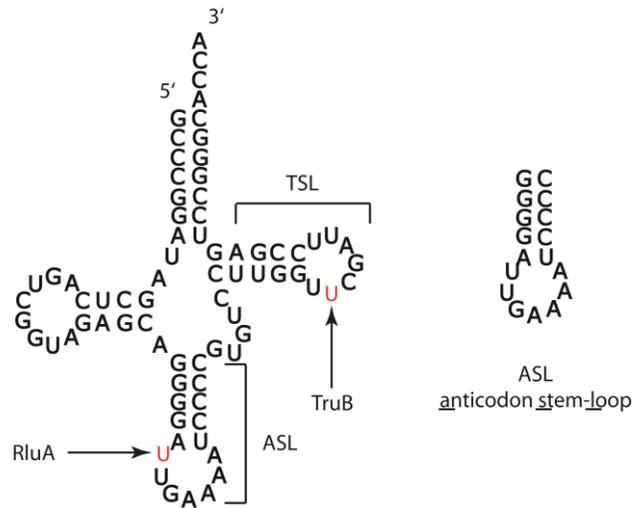
be irreversibly inhibited by [F<sup>5</sup>U]RNA until TruB was examined.<sup>15, 16</sup> RluA and TruA form heat-sensitive and apparently covalent adducts with [F<sup>5</sup>U]RNA,<sup>17, 18</sup> but TruB handles [F<sup>5</sup>U]RNA as a substrate to make hydrated products from F<sup>5</sup>U.<sup>19</sup> Upon heating, hydrated F<sup>5</sup>U products are also formed by the collapse of the adducts of TruA and RluA.<sup>20, 21</sup> The hydrated products were reasonably assumed to derive from the hydrolysis of ester intermediates of the Michael mechanism and thus favor it over the acylal mechanism.<sup>22, 23</sup> An <sup>18</sup>O labeling experiment was carried out in Mueller group (McDonald *et al*, in press, DOI: 10.1021/bi101737z) to monitor whether oxygen from <sup>18</sup>O water was incorporated into the active site Asp-60 of TruA or into the hydrated product of F<sup>5</sup>U by incubating the TruA-[F<sup>5</sup>U]tRNA adduct with 50% [<sup>18</sup>O]water followed by MALDI-MS analysis. The <sup>18</sup>O labeling results clearly show that <sup>18</sup>O was detected in pyrimidine ring of RNA but not in Asp-60 of TruA. Therefore, F<sup>5</sup>U is directly hydrated instead of being generated through ester hydrolysis. This result agrees with similar studies of TruB and RluA reported earlier by Mueller group.<sup>16, 24</sup> However, this result still cannot distinguish the two mechanisms proposed for Ψ synthases because both mechanisms generate rearranged F<sup>5</sup>U.

The behavior of RluA, TruA and TruB can all be accommodated by one scheme (Figure 1.3). By either mechanism, F<sup>5</sup>U is rearranged to the depicted C-glycoside. For RluA and TruA, the equilibrium in Figure 1.3 lies to the right, giving the observed adduct; heating causes the reversal of adduct formation by elimination of the Asp. For TruB, that equilibrium lies far to the left, and no adduct is observed. The different behavior for RluA, TruA and TruB towards [F<sup>5</sup>U]RNA can be rationalized by the relative position of their catalytic Asp to rearranged F<sup>5</sup>U. The results to date, however, still cannot tell the mechanism the Ψ synthases follow.



**Figure 1.3: Schemes showing the different behavior of RluA, TruA and TruB towards [F<sup>5</sup>U]RNA.**

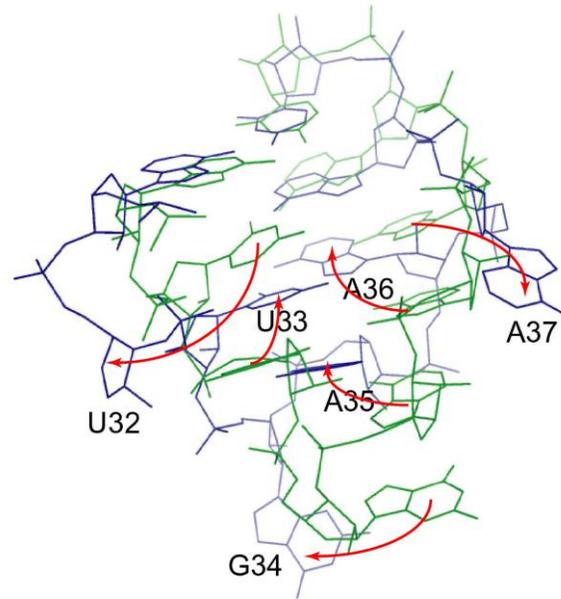
$\Psi$  synthases display diverse substrate specificities. RsuA exclusively modifies a single position, U516, in 16S ribosomal RNA.<sup>17</sup> TruB isomerizes the universally conserved U55 in the T-arm of all the elongator tRNAs in the cell.<sup>25</sup> TruD forms  $\Psi$ 13 in *E. coli* tRNA.<sup>26</sup> TruA is responsible for modifying positions 38, 39 and 40 in tRNA.<sup>27</sup> RluA is capable of  $\Psi$ 32 formation in four tRNAs (Figure 1.4) of *E. coli*: tRNA<sup>Phe</sup>, tRNA<sup>Cys</sup>, tRNA<sub>4</sub><sup>Leu</sup>, and tRNA<sub>5</sub><sup>Leu</sup>.<sup>20, 21</sup> RluA also forms  $\Psi$ 746 in 23S ribosomal RNA.<sup>22</sup> These RNA substrates of RluA have no related global tertiary structures, but they do share local sequence ( $\Psi$ UXXAAA, X is any nucleotide) and in a stem-loop structure at the sites of modification.<sup>23, 28</sup> Pus10 can produce both  $\Psi$ 55 and  $\Psi$ 54, and the two modifications occur independently.<sup>29</sup>



**Figure 1.4:** Cloverleaf schematic of tRNA<sup>Phe</sup> showing uridine residues that are isomerized to Ψ by RluA and TruB. ASL is anticodon stem-loop and TSL is T-armed stem-loop.

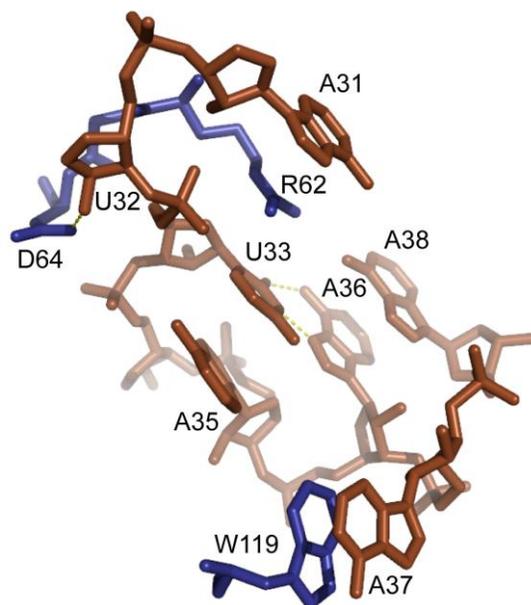
Kinetic studies have been carried out to investigate the molecular recognition of stem-loop RNA by RluA. TruB was the first Ψ synthase for which a structure in complex with RNA was reported, and TruB recognizes its substrate primarily through shape complementarity.<sup>30</sup> To better understand whether the RNA recognition strategies employed by TruB are shared by Ψ synthases belonging to other families, the cocrystal structure of the *E. coli* Ψ synthase RluA bound to an anticodon stem-loop (ASL) RNA was determined by the Ferré-D'Amaré group in 2006 in collaboration with the Mueller group.<sup>25</sup> The cocrystal structure showed that binding to RluA results in a dramatic reorganization of the structure of ASL (Figure 1.5). Free ASL adopts a canonical U turn conformation.<sup>31, 32</sup> Once bound to RluA, ASL undergoes several major changes. U32, G34 and A37 are flipped out from the helical stack. U32 moves into the active site of RluA. A36 occupies the position vacated by A37 and makes a reverse-Hoogsteen base pair with U33. The cocrystal structure (Figure 1.6) shows that the reverse-Hoogsteen base pair formed between U33 and A36 plays an important role in

sequence specificity of RluA. A35 rotates into the helical stack to lie under U33 and the space between A31 and U33 vacated by U32 is taken by Arg-62 of RluA. Arg-62 also makes two water-mediated hydrogen bonds with A38 and a bidentate salt bridge with U33. A37 and A38 are conserved among all known RluA substrates. The structure shows A37 is solvent-exposed on one face, and the other face stacks on the the indole

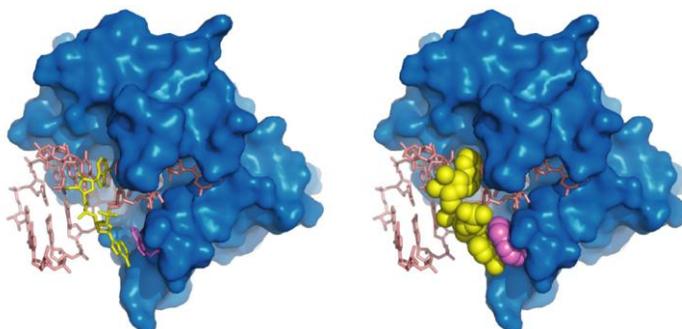


**Figure 1.5: Superposition of the free (green) and RluA-bound (blue) ASLs.**

ring of Trp-119 (Figure 1.7). Based on various interactions seen in the RluA-ASL cocrystal structure, I characterized the enzymatic activity of RluA acting on seven mutated ASLs: U33C, U33G, A36C, A37C, A37G, A38C, and A38G. Also, the reactivity of A36C [ $F^5U$ ]ASL and U33C [ $F^5U$ ]ASL towards RluA was investigated.



**Figure 1.6:** Various interactions observed from the cocrystal structure of RluA (blue) bound with ASL (brown) containing 5-fluorouridine (PDB ID: 2I82). The water-mediated hydrogen bonding between the side chain of Arg-62 and N6 of A38 is not shown because the water molecules are not included in the structure file. Dashed lines denote hydrogen bonds.

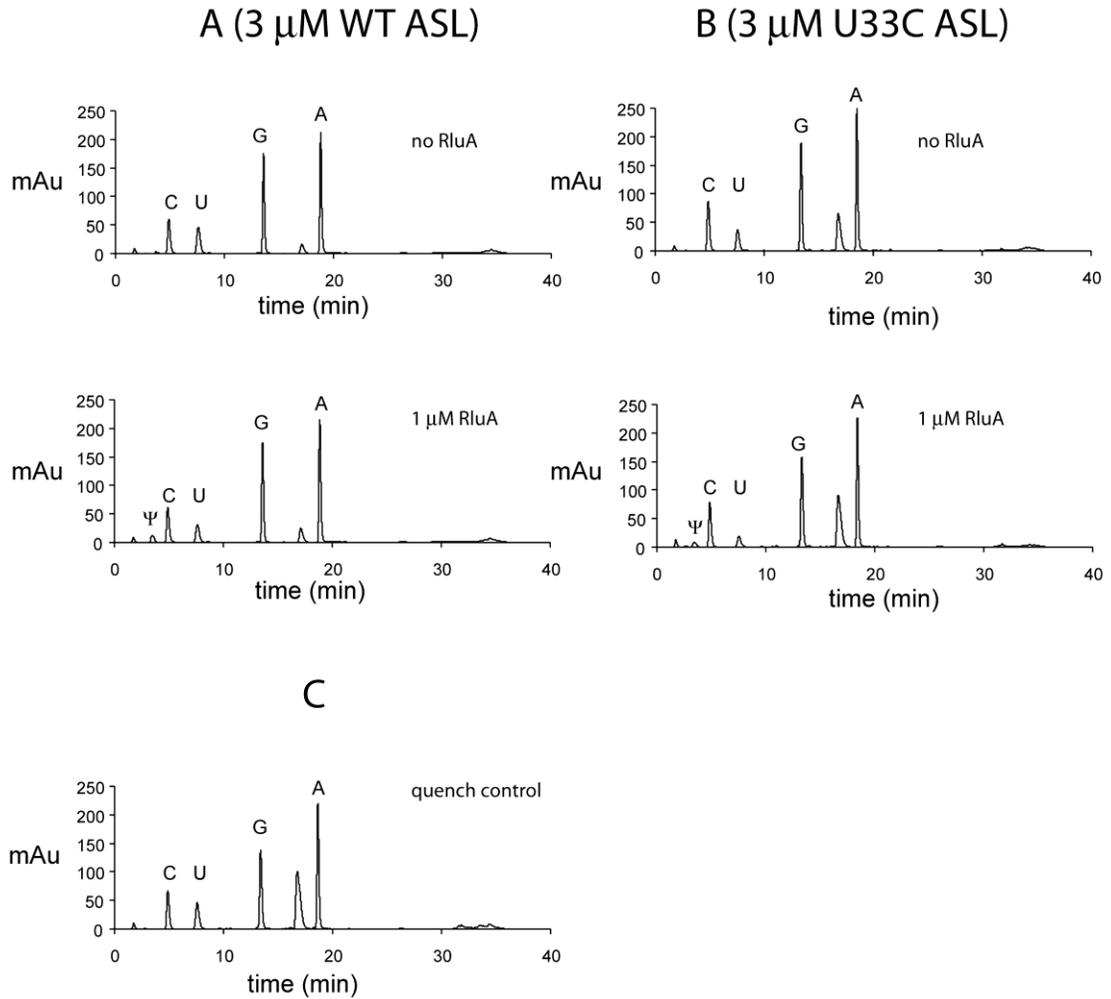


**Figure 1.7:** The relative positions of A37 and A38 (yellow) with RluA (blue). The pink sphere denotes for Trp-119.

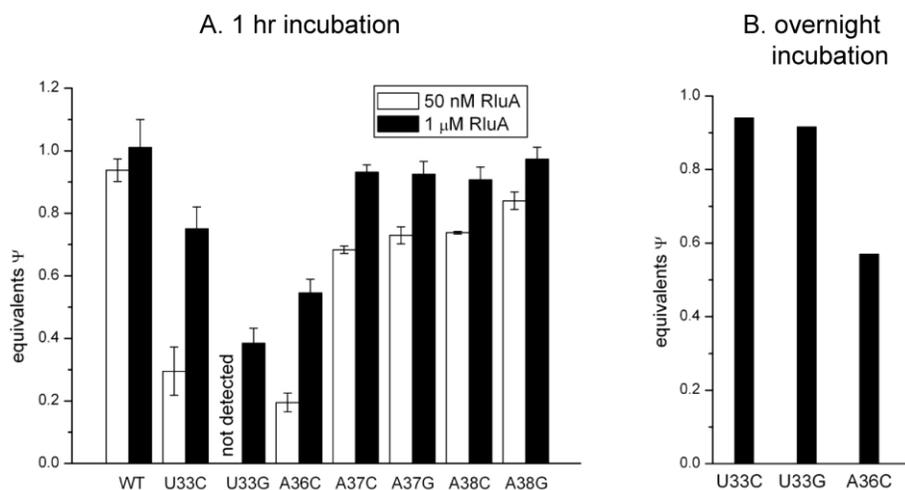
## 1.2 Results and Discussion

### 1.2.1 Preliminary kinetic characterization of substrates ASLs with wild-type RluA

In order to further investigate the relative importance of various interactions seen in the cocrystal structure, single time point enzyme activity assays were carried out with 3  $\mu\text{M}$  ASL at two fixed enzyme concentrations (50 nM and 1  $\mu\text{M}$ ). Using the kinetic parameters  $k_{\text{cat}}$  ( $0.068 \text{ s}^{-1}$ ) and  $K_m$  (308 nM) determined for RluA and ASL and the integrated Briggs-Haldane equation for an irreversible reaction with no product inhibition,<sup>16</sup> the lower concentration of RluA (50 nM) will achieve 99% conversion of U32 to  $\Psi$  in 21.5 min. At 1 h time point, the higher RluA concentration (1  $\mu\text{M}$ ) only requires 1.1 min to achieve 99% conversion. Representative HPLC traces for S1 digestion samples are shown in Figure 1.8.

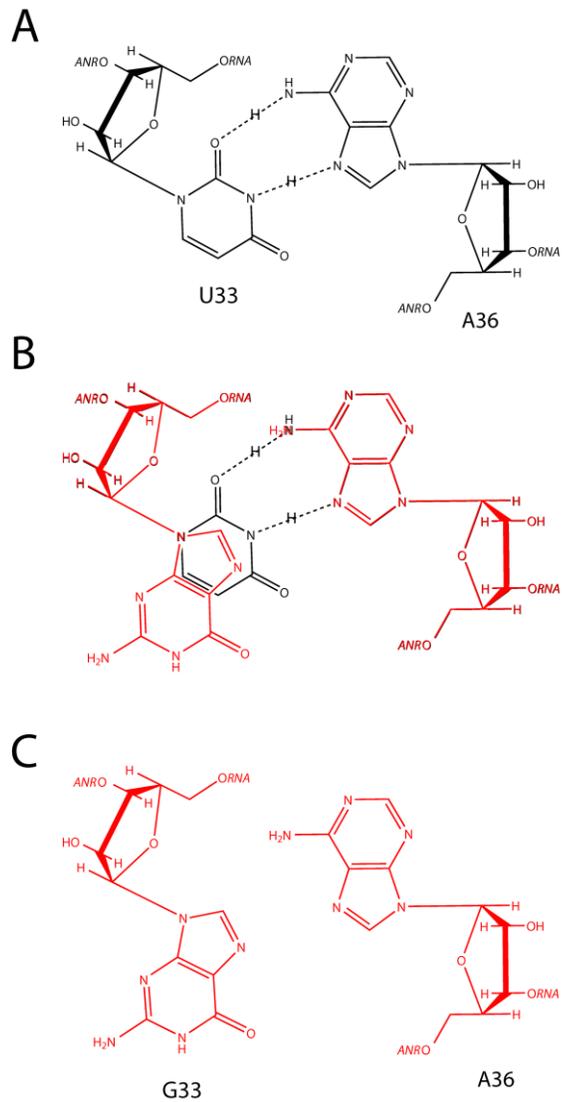


**Figure 1.8:** HPLC analysis of the nuclease S1/alkaline phosphatase digestion of mixture of (A) 3  $\mu\text{M}$  ASL incubated in the absence (above) and presence (below) of 1  $\mu\text{M}$  RluA (B) 3  $\mu\text{M}$  U33C ASL incubated in the absence (above) or presence (below) of 1  $\mu\text{M}$  RluA, and (C) 3  $\mu\text{M}$  ASL incubated with 1  $\mu\text{M}$  RluA that had been heat denatured. All incubations were 1 h at 37  $^{\circ}\text{C}$ .



**Figure 1.9: RluA activity towards ASL and seven variant ASLs. (A) 3  $\mu$ M ASL incubated with 50 nM RluA (white bar) and 1  $\mu$ M RluA (black bar) for 1 h. (B) 3  $\mu$ M ASL variants (U33C, U33G, A36C) incubated with 1  $\mu$ M RluA overnight.**

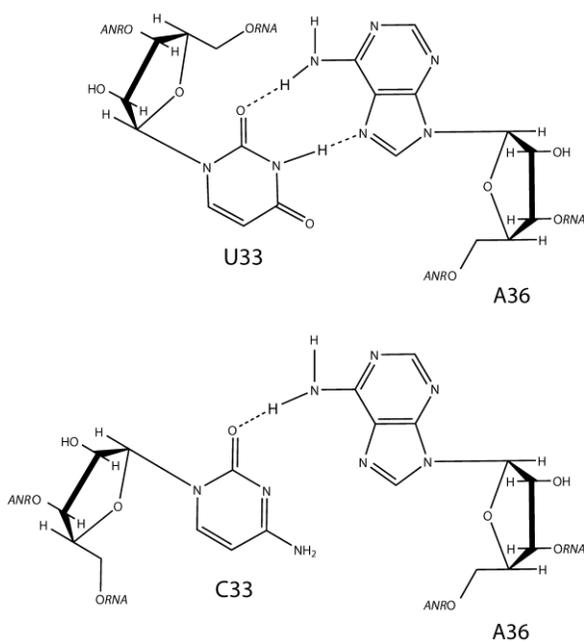
The assay results (Figure 1.9) indicate that the U33 A36 reverse-Hoogsteen base pair plays an important part in substrate recognition by RluA. U33G did not produce any  $\Psi$  with 50 nM RluA and only 0.38 equivalent of  $\Psi$  was produced with 1  $\mu$ M RluA after 1 hr. The effect of U33C change is not as severe as the U33G mutation, but only 0.29 equivalent of  $\Psi$  was produced with 50 nM RluA. A36C was also a poor substrate, producing only 0.2 equivalent of  $\Psi$  with 50 nM RluA. The deleterious effect of mutations at position 33 and 36 are consistent with the expectation from crystal structure. If U33 is replaced by G, the bulky purine ring would be sterically excluded from the helical stem (Figure 1.10), and a C at position 33 can form



**Figure 1.10: The structure comparison between U and G. Panel B is the overlap of panel A and panel C.**

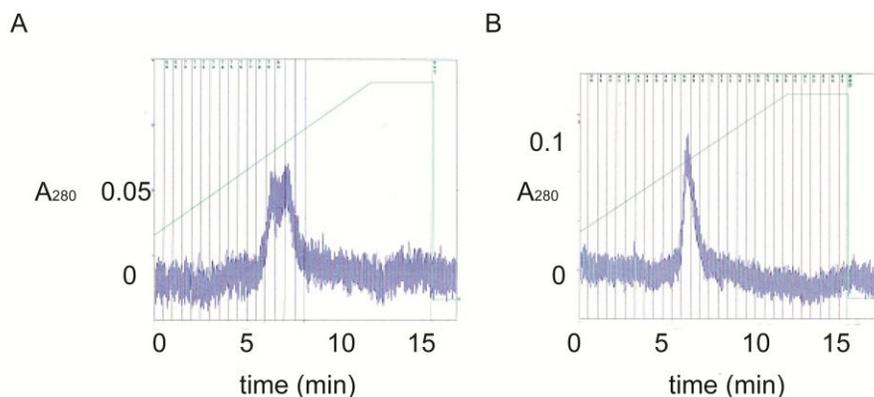
only one hydrogen bond with A36 whereas the normal U33 can make a reverse-Hoogsteen base pair with two hydrogen bonds (Figure 1.11). Replacing A36 with C disallows a reverse-Hoogsteen base pair with U33. In contrast with the dramatic effect

caused by mutations at positions 33 and 36, the alterations at position 37 and 38 were only very mildly deleterious. The results agree with the structural observation that A38 is solvent exposed and A37 is solvent exposed on one face while its other face stacks on the indole ring of Trp-119.



**Figure 1.11: The hydrogen bonds made by U33 A36 and C33 A36.**

### **1.2.2 Assessing the fraction of ASL substrates in an incompetent conformation by chromatography analysis**



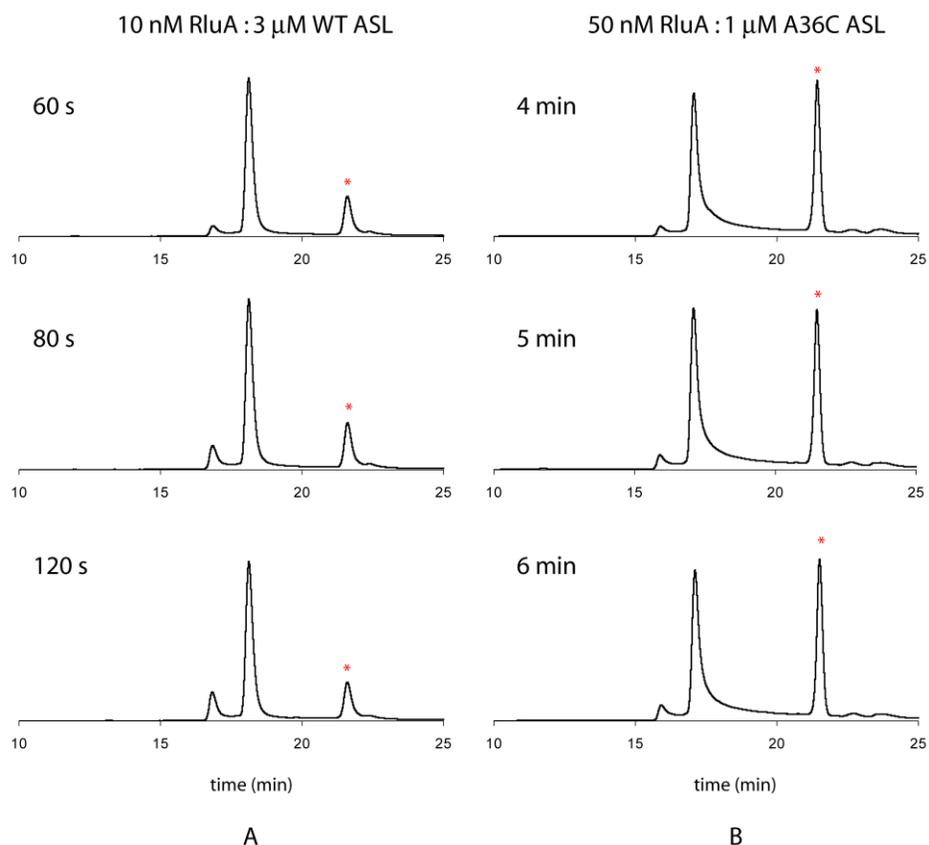
**Figure 1.12: Wild-type ASL on POROS HQ strong anion exchange column before (A) and after (B) heating and cooling.**

The failure to achieve complete turnover of A36C ASL even overnight immediately suggested that a high proportion of the RNA was in a conformation not recognized by RluA, most likely due to a noncanonical secondary structure. To test this hypothesis, ASL was heated to 100 °C for 5 min and then cooled on ice. Figure 1.12 shows that two peaks were present before heating, but only one peak remained after heating, in agreement with the hypothesis that some ASL started in an alternative secondary structure. Subsequently, all ASLs were heated to 100 °C and then cooled on ice immediately before use.

### 1.2.3 Full kinetic characterization of ASLs via HPLC analysis

Product formation was monitored by reverse phase HPLC (Figure 1.13). The product was eluted 1.2 min earlier than substrate ASL. The percentage of the conversion was determined by integrating both peaks and used to calculate the reaction rate. Data was plotted and fit to Briggs-Haldane equation (Figure 1.14). The kinetic data indicated that the C<sub>8</sub> purified ASL is nearly as good a substrate as full-length

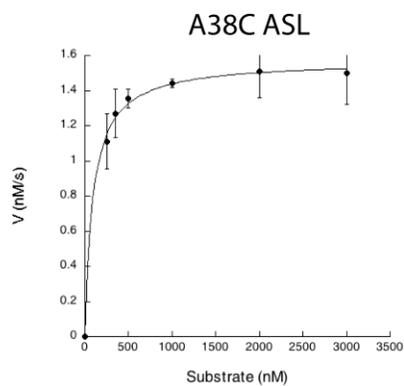
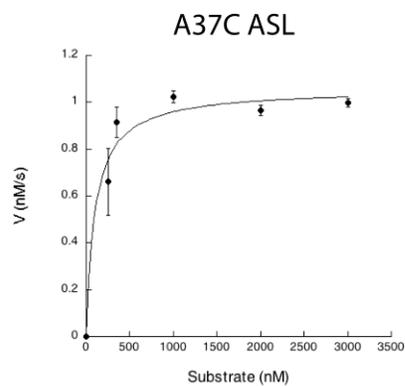
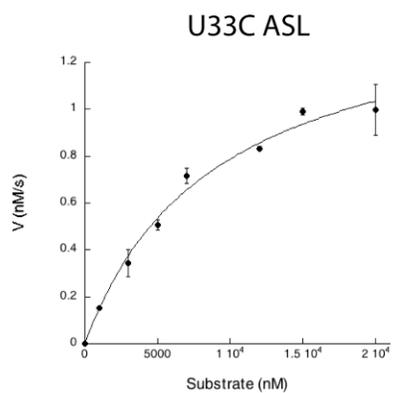
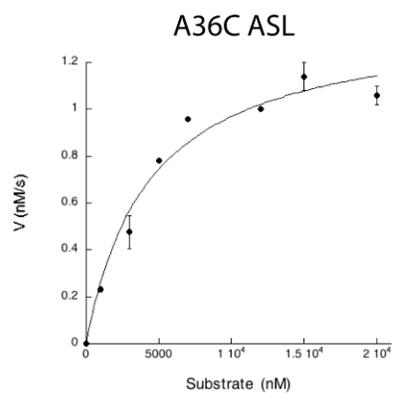
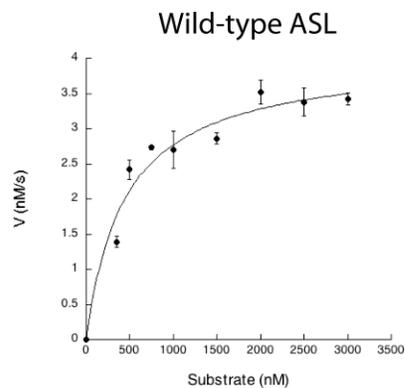
tRNA (Table 1.1), with  $k_{\text{cat}} = 0.402 \text{ s}^{-1}$  and  $K_{\text{m}} = 450 \text{ nM}$  (both 4-fold higher than full-length tRNA).<sup>26</sup> Therefore, the  $k_{\text{cat}}/K_{\text{m}}$  for RluA with C<sub>8</sub>-purified ASL is  $8.9 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ , which is the same as full length tRNA.



**Figure 1.13: Representative HPLC traces of (A) three time points of 3  $\mu\text{M}$  WT ASL incubated with 10 nM RluA (B) three time points of 1  $\mu\text{M}$  A36C ASL incubated with 50 nM RluA.**

C<sub>8</sub>-purified ASL has a 6-fold higher  $k_{\text{cat}}$  compared with the unpurified ASL, which may be caused by mixed inhibition from the impurities present in the unpurified ASL. The ASL and impurities likely bind at different sites on RluA, and the binding of the

impurities slightly affects the binding of ASL. Also, the binding from the impurities may affect the interaction between RluA and its substrate and cause misalignment in the active site. Therefore, this mixed inhibition can result in a slightly decreased  $K_m$  and significantly decreased  $k_{cat}$ .



**Figure 1.14: Briggs-Haldane plots of the kinetic data for RluA with variant ASLs as substrates. The RluA concentration was 10 nM except for A36C ASL and U33C ASL, for which 50 nM RluA was used. Each point is the average of at least two independent determinations. Error bars smaller than the data point are not shown.**

**Table 1.1 Kinetic parameters for RluA with tRNA and variant ASLs**

Substrate	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ (nM)	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{M}^{-1}\text{s}^{-1}$ )
tRNA <sup>a</sup>	$0.099 \pm 0.003$	$108 \pm 20$	$9.2 \times 10^5$
WT ASL <sup>b</sup>	$0.068 \pm 0.003$	$308 \pm 66$	$2.2 \times 10^5$
WT ASL	$0.40 \pm 0.03$	$450 \pm 110$	$8.9 \times 10^5$
A36C ASL	$0.028 \pm 0.002$	$4400 \pm 1000$	$6.4 \times 10^3$
U33C ASL	$0.030 \pm 0.002$	$9000 \pm 1500$	$3.32 \times 10^3$
A37C ASL	$0.11 \pm 0.006$	$100 \pm 41$	$1.05 \times 10^6$
A38C ASL	$0.16 \pm 0.002$	$92 \pm 8.8$	$1.7 \times 10^6$

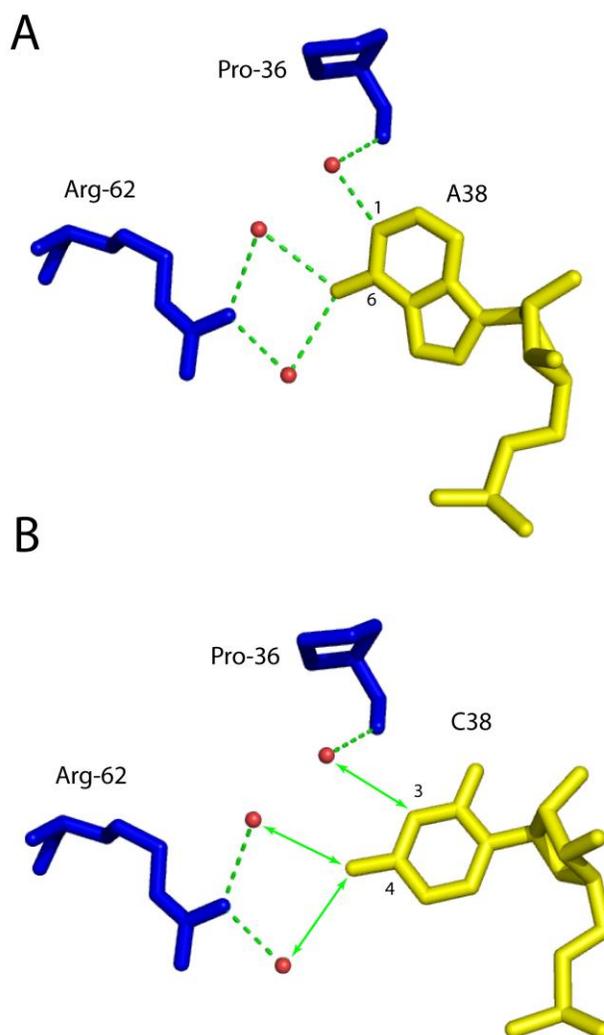
<sup>a</sup> From Ramamurthy *et al.*<sup>26</sup>

<sup>b</sup> From Hamilton *et al.*<sup>16</sup>

Substitution of A36 of ASL with C reduced  $k_{\text{cat}}$  by 14-fold and caused a 10-fold increase in  $K_{\text{m}}$  value. Substitution of U33 of ASL with C reduced  $k_{\text{cat}}$  by 13-fold and caused a 20-fold increase in  $K_{\text{m}}$ . Both A36C and U33C, then, are two orders of magnitude less efficient than WT ASL based on the  $k_{\text{cat}}/K_{\text{m}}$  values. Substitution of A37 or A38 of ASL with C caused almost parallel decreases in  $k_{\text{cat}}$  and  $K_{\text{m}}$  values. This leads A37C ASL to have the same efficiency as WT ASL. A38C ASL behaves as a somewhat better substrate than WT ASL as assessed by  $k_{\text{cat}}/K_{\text{m}}$  values.

Because the enzyme is not diffusion-limited,  $K_{\text{m}}$  largely reflects substrate binding, and the significantly increased  $K_{\text{m}}$  in the altered stem-loop U33C and A36C support the assumptions drawn from the RluA•ASL cocrystal structure that the

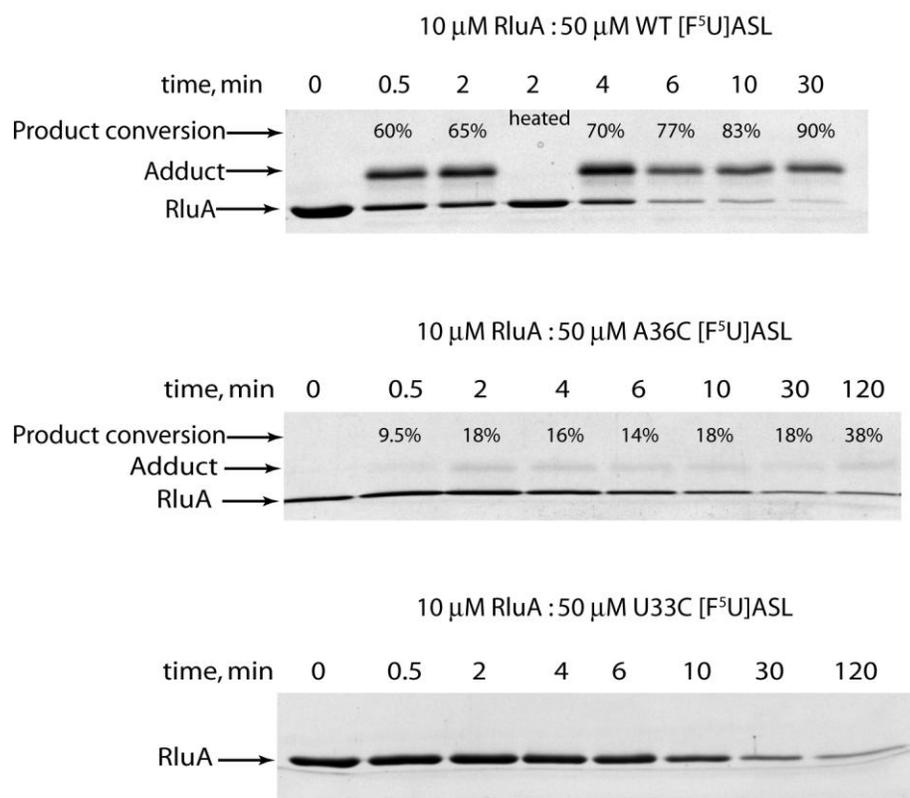
U33 A36 reverse-Hoogsteen base pair plays a key role in substrate recognition by RluA. The reduced  $k_{\text{cat}}$  may arise from the inappropriate positioning of nucleobases in the active site. A37C ASL has the same specificity constant as WT ASL, and A38C ASL is an even more efficient substrate than WT ASL. This is also consistent with the observation that A37 is solvent-exposed on one face and its other face stacks on Trp-119; A37 is therefore not involved in active site interactions. A38 makes three water-mediated hydrogen bonds to RluA. Two of them are from N<sup>6</sup> of A38 to the side chain of Arg-62, and the other is from N1 of A38 to the carbonyl group of Pro-36. When A38 is replaced by C, N<sup>4</sup> and N3 of C38 can also form the three water-mediated hydrogen bonds with RluA (Figure 1.15). However, C38 can fit into the helical stack between A36 and U39 with less steric hindrance compared with A at position 38, which may explain why A37C and A38C have reduced  $K_m$  values compared with WT ASL.



**Figure 1.15:** (A) The water mediated hydrogen bonds formed between Arg-62, Pro-36 of RluA and A38 of ASL. (B) The hypothetical water mediated hydrogen bonds formed between Arg-62, Pro-36 of RluA and C38 of ASL. The pictures are generated by Pymol software<sup>33</sup> using PDB file: 2I82. The depicted water molecules are not included in the pdb file, so they were added manually during the preparation of the graphic based on their position as presented in the report of the cocrystal.<sup>25</sup>

### 1.2.4 Investigation of the reactivity between RluA and two altered [F<sup>5</sup>U]ASLs

To study substrate recognition by RluA further, A36C [F<sup>5</sup>U]ASL and U33C [F<sup>5</sup>U]ASL were incubated with RluA. RluA forms an adduct with wild-type [F<sup>5</sup>U]ASL that can be detected by a gel shift in both SDS-PAGE (protein gel) and urea-PAGE (RNA gel).<sup>15</sup> The adduct can be disrupted by heat, after which free and RNA bands can be seen on gels.<sup>15</sup>



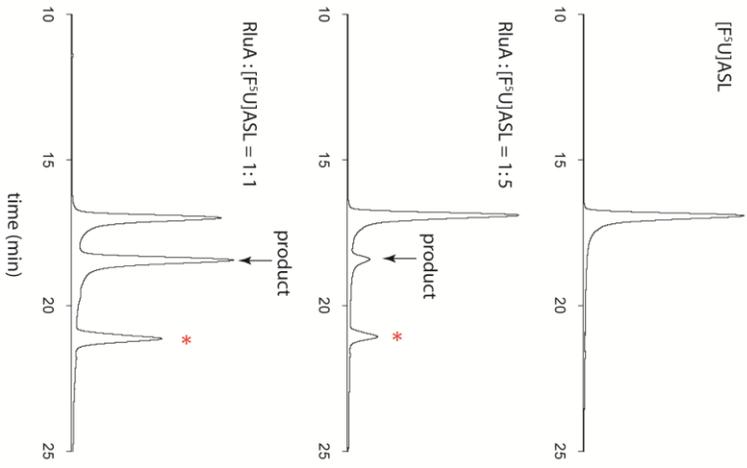
**Figure 1.16: SDS-PAGE showing a time course of adduct formation between RluA and variant [F<sup>5</sup>U]ASLs.**

A very faint shifted band appeared with A36C [F<sup>5</sup>U]ASL, but no shifted band was seen with U33C [F<sup>5</sup>U]ASL; the control incubated with wild-type RluA

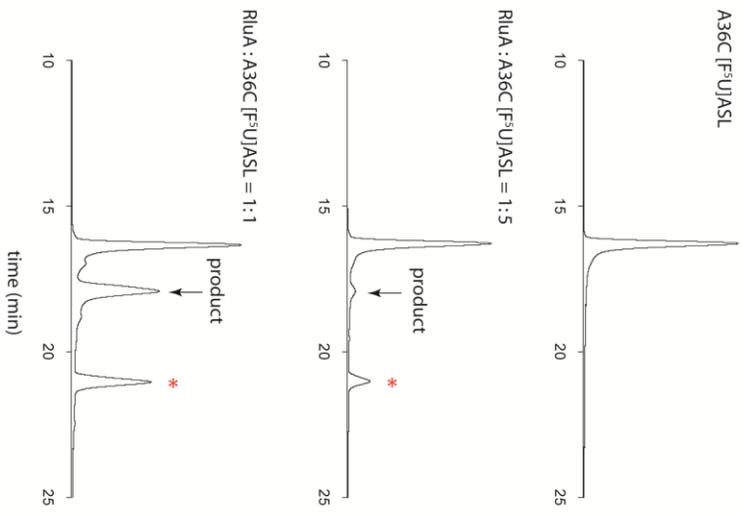
shows normal adduction, verifying the methodology (Figure 1.16). Apparently, adduct formation is much slower with A36C [F<sup>5</sup>U]ASL than with [F<sup>5</sup>U]ASL. This difference may arise because A36C [F<sup>5</sup>U]ASL binds RluA less efficiently than [F<sup>5</sup>U]ASL.

Reverse phase HPLC of intact RNA clearly showed that RluA altered A36C [F<sup>5</sup>U]ASL identically to [F<sup>5</sup>U]ASL (Figure 1.17). Interestingly, RluA-modified U33C [F<sup>5</sup>U]ASL behaved similarly to TruB-modified [F<sup>5</sup>U]TSL, the intact RNA products appearing as doubled peaks.

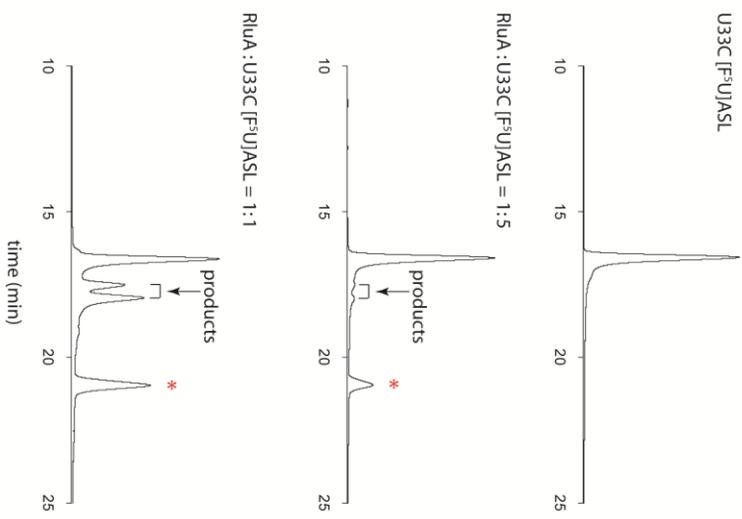
A



B



C

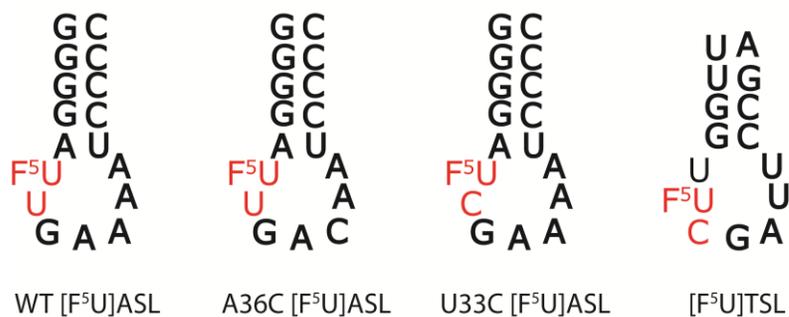


**Figure 1.17: HPLC traces of the intact variant [F<sup>5</sup>U]ASLs and the [F<sup>5</sup>U]ASLs modified after incubation with different amount of RluA.**

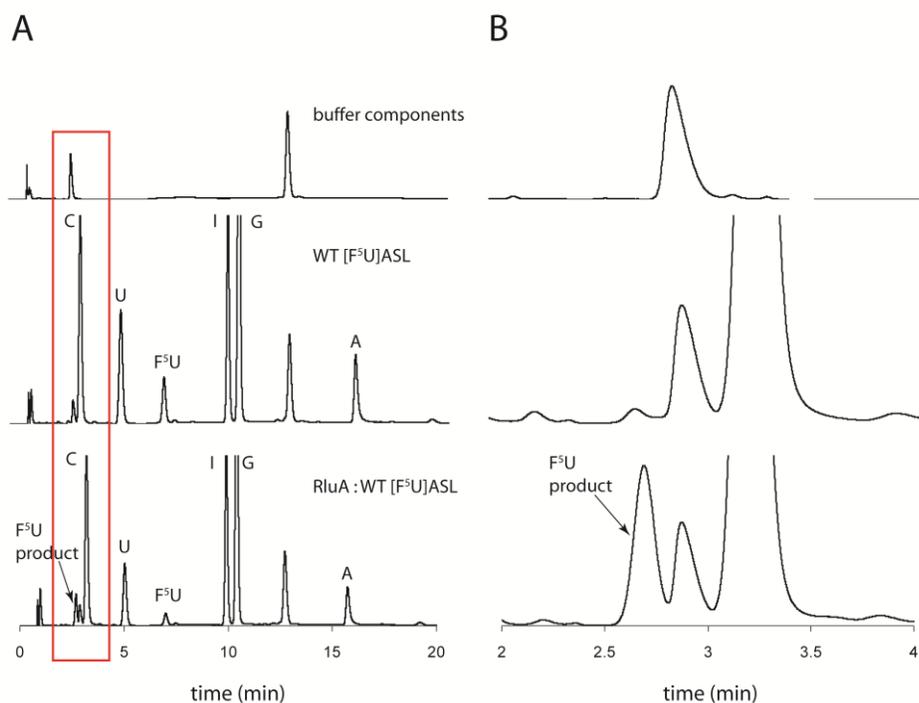
Substrate and product [F<sup>5</sup>U]ASLs were also digested to free nucleosides and analyzed by HPLC (Figure 1.19 – 1.21). RluA converted F<sup>5</sup>U in [F<sup>5</sup>U]ASL and its A36C variant into the same product peak (Figure 1.19 – 1.20).<sup>16</sup> After the digestion of U33C [F<sup>5</sup>U]ASL that had been incubated with RluA, on the other hand, two partially resolved product peaks appeared (Figure 1.21).

It was recently reported by the Mueller group (McDonald *et al*, in press, DOI: 10.1021/bi101737z) that the partially resolved digestion products formed by the action of TruB on [F<sup>5</sup>U]TSL are actually dinucleotides resulting from the inability of S1 nuclease to cleave phosphodiester bonds following a nonplanar nucleobase<sup>27, 28, 30</sup>. The apparently different products observed after the action of RluA on [F<sup>5</sup>U]ASL and the A36C variant versus U33C [F<sup>5</sup>U]ASL, then, can be readily ascribed to the product dinucleotide differing in the nucleoside following the product of F<sup>5</sup>U (Figure 1.18). In [F<sup>5</sup>U]TSL, C follows F<sup>5</sup>U, but in WT [F<sup>5</sup>U]ASL, U follows F<sup>5</sup>U. If this scenario is correct, then WT and A36C [F<sup>5</sup>U]ASL would give identical digestion products (as observed) and U33C [F<sup>5</sup>U]ASL would give the same dinucleotide products as [F<sup>5</sup>U]TSL since C follows F<sup>5</sup>U in both. The overlaid HPLC traces of digested products from [F<sup>5</sup>U]TSL incubated with TruB and U33C [F<sup>5</sup>U]ASL incubated with RluA show that the digested U33C [F<sup>5</sup>U]ASL products coelute with the digested [F<sup>5</sup>U]TSL products (Figure 1.22). The only difference is the proportion between the earlier peak and later peak, with the later peak in the digested [F<sup>5</sup>U]TSL products showing a higher preponderance relative to the earlier peak than the digested U33C [F<sup>5</sup>U]ASL products. Even though the exact structures of the two dinucleotides have not been determined,

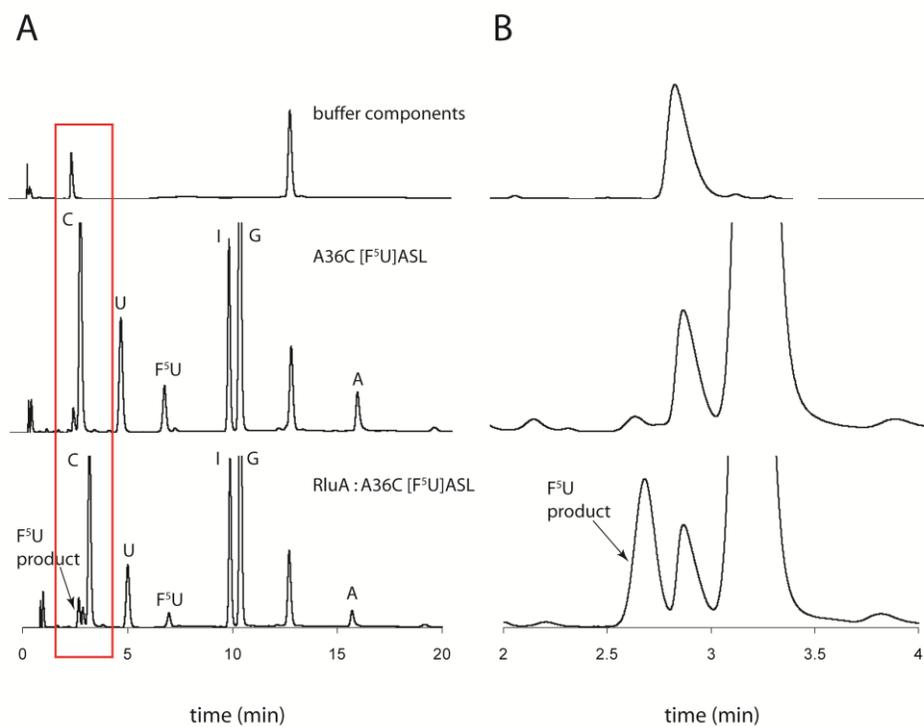
the discrepancy in the relative amounts of the two products may arise from the different equilibrium of the isoforms at the two different active sites or in different rates of conversion of the intermediates to the two products in the different active sites.



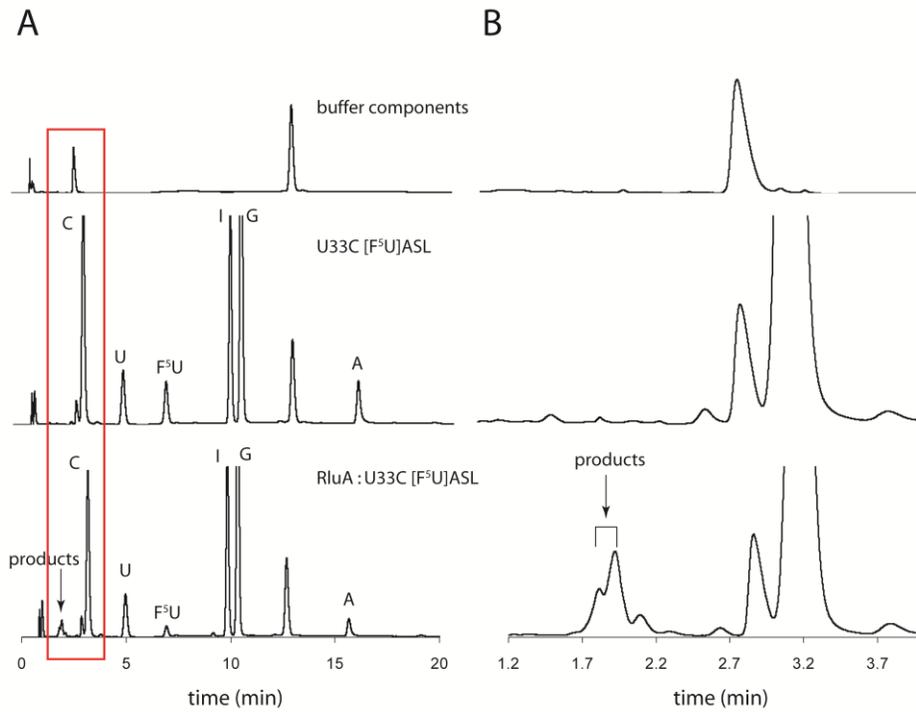
**Figure 1.18: The sequence comparison of [F<sup>5</sup>U]ASL, A36C [F<sup>5</sup>U]ASL, U33C [F<sup>5</sup>U]ASL and [F<sup>5</sup>U]TSL.**



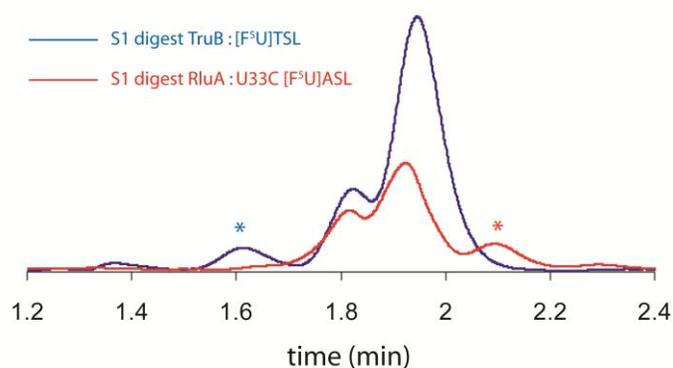
**Figure 1.19: HPLC analysis of the products of F<sup>5</sup>U. (A) buffer components (upper), [F<sup>5</sup>U]ASL incubated in the absence (middle) and presence (bottom) of RluA. (B) Blow-ups of the boxed areas in panel A. Inosine (I) appears as the result of the presence of adenosine deaminase in the S1 nuclease.**



**Figure 1.20: HPLC analysis of the products of F<sup>5</sup>U. (A) buffer components (upper), A36C [F<sup>5</sup>U]ASL incubated in the absence (middle) and presence (bottom) of RluA. (B) Blow-ups of the boxed areas in panel A. Inosine (I) appears as the result of the presence of adenosine deaminase in the S1 nuclease.**



**Figure 1.21: HPLC analysis of the products of  $F^5U$ . (A) buffer components (upper), U33C [ $F^5U$ ]ASL incubated in the absence (middle) and presence (bottom) of RluA. (B) Blow-ups of the boxed areas in panel A. Inosine (I) appears as the result of the presence of adenosine deaminase in the S1 nuclease.**



**Figure 1.22:** Overlaid HPLC traces of nuclease S1/alkaline phosphatase digestion of [F<sup>5</sup>U]TSL incubated with TruB (blue) and U33C [F<sup>5</sup>U]ASL incubated with RluA (red). The blue star denotes the buffer peak that always elutes before the digested [F<sup>5</sup>U]TSL products, and the red star denotes the buffer peak that always elutes after the digested U33C [F<sup>5</sup>U]ASL products.

The kinetic characterization of RluA with several ASL variants was carried out to probe the importance of interactions seen in the cocrystal of RluA and RNA.<sup>25</sup> The recognition of RNA as a substrate by RluA relies critically on the reverse Hoogsteen interaction between U33 and A36, and the replacement of either of them with C dramatically decreased the catalytic efficiency of RluA. However, interactions with A37 and A38 are largely dispensable since either can be substituted with C without compromising the efficiency with which RluA handles ASL as a substrate.

The reactivity between RluA and two variant [F<sup>5</sup>U]ASLs was also investigated. A single product peak suggested that RluA generates a single product of F<sup>5</sup>U, but NMR and MS analysis later revealed that RluA generates two isomeric products with [F<sup>5</sup>U]ASL. (McDonald *et al*, in press, DOI: 10.1021/bi101737z) The difference in HPLC migration for the products of F<sup>5</sup>U from the action of the two enzymes results from a different trailing nucleoside in the dinucleotide products

analyzed. To test this hypothesis, the U following F<sup>5</sup>U in ASL was replaced with C. RluA acts on this U33C [F<sup>5</sup>U]ASL to generate products identical to those of TruB action on F<sup>5</sup>U in its substrate stem-loop as judged by their comigration under the HPLC conditions. When RluA handles A36C [F<sup>5</sup>U]ASL, on the other hand, a single product peak is observed that migrates identically to the product peak from WT [F<sup>5</sup>U]ASL, as expected since U follows F<sup>5</sup>U in both.

Together, then, the data presented in this chapter confirm both the importance of key protein-RNA interactions seen in the cocrystal of RluA and RNA and the role of the trailing nucleoside in the disparate HPLC migration behavior of the products of F<sup>5</sup>U from the action of RluA and TruB.

### **1.3 Materials and methods**

#### **1.3.1 General**

Competent BLR(DE3) pLysS *E. coli* cells were purchased from Novagen (San Diego, CA) and transformed with the plasmid pΨ746, which is pET15b (Novagen) containing the gene *rluA*. Ni-NTA superflow resin was purchased from QIAGEN (Valencia, CA), and POROS SP and POROS HQ resins were purchased from Applied Biosystems (Foster City, CA). The sequences of (wild-type) ASL and the variant substrates examined in this work are:

ASL (G.G.G.G.A.U.U.G.A.A.A.A.U.C.C.C.C),

U33C ASL (G.G.G.G.A.U.C.G.A.A.A.A.U.C.C.C.C),

U33G ASL (G.G.G.G.A.U.G.G.A.A.A.A.U.C.C.C.C),

A36C ASL (G.G.G.G.A.U.U.G.A.C.A.A.U.C.C.C.C),

A37C ASL (G.G.G.G.A.U.U.G.A.A.C.A.U.C.C.C.C),

A37G ASL (G.G.G.G.A.U.U.G.A.A.G.A.U.C.C.C.C),  
A38C ASL (G.G.G.G.A.U.U.G.A.A.A.C.U.C.C.C.C),  
A38G ASL (G.G.G.G.A.U.U.G.A.A.A.G.U.C.C.C.C), and  
[F<sup>5</sup>U]ASL (G.G.G.G.A. **F<sup>5</sup>U**.U.G.A.A.A.U.C.C.C.C),  
A36C [F<sup>5</sup>U]ASL (G.G.G.G.A. **F<sup>5</sup>U**.U.G.A.C.A.A.U.C.C.C.C) and  
U33C [F<sup>5</sup>U]ASL (G.G.G.G.A. **F<sup>5</sup>U**.C.G.A.A.A.U.C.C.C.C) the underlined position  
is the one isomerized by RluA, and bold type indicates the change from parent ASL.  
All oligonucleotides were purchased from Dharmacon (Lafayette, CO). S1 nuclease  
was purchased from Promega (Madison, WI) and Prime RNase inhibitor was  
purchased from Eppendorf (Hauppauge, NY). All HPLC analysis was operated on an  
Agilent 1100 binary systems with a degassing device (Santa Clara, CA).

### 1.3.2 Overexpression and Purification of RluA

For the expression of wild-type RluA, BLR(DE3) pLysS/pΨ746 was used to inoculate LB broth (1 L) in a baffled flask and shaken vigorously at 37 °C. When OD<sub>600</sub> reached 0.4 – 0.6, 1 mM IPTG (4 ml, 100 mM) was added. The cells were harvested when OD<sub>600</sub> reaches 1.1 – 1.2, approximately 3 h after induction. The cells were centrifuged at 10,000g for 30 min at 4 °C and resuspended in 50 mM sodium phosphate buffer, pH 8.0, containing imidazole (10 mM) and sodium chloride (300 mM). The suspensions were quick-frozen and stored at –80 °C.

For purification of RluA, the cells were thawed and sonicated, and the lysate was centrifuged at 20,000g for 30 min to pellet cell debris. A slurry of Ni-NTA superflow resin (3 ml settled resin volume) was added to the supernatant and nutated for 1 – 2 h at 4 °C to bind the His-tagged RluA. The resin was pelleted at 2000g for 5 min, resuspended in lysis buffer (3 ml), and packed into a column (0.9 × 8.5 cm). The

column was washed ( $3 \times 5$  ml) with 50 mM sodium phosphate buffer, pH 8.0, containing imidazole (20 mM) and sodium chloride (300 mM). The bound RluA was eluted ( $2 \times 5$  ml) with 50 mM sodium phosphate buffer, pH 8.0, containing imidazole (250 mM) and sodium chloride (300 mM). The elutions were dialyzed ( $2 \times 2$  h) against 20 mM HEPES buffer (1000 ml), pH 7.5, containing potassium chloride (100 mM), EDTA (5 mM),  $\beta$ -mercaptoethanol (1 mM), and glycerol (5% v/v). The elutions were concentrated using either Centricon-10 or Amicon ultrafiltration units with DiaFlo YM-10 membrane (final volume, 2 ml). The RluA was then further purified by POROS SP cation exchange chromatography, eluting with a linear gradient of potassium chloride (0 – 1 M) in 20 mM HEPES buffer, pH 7.5, containing EDTA (5 mM),  $\beta$ -mercaptoethanol (1 mM), and glycerol (5% v/v). Each protein-bearing fraction (1 ml) was analyzed for purity by SDS-PAGE.

### **1.3.3 Deprotection and Purification of 2'-ACE RNA Oligonucleotides**

All purchased ASLs were deprotected according to the manufacturer's protocol. Lyophilized ASL was dissolved in 100 mM acetate buffer, pH 3.8, made by the addition of tetramethylethylenediamine (TEMED) to a solution of acetic acid, and incubated at 60 °C for 1 h. The deprotected ASL was lyophilized using a Speed Vac concentrator and then dissolved in doubly distilled water (100  $\mu$ l). ASL was precipitated by the addition of 0.1 volume of 3 M sodium acetate buffer, pH 5.2, and then 3 volumes of absolute ethanol. After at least 2 h at -20 °C, the ASL was pelleted by centrifugation and washed twice with cold 70% aqueous ethanol. The pellet was dissolved in double distilled water, and the ASL was injected onto a C<sub>8</sub> HPLC column (CLIQUEUS C8 5 $\mu$ m column, 10  $\times$  250 mm, Higgins Analytical, Mountain View, CA). The ASL was eluted with an acetonitrile gradient using the following program (the

first number is the percentage of aqueous acetonitrile, 40% v/v, in 5 mM ammonium acetate buffer, pH 6.0; the second number is the elapsed time): 0, 0; 0, 3; 5, 8; 15, 18; 15, 23; 30, 26; 50, 27.5; 50, 29; 100, 30; 100, 31; 0, 32; 0, 40. The ASL was fractionally collected, and all of the RNA-bearing fractions were taken to dryness in a Speedvac. Each ASL fraction was redissolved in doubly distilled water and analyzed using an analytical C<sub>18</sub> column (CLPEUS C18 5 μm column, 4.6 × 150 mm, Higgins Analytical, Mountain View, CA). Pure ASL fractions were combined and washed 1000-fold with doubly distilled water using a Centricon-3 device.

#### **1.3.4 Preliminary kinetic characterization for wild-type RluA**

RluA activity was assayed by incubating ASL and variants (U33C, U33G, A36C, A37C, A37G, A38C and A38G ASL) with the enzyme and quantifying Ψ formation by the total digestion of the RNA to nucleosides followed by HPLC analysis.<sup>34, 35</sup> The reaction mixture (230 μl) was 50 mM HEPES buffer, pH 7.5, containing ammonium chloride (100 mM), DTT (5 mM), EDTA (1 mM), RNase inhibitor (30 units), and stem-loop RNA (3 μM). After pre-equilibration at 37 °C for 5 min, reactions were initiated by the addition of RluA (50 nM or 1 μM). After 1 h at 37 °C, the reaction was quenched by denaturing the enzyme in a 70 °C water bath for 5 min; control reactions in which ASL was added to heat denatured RluA showed that the heating inactivated enzyme activity below levels detectable by this assay method. Precipitated protein was pelleted by centrifugation, and the supernatant was transferred to a fresh tube and heated to 100 °C for 5 min. After the addition of concentrated nuclease buffer to make the solution 100 mM sodium acetate buffer, pH 4.5, containing sodium chloride (560 mM) and zinc chloride (9 mM), S1 nuclease (100 units) was added to digest the RNA. After 1 h at 37 °C, the reaction mixture was

heated to 100 °C for 5 min. Additional S1 nuclease (100 units) was added, and the reaction mixture was incubated for 1 h at 37 °C. After 1 h at 37 °C, alkaline phosphatase (5 units) was added and the reaction mixture was incubated at 37 °C for 1 h. The digested ASLs were analyzed by HPLC over a C<sub>18</sub> column (Eclipse XDB-C18 5 µm column, 4.6 × 250 mm, Agilent). The free nucleosides were eluted with an acetonitrile gradient using the following program (the first number is the percentage of aqueous acetonitrile, 40% v/v, in 25 mM ammonium acetate buffer, pH 6.0; the second number is the elapsed time in minutes): 0, 0; 0, 3; 5, 8; 15, 18; 15, 23; 30, 26; 50, 27.5; 50, 29; 100, 30; 100, 31; 0, 32; 0, 40. Nucleosides were detected by A<sub>260</sub>, and the peak areas were integrated and adjusted by the extinction coefficient for each nucleoside to determine the relative concentrations of Ψ, U, C, G, and A,<sup>36</sup> from which the extent of conversion from U to Ψ was calculated. ASL and each variant ASL was analyzed before incubation with RluA, and the ratio of the nucleosides agreed within 10% of the expected values. Each stem-loop was assayed in duplicate at both 50 nM and 1 µM RluA, and the two values for the extent of conversion of U to Ψ agreed within 10% except for U33C and A36C ASL, which showed lower extents of conversion and agreed within 20% and 15%, respectively. After incubation with 1 µM RluA for 1 h, only U33C, U33G, and A36C ASL generated substantially less than one equivalent of Ψ. To test whether the incomplete conversion was due to slow reaction or a substantial fraction of incompetent stem-loop substrate, they were incubated overnight with 1 µM RluA and subjected to the total digestion/HPLC analysis. To test the noncanonical secondary structure hypothesis mentioned in section 1.1.2, ASL was heated to 100 °C for 5 min and then cooled on ice. The ASL was analyzed by anion

exchange chromatography over a column of POROS HQ resin. ASL was eluted with a linear gradient of potassium chloride (0–1.5 M) in 50 mM Tris•HCl buffer, pH 7.6.

### 1.3.5 Kinetic assays

The reaction mixture (260 – 1400  $\mu$ l) was 50 mM HEPES buffer, pH 7.5, containing ammonium chloride (100 mM),  $\beta$ -mercaptoethanol (5 mM), EDTA (1 mM), RNase inhibitor (30 units), and ASLs (250 nM–20  $\mu$ M). DTT was replaced by  $\beta$ -mercaptoethanol in this assay because impurities in the DTT eluted at almost the same time as the ASL peak in HPLC analysis. After pre-equilibration at 37  $^{\circ}$ C for 5 min, reactions were initiated by the addition of RluA (to 10 nM or 50 nM). It had previously been found that sodium phosphate buffers inhibit RluA,<sup>37</sup> so aliquots (80  $\mu$ l–400  $\mu$ l) were removed at various times and immediately quenched into an equal volume of 500 mM sodium phosphate buffer, pH 7.5, containing sodium chloride (500 mM). Samples were analyzed by reverse phase HPLC over a C<sub>18</sub> column (CLIQUEUS C18 5  $\mu$ m column, 4.6  $\times$  150 mm, Higgins Analytical, Mountain View, CA). Substrate and product peaks were integrated to determine the fraction of product, and the product concentration was calculated based on the total ASL concentration. Data were plotted and fit to the Briggs-Haldane equation using KaleidaGraph (v. 3.6, Synergy Software, Reading, PA).

### 1.3.6 Band shift analysis

[F<sup>5</sup>U]ASL and its A36C and U33C variants (50  $\mu$ M) were incubated with RluA (10  $\mu$ M) in assay buffer at 37  $^{\circ}$ C. Aliquots (10  $\mu$ l) were removed at various time points and quenched into an equal volume of 2 $\times$  Laemmli sample buffer.

### **1.3.7 Intact and S1 digestion HPLC analysis**

[F<sup>5</sup>U]ASL and its A36C and U33C variants (50 μM or 10 μM) were incubated with RluA (10 μM) in assay buffer at 37 °C. Aliquots (50 μl) were removed at various time points and quenched immediately into an equal volume of 500 mM sodium phosphate buffer, pH 7.5, containing sodium chloride (500 mM). Samples were analyzed by reverse phase HPLC over a C<sub>18</sub> column as described in section 1.4.5, or the samples were subjected to digestion with S1 nuclease and alkaline phosphatase coupled with HPLC analysis as described in section 1.4.4.

## REFERENCES

1. Grosjean, H.; Benne, R., Modification and editing of RNA. ASM Press: Washington, DC, 1998; p xiii, 596 p.
2. Ofengand, J.; Bakin, A.; Wrzesinski, J.; Nurse, K.; Lane, B. G., The pseudouridine residues of ribosomal RNA. *Biochem Cell Biol* 1995, 73 (11-12), 915-24.
3. Motorin, Y.; Helm, M., tRNA stabilization by modified nucleotides. *Biochemistry* 2010, 49 (24), 4934-44.
4. Charette, M.; Gray, M. W., Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* 2000, 49 (5), 341-51.
5. Davis, D. R., Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res* 1995, 23 (24), 5020-6.
6. Mueller, E. G.; Ferré-D'Amaré A. R., Pseudouridine Formation, the Most Common Transglycosylation in RNA, In *DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution* (Grosjean, H., Ed.). Landes Bioscience: Austin, TX, 2009; p p 363-376.
7. Mueller, E. G., Chips off the old block. *Nat Struct Biol* 2002, 9 (5), 320-2.
8. Huang, L.; Pookanjanatavip, M.; Gu, X.; Santi, D. V., A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. *Biochemistry* 1998, 37 (1), 344-51.
9. McCleverty, C. J.; Hornsby, M.; Spraggon, G.; Kreuzsch, A., Crystal structure of human Pus10, a novel pseudouridine synthase. *J Mol Biol* 2007, 373 (5), 1243-54.
10. Heiss, N. S.; Knight, S. W.; Vulliamy, T. J.; Klauck, S. M.; Wiemann, S.; Mason, P. J.; Poustka, A.; Dokal, I., X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* 1998, 19 (1), 32-8.

11. Yu, Y. T.; Shu, M. D.; Steitz, J. A., Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J* 1998, 17 (19), 5783-95.
12. Phillips, B.; Billin, A. N.; Cadwell, C.; Buchholz, R.; Erickson, C.; Merriam, J. R.; Carbon, J.; Poole, S. J., The Nop60B gene of *Drosophila* encodes an essential nucleolar protein that functions in yeast. *Mol Gen Genet* 1998, 260 (1), 20-9.
13. Giordano, E.; Peluso, I.; Senger, S.; Furia, M., minifly, a *Drosophila* gene required for ribosome biogenesis. *J Cell Biol* 1999, 144 (6), 1123-33.
14. Gu, X.; Liu, Y.; Santi, D. V., The mechanism of pseudouridine synthase I as deduced from its interaction with 5-fluorouracil-tRNA. *Proc Natl Acad Sci U S A* 1999, 96 (25), 14270-5.
15. Spedaliere, C. J.; Mueller, E. G., Not all pseudouridine synthases are potently inhibited by RNA containing 5-fluorouridine. *RNA* 2004, 10 (2), 192-9.
16. Hamilton, C. S.; Greco, T. M.; Vizthum, C. A.; Ginter, J. M.; Johnston, M. V.; Mueller, E. G., Mechanistic investigations of the pseudouridine synthase RluA using RNA containing 5-fluorouridine. *Biochemistry* 2006, 45 (39), 12029-38.
17. Wrzesinski, J.; Bakin, A.; Nurse, K.; Lane, B. G.; Ofengand, J., Purification, cloning, and properties of the 16S RNA pseudouridine 516 synthase from *Escherichia coli*. *Biochemistry* 1995, 34 (27), 8904-13.
18. Nurse, K.; Wrzesinski, J.; Bakin, A.; Lane, B. G.; Ofengand, J., Purification, cloning, and properties of the tRNA psi 55 synthase from *Escherichia coli*. *RNA* 1995, 1 (1), 102-12.
19. Kammen, H. O.; Marvel, C. C.; Hardy, L.; Penhoet, E. E., Purification, structure, and properties of *Escherichia coli* tRNA pseudouridine synthase I. *J Biol Chem* 1988, 263 (5), 2255-63.
20. Sprinzl, M.; Horn, C.; Brown, M.; Ioudovitch, A.; Steinberg, S., Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* 1998, 26 (1), 148-53.
21. Horie, N.; Yamaizumi, Z.; Kuchino, Y.; Takai, K.; Goldman, E.; Miyazawa, T.; Nishimura, S.; Yokoyama, S., Modified nucleosides in the first positions of the anticodons of tRNA<sup>4Leu</sup> and tRNA<sup>5Leu</sup> from *Escherichia coli*. *Biochemistry* 1999, 38 (1), 207-17.

22. Wrzesinski, J.; Nurse, K.; Bakin, A.; Lane, B. G.; Ofengand, J., A dual-specificity pseudouridine synthase: an *Escherichia coli* synthase purified and cloned on the basis of its specificity for psi 746 in 23S RNA is also specific for psi 32 in tRNA(phe). *RNA* 1995, 1 (4), 437-48.
23. Raychaudhuri, S.; Niu, L.; Conrad, J.; Lane, B. G.; Ofengand, J., Functional effect of deletion and mutation of the *Escherichia coli* ribosomal RNA and tRNA pseudouridine synthase RluA. *J Biol Chem* 1999, 274 (27), 18880-6.
24. Spedaliere, C. J.; Ginter, J. M.; Johnston, M. V.; Mueller, E. G., The pseudouridine synthases: revisiting a mechanism that seemed settled. *J Am Chem Soc* 2004, 126 (40), 12758-9.
25. Hoang, C.; Chen, J.; Vizthum, C. A.; Kandel, J. M.; Hamilton, C. S.; Mueller, E. G.; Ferré-D'Amaré A. R., Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure. *Mol Cell* 2006, 24 (4), 535-45.
26. Ramamurthy, V.; Swann, S. L.; Spedaliere, C. J.; Mueller, E. G., Role of cysteine residues in pseudouridine synthases of different families. *Biochemistry* 1999, 38 (40), 13106-11.
27. Romier, C.; Dominguez, R.; Lahm, A.; Dahl, O.; Suck, D., Recognition of single-stranded DNA by nuclease P1: high resolution crystal structures of complexes with substrate analogs. *Proteins* 1998, 32 (4), 414-24.
28. Weinfeld, M.; Liuzzi, M.; Paterson, M. C., Selective hydrolysis by exo- and endonucleases of phosphodiester bonds adjacent to an apurinic site. *Nucleic Acids Res* 1989, 17 (10), 3735-45.
29. Gurha, P.; Gupta, R., Archaeal Pus10 proteins can produce both pseudouridine 54 and 55 in tRNA. *RNA* 2008, 14 (12), 2521-7.
30. Weinfeld, M.; Soderlind, K. J.; Buchko, G. W., Influence of nucleic acid base aromaticity on substrate reactivity with enzymes acting on single-stranded DNA. *Nucleic Acids Res* 1993, 21 (3), 621-6.
31. Kim, S. H.; Suddath, F. L.; Quigley, G. J.; McPherson, A.; Sussman, J. L.; Wang, A. H.; Seeman, N. C.; Rich, A., Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* 1974, 185 (149), 435-40.

32. Robertus, J. D.; Ladner, J. E.; Finch, J. T.; Rhodes, D.; Brown, R. S.; Clark, B. F.; Klug, A., Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* 1974, 250 (467), 546-51.
33. The PyMOL Molecular Graphics System, V. r. p., Schrödinger, LLC.
34. Buck, M.; Connick, M.; Ames, B. N., Complete analysis of tRNA-modified nucleosides by high-performance liquid chromatography: the 29 modified nucleosides of *Salmonella typhimurium* and *Escherichia coli* tRNA. *Anal Biochem* 1983, 129 (1), 1-13.
35. Gehrke, C. W.; Kuo, K. C.; McCune, R. A.; Gerhardt, K. O.; Agris, P. F., Quantitative enzymatic hydrolysis of tRNAs: reversed-phase high-performance liquid chromatography of tRNA nucleosides. *J Chromatogr* 1982, 230 (2), 297-308.
36. Dawson, R. M. C., *Data for biochemical research*. 3rd ed.; Clarendon Press: Oxford, 1986; p xii, 580 p.
37. Greco, T. M., *Senior Thesis in Chemistry and Biochemistry*. University of Delaware: Newark, DE 2003.

## Chapter 2

### CHEMICALLY UBIQUITINATED PCNA AS A PROBE FOR EUKARYOTIC TRANSLESION DNA SYNTHESIS

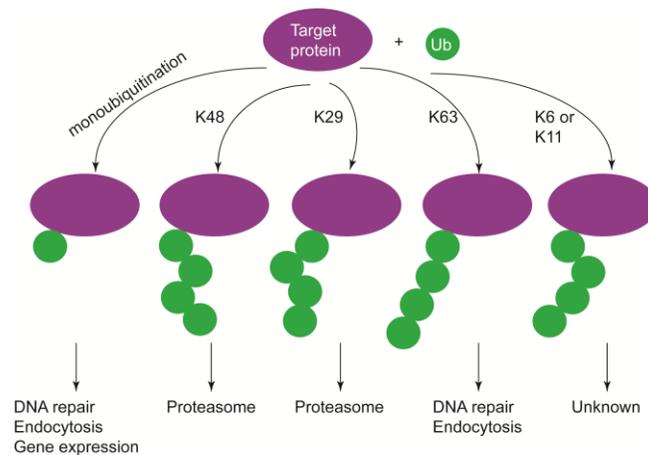
#### 2.1 Introduction

##### 2.1.1 Ubiquitin is a versatile signaling molecule

Ubiquitin is an 8.6 kDa small protein in which the C-terminal glycine can be covalently conjugated to the target protein via an isopeptide bond. Ubiquitination requires a cascade of three enzyme reactions to conjugate the ubiquitin moiety to the  $\epsilon$ -amino of lysine side chain site-specifically. Ubiquitination was initially discovered as a signaling mechanism for proteasome-mediated protein degradation.<sup>1</sup> In recent years, it has become clear that ubiquitination plays a far broader role in eukaryotes, controlling a vast network of physical and functional interactions in living cells. New pathways regulated by ubiquitin are being discovered at a fast pace in many different aspects of cell biology, including DNA damage repair/tolerance, signal transduction, transcription, nuclear transport, autophagy and immune response.<sup>2</sup> Moreover, a group of ubiquitin-like proteins (Ubls) have been discovered in various fundamentally important cellular pathways.<sup>3,4</sup>

The scale and importance of ubiquitin modification rivals protein phosphorylation, by far the most studied post-translational protein modification. However, ubiquitination possesses higher complexity and greater versatility than phosphorylation, in view that ubiquitin can be conjugated to proteins in the forms of either monoubiquitin or polyubiquitin. Further, polyubiquitin can be formed with

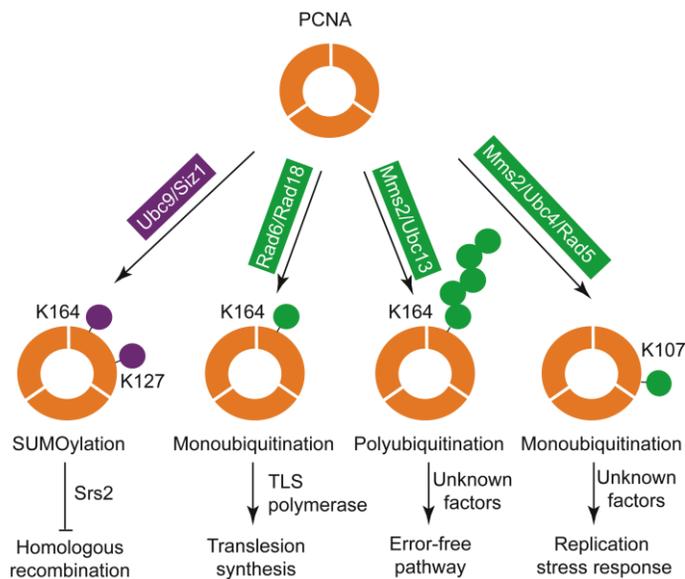
distinct chain structures linked through one of the seven lysine residues (K6, K11, K27, K29, K33, K48 and K63) on ubiquitin (Figure 2.1).<sup>5</sup> It is generally believed that K48 or K29 linked polyubiquitin chain targets protein for proteasome degradation.<sup>6</sup> However, monoubiquitination or polyubiquitination formed through other lysine linkages have nonproteolytic functions, such as endocytosis and DNA damage tolerance. Ubiquitination, like phosphorylation, is a reversible process. Deubiquitinating enzymes (DUB) are responsible for cleaving the isopeptide bond between ubiquitin and a substrate protein, or between ubiquitin molecules within a polyubiquitin chain. The important cellular role of DUBs is manifested by the large number of DUBs identified in eukaryotes (there are approximately 100 putative DUBs in the human proteome).<sup>7</sup>



**Figure 2.1: Cellular pathways associated with different forms of ubiquitin modification. K48 or K29 linked polyubiquitin chain targets protein for proteasome degradation. Monoubiquitination or polyubiquitination formed through other lysine linkages have nonproteolytic functions, such as endocytosis and DNA repair.**

### 2.1.2 Ubiquitination and SUMOylation of PCNA

Proliferating cell nuclear antigen (PCNA) is a processivity clamp protein for the replicative DNA polymerases and physically interacts with various proteins involved in DNA replication, repair, cell cycle regulation and chromatin assembly.<sup>8</sup> It has emerged as an important mediator of eukaryotic DNA damage tolerance. PCNA is one of the prominent examples that undergo ubiquitination in response to genotoxic stress. PCNA undergoes monoubiquitination, polyubiquitination and SUMOylation (Figure 2.2). Strikingly, the same lysine residue (K164) of PCNA is



**Figure 2.2:** PCNA ubiquitination and SUMOylation pathways. PCNA (orange ring) can be modified by SUMO (purple sphere), monoubiquitin or ubiquitin chains (green spheres) at different sites. SUMOylation on PCNA can occur either on K164 or K127. Monoubiquitination was observed on K164. Polyubiquitination on PCNA K164 is critical for an error-free DNA damage tolerance pathway.

modified by either a single ubiquitin (monoubiquitination) or by a polyubiquitin chain (polyubiquitination). Notably, the polyubiquitin chain is linked through the K63 of ubiquitin, which is distinct from the K48-linked polyubiquitin chain involved in proteasome-mediated protein degradation. Available evidences suggest that monoubiquitination and polyubiquitination of PCNA functions to channel the DNA damage response to different branches by recruiting specific replication factors. The stalling of a replication fork triggers monoubiquitination of PCNA at Lys164, which in turn recruits a specialized DNA polymerase for the effective translesion synthesis. Remarkably, many TLS polymerase contain either ubiquitin binding zinc finger (UBZ) domain or ubiquitin binding motif (UBM) that helps to recruit the specific TLS polymerase to the lesion site.<sup>9-13</sup> In contrast, polyubiquitination of PCNA is thought to trigger the error-free lesion bypass and restart the stalled replication fork.<sup>14-22</sup> SUMOylation of PCNA occurs at Lys164 and to a less extent at Lys127.<sup>22</sup> SUMOylation of PCNA was found to recruit the helicase Srs2 and exert an anti-recombinogenic effect during DNA synthesis to prevent untimely recombination.<sup>23, 24</sup> Moreover, SUMOylation of PCNA at Lys127 was shown to prevent the binding of Eco1 through its PIP motif, thus abolishing sister-chromatid cohesion during S phase.<sup>25</sup> Rad6 and Rad18 are the ubiquitin conjugating enzyme and the ubiquitin ligase that catalyze the monoubiquitination of PCNA on Lys164.<sup>22, 26</sup> Polyubiquitination of PCNA requires Mms2-Ubc13 and the ubiquitin ligase Rad5.<sup>14-16</sup>

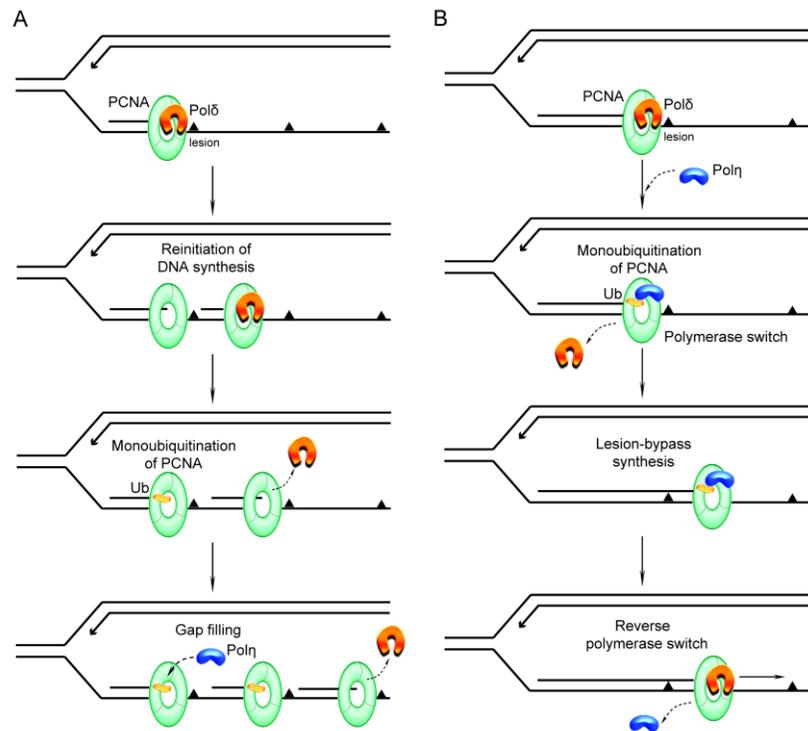
Rad6/Rad18 is not the only ubiquitin ligase complex that monoubiquitinates PCNA. Recently, RNF8 was discovered to catalyze monoubiquitination of PCNA with the assistance of the E2 ubiquitin conjugating enzyme, UbcH5c.<sup>27</sup> RNF8 also mediated polyubiquitination of PCNA in the presence

of Mms2-Ubc13. Depletion of RNF8 by shRNA suppressed PCNA monoubiquitination when the cells were treated with UV or MNNG (N-methyl-N'-nitro-N-nitrosoguanidine).<sup>27</sup> It was also found that CRL4<sup>Cdt2</sup> monoubiquitinates PCNA at Lys164 both *in vivo* and *in vitro*.<sup>28</sup> Interestingly, CRL4<sup>Cdt2</sup>-mediated monoubiquitination of PCNA occurs in the absence of external DNA damage. It is likely that CRL4<sup>Cdt2</sup> regulates PCNA-dependent TLS to cope with stresses accompanying DNA replication. Moreover, monoubiquitination of PCNA at Lys107 was recently discovered and linked to the DNA-ligase I deficiency in yeast cells.<sup>29</sup> Instead of activating TLS, K107 ubiquitination might be responsible for providing a signal for DNA damage on newly synthesized DNA. The authors proposed that monoubiquitination of PCNA may be a "DNA damage code" that allows cells to respond to different types DNA defects during DNA replication.

### **2.1.3 Translesion synthesis (TLS)**

Cells have developed elegant mechanisms to replicate its genome with high efficiency and fidelity. The replicative DNA polymerase can copy genomic DNA with high fidelity and processivity.<sup>30-32</sup> Faithful replication of the genome is a formidable task in view that genomic DNA is subjected to constant damage by both endogenous and exogenous agents. Most DNA lesions can be repaired by the cellular DNA repair machineries. However, in some cases the damage is not repairable or repair cannot be carried out in a timely manner. Consequently, the unrepaired DNA lesion blocks the progression of the replication fork. Prolonged stalling leads to fork collapse, thus jeopardizing the genome integrity. This is particularly perilous during the S phase, in which timely replication of genome is of paramount importance for the normal cell cycle progression and cell division. Mechanisms of moving replication

fork past unrepaired lesion exist in cells to overcome the blockade. Three mechanisms have been proposed to account for the DNA damage tolerance. In my dissertation, I'm going to focus on translesion synthesis (TLS). In TLS, specialized DNA polymerases are required to carry out lesion bypass synthesis.<sup>33</sup> Compared to the replicative DNA polymerases, these specialized polymerases operate with low fidelity and processivity on normal DNA template. Therefore, translesion synthesis is considered to be an error-prone mechanism.



**Figure 2.3: Two models of translesion synthesis. A, postreplicative gap-filling model. B, polymerase switch model. Note: this figure was generated by William Bozza and Dr. Yongxing Ai.**

The molecular details of translesion synthesis have been subjected to intensive studies in recent years. However, despite great progress in our understanding of the structure and function of both prokaryotic and eukaryotic TLS polymerases, the timing of cellular TLS remains to be fully elucidated.<sup>34, 35</sup> Two prevalent models were proposed, *i.e.* the postreplicative gap-filling model and the polymerase switch model (Figure 2.3). In the former model, lesion-bypass synthesis is uncoupled from the normal genomic DNA replication. Specifically, when the replicative DNA polymerase encounters a DNA lesion in either leading or lagging strands, the replicative DNA polymerases skip the lesion and reinitiate DNA synthesis downstream of the DNA damage site (Figure 2.3A). The TLS polymerases are later recruited to the lesion site to promote gap-filling DNA synthesis. In the latter model (Figure 2.3B), when the DNA replication fork encounters a lesion that stalls the replicative polymerase, a TLS polymerase will be recruited to the lesion site. Following a polymerase switch process, the newly recruited TLS polymerase can synthesize past the DNA lesion. After the lesion-bypass synthesis, the replicative polymerase replaces the TLS polymerase and normal DNA synthesis resumes. In this model the normal and lesion-bypass DNA syntheses are coupled and the replicative and TLS polymerases cooperate in moving the replication fork past the lesion.

#### **2.1.4 Chemical ubiquitination**

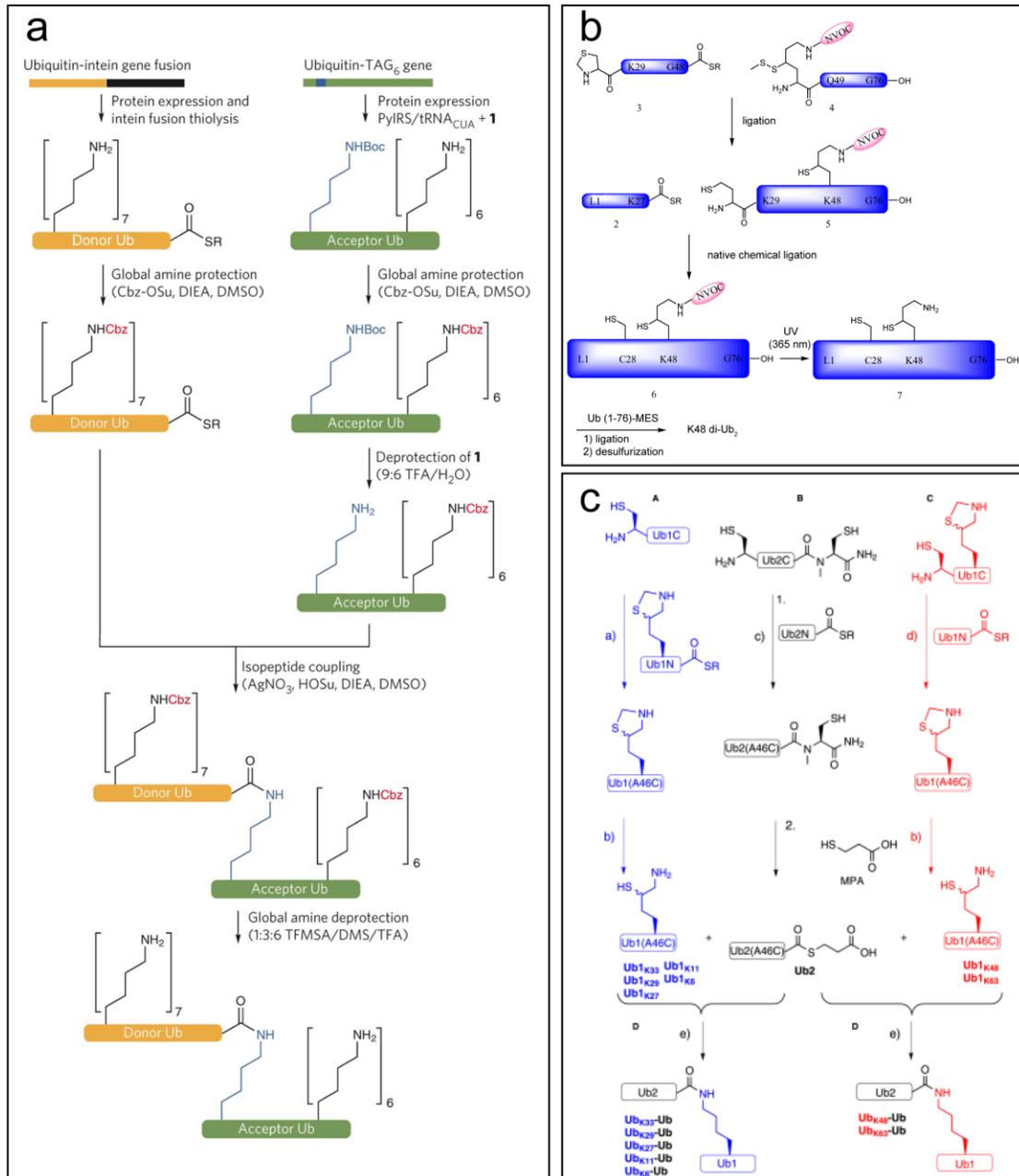
Although *in vitro* enzymatic ubiquitination of various proteins has been demonstrated, the yield of enzymatic ubiquitination is in general low. Given the importance of obtaining large amounts of ubiquitinated proteins, more efficient chemical approaches for protein ubiquitination are required. Several groups have reported chemical methods for preparing diubiquitin with different linkages. Virdee

*et al* combined genetic code expansion and chemoselective protein chemistry to generate two previously inaccessible homogeneous ubiquitin chains, K6 and K29 diubiquitin.<sup>36</sup> This specific chemical ligation was achieved in five steps (Figure 2.4a), by (i) protecting all lysines and the N terminus on the donor ubiquitin and all the lysines but one and the N terminus on the acceptor ubiquitin, resulting a single free amine on the acceptor ubiquitin, (ii) activating the C terminus of the donor ubiquitin as a thioester, (iii) forming a specific isopeptide bond between donor and acceptor ubiquitins by selectively coupling the free amine and the thioester, (iv) removing all the protecting groups on other lysines and N terminus, (v) refolding the ubiquitin chain. Also, the K6 and K29 linked diubiquitins were tested on a panel of deubiquitinases. Surprisingly, they found ovarian tumor (OTU) family deubiquitinase TRABID has a 40-fold specificity on cleaving K29 diubiquitin than K63 linked diubiquitin. K63 diubiquitin was reported to have higher specificity over K48 linkage by TRABID.

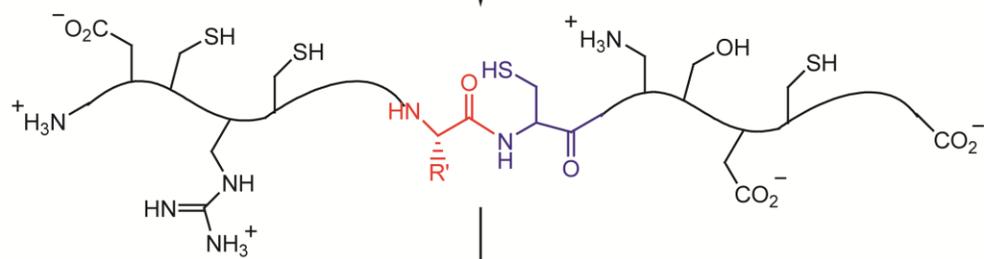
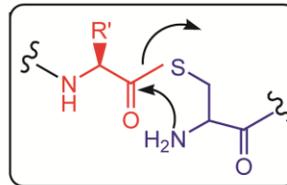
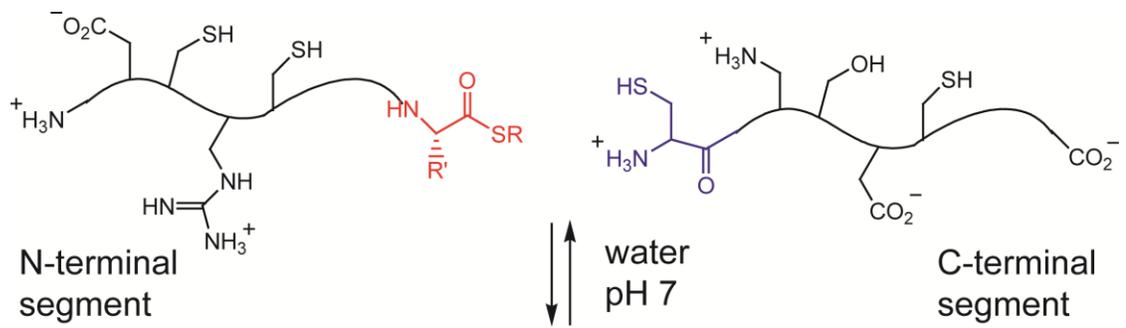
The Liu group used native chemical ligation (Figure 2.5) to synthesize K48 linked diubiquitins.<sup>37-39</sup> The first step involves installation of 4-mercaptolysine at position 48 (K48(4-SH)) of the monoubiquitin (Figure 2.4b). A photolabile protection group, *o*-nitroveratryloxycarbonyl (NVOC) was used to protect the K48(4-SH) amine. K48(4-SH)-containing ubiquitin ( ubiquitin 6 in Figure 2.4b) was generated using C-to-N sequential ligation with Ala28Cys and K48(4-SH) as two ligation pieces. The NVOC group was later removed by 365 nm UV irradiation for ubiquitin 6 to produce ubiquitin 7. Monoubiquitin 7 with free  $\epsilon$ -NH<sub>2</sub> on K48 was then reacted with ubiquitin thioester Ub(1-76)-MES. Ub(1-76)-MES was generated by thiolysis of ubiquitin-intein with sodium mercaptoethanesulfonate (MESNa). In the final step, a

desulfurization was performed for the ligation product to convert Cys28 to Ala and K48(4-SH) to Lys.

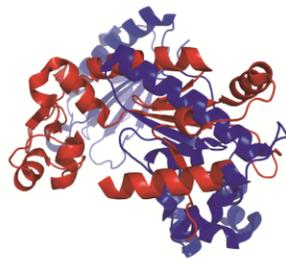
The Brik group also utilized native chemical ligation to synthesize all seven homogeneous ubiquitin chains in which K6, K27, K29 and K33 linked diubiquitins were not reported before.<sup>40</sup> Three main steps were used to assemble the specific linked diubiquitins. First, native chemical ligation was used to synthesize the proximal Ub molecule (Ub1) with the  $\delta$  mercaptolysine at the desired position by dividing the polypeptide sequence of Ub into two fragments (Figure 2.4c, panel A), the C-terminal peptide Ub1C (Ub46-76) and the N-terminal peptide Ub1N (Ub1-45), to which the orthogonally protected thiazolidine  $\delta$  mercaptolysine was incorporated at a specific location of the isopeptide bond (Figure 2.4c, panel A and C). In the second step, Ub1 was conjugated with the Ub thioester (Ub2) through the  $\delta$  mercaptolysine. Ub2 was chemically synthesized from two peptides (Ub2C and Ub2N) wherein the C-terminal fragment Ub2C would bear the N-methylcysteine as a masked thioester group (Figure 2.4c, panel B). Last, Ub1 was ligated with Ub2 (Figure 2.4c, panel D) and followed by a desulfurization step to convert the Cys46 residues in each Ub1–Ub2 analogue into their native Ala46.



**Figure 2.4:** Three different strategies to generate homogenous diubiquitins. (a) Genetic engineering and synthetic chemistry was used to create K6 and K29 di-Ub. (b) Native chemical ligation to synthesize K48 di-Ub. (c) Native chemical ligation to synthesize all seven homogenous di-Ub. Figure a and c were originally published in references.<sup>36, 40</sup> Permission is granted by publishers.

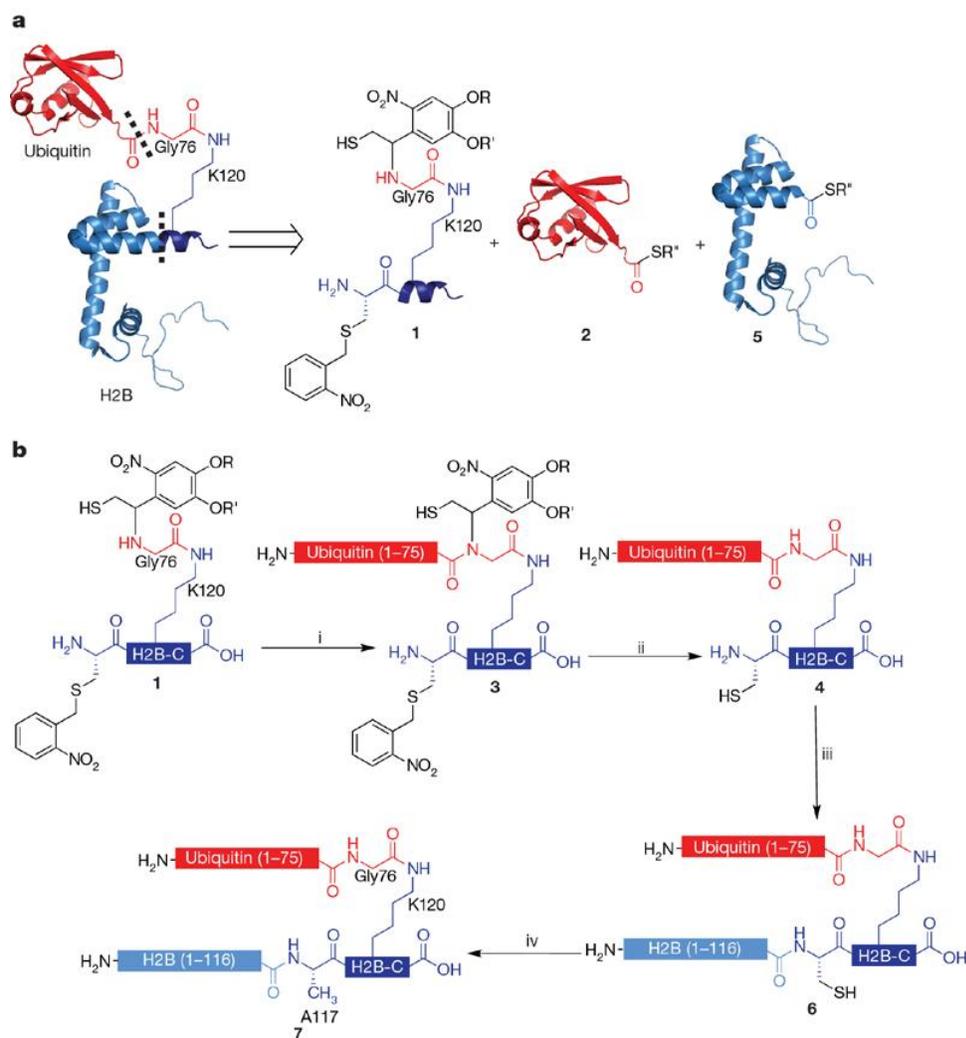


Folding gives active proteins



**Figure 2.5: Native chemical ligation. Unprotected peptide segments are reacted by means of reversible thiol/thioester exchange to give thioester-linked initial reaction products. Uniquely, the thioester-linked intermediate involving an N-terminal Cys residue (boxed) is able to undergo nucleophilic rearrangement by a highly favored intramolecular mechanism; this step is irreversible (under the conditions used) and gives a polypeptide product that is linked by a native amide (i.e. peptide) bond. Only a single reaction product is obtained, even in the presence of additional Cys residues in either segment. The product polypeptide is subsequently folded to give the desired synthetic protein molecule.**

Ubiquitin is relatively small in size and possesses exceptional stability in organic solvent. These properties made it possible to prepare diubiquitin following the above described chemical methods. More challenges exist for chemical ubiquitination of other target proteins that are either too large for peptide synthesis or unstable in organic solvent. McGinty *et al* recently reported the chemical ubiquitination of histone H2B following a semisynthetic approach.<sup>41</sup> The reported chemical method involves two orthogonal expressed protein ligation in combination with photolytic deprotection and Raney-nickel-mediated desulphurization (Figure 2.6). A final yield of 20% in H2B ubiquitination was afforded. This reported approach is particularly suitable for monoubiquitination of histone H2B because the ubiquitination site in histone is very close to the C-terminus of the 131 a. a. polypeptide. Thus a short peptide can be conveniently synthesized and chemically derivatized for the subsequent expressed protein ligation. However, this method is not suitable for large protein and ubiquitin-site in the interior of the protein. Our target, proliferating cell nuclear antigen (PCNA), is a homotrimeric, toroid-shaped protein with 258 amino acids in each subunit.<sup>42</sup> The site of ubiquitination (Lys164) is distant from both termini, which presents further challenges for chemical ubiquitination.



**Figure 2.6:** (a) Retrosynthetic analysis of uH2B synthesis. uH2B was generated by a three-piece ligation with several polypeptides. (b) Synthetic scheme for the ubiquitination of H2B. This figure was originally published in the reference.<sup>41</sup> Permission is granted by the publisher.

### **2.1.5 New approaches for PCNA ubiquitination is imperative**

Enzymatic monoubiquitination of PCNA has been achieved in a reconstituted system that contains the *S. cerevisiae* Uba1, Rad6, Rad18, ubiquitin, RFC and duplex DNA.<sup>17, 43, 44</sup> Uba1 is the ubiquitin activating enzyme; Rad6 and Rad18 are ubiquitin conjugating enzyme and ubiquitin ligase respectively. Rad18, like many other ubiquitin ligases, is highly specific for PCNA. Besides the requirement of Uba1 and Rad6/Rad18, PCNA needs to be loaded onto a primer/template DNA blocked at both ends for efficient ubiquitination. Replication factor C (RFC), a heteropentameric protein complex, is required to load PCNA onto DNA with the consumption of ATP.<sup>45, 46</sup> The overall yield of enzymatic ubiquitination of PCNA is usually low, which stems from the requirement of a large number of purified enzymes and the relatively low catalytic turnover of the Rad6/Rad18 complex. The complexity of enzymatic PCNA ubiquitination and the low yield of pure product have severely hindered our understanding of the many processes involving PCNA ubiquitination.

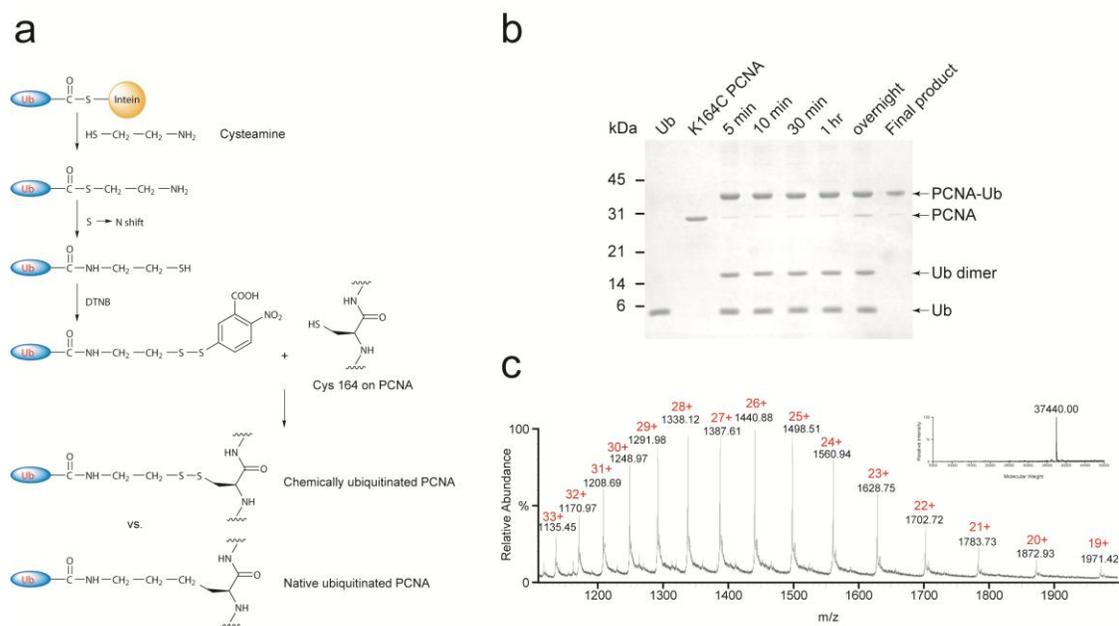
## **2.2 Results**

### **2.2.1 Chemical ubiquitination of PCNA**

We engineered a PCNA mutant with a unique cysteine introduced to replace Lys164 (for brevity we refer to the mutant K164C/C22S/C30S/C62S/C81S PCNA as K164C PCNA). The mutant PCNA retains normal activity as the wild-type PCNA in supporting the processive DNA synthesis by Pol $\delta$ .<sup>47</sup> In the first step of chemical ubiquitination of PCNA, the ubiquitin C-terminal carboxylate was activated by the formation of a thioester through intein chemistry (Figure 2.7a). Cysteamine was used to cleave ubiquitin off the chitin column. Following an intramolecular S to N

shift, a modified ubiquitin (designated as Ub-SH) was generated with a unique thiol introduced at its C-terminus (native ubiquitin contains no cysteine residue). Ub-SH was then treated with 5, 5'-dithiobis-(2-nitrobenzoic acid) (DTNB). This step activates the C-terminal thiol of Ub-SH and facilitates the subsequent disulfide bond formation with Cys164 in PCNA. The reaction proceeded at a fast rate and reached completion within one hour (Figure 2.7b). In a typical reaction ca. 4.5 mg of monoubiquitinated PCNA was prepared with a final yield of 80%. The reactions usually generate 89-94% ubiquitinated PCNA subunits in the final product.

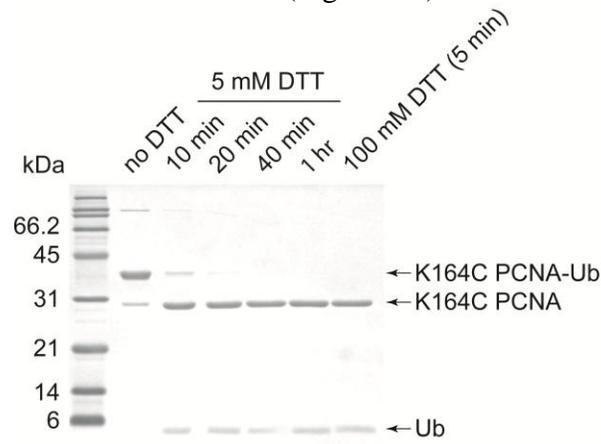
Electrospray ionization (ESI) mass spectrometry was used to determine the molecular weight of the chemically monoubiquitinated PCNA subunit (Figure 2.7c). The measured molecular weight of 37,440.00 Da is almost identical to the theoretical molecular weight (37,440.60 Da).



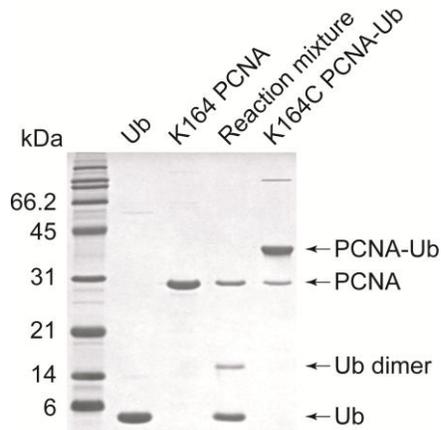
**Figure 2.7: Chemical ubiquitination of PCNA. (a) The disulfide-mediated ligation between ubiquitin-cysteamine, generated by cleaving ubiquitin off chitin resin with cysteamine, and K164C PCNA. The chemically ubiquitinated PCNA mimics the native ubiquitinated PCNA. (b) Denaturing SDS-PAGE gel showing the progression of the chemical ligation reaction and the final product following an anion exchange chromatography. (c) The multiple charge states of the ubiquitinated PCNA determined by electrospray ionization mass spectrometry. The inset shows the deconvoluted peaks of K164C PCNA-Ub (37,440 Da).**

### 2.2.2 Disulfide linkage characterization

The nature of the linkage was confirmed by the complete cleavage of the disulfide by DTT (Figure 2.8). A complete cleavage of the disulfide bond converted K164C PCNA-Ub back to free K164C PCNA and ubiquitin was observed. We also carried out a control reaction using a cysteine-free PCNA mutant that retains Lys164 and no ubiquitylated PCNA was formed (Figure 2.9).



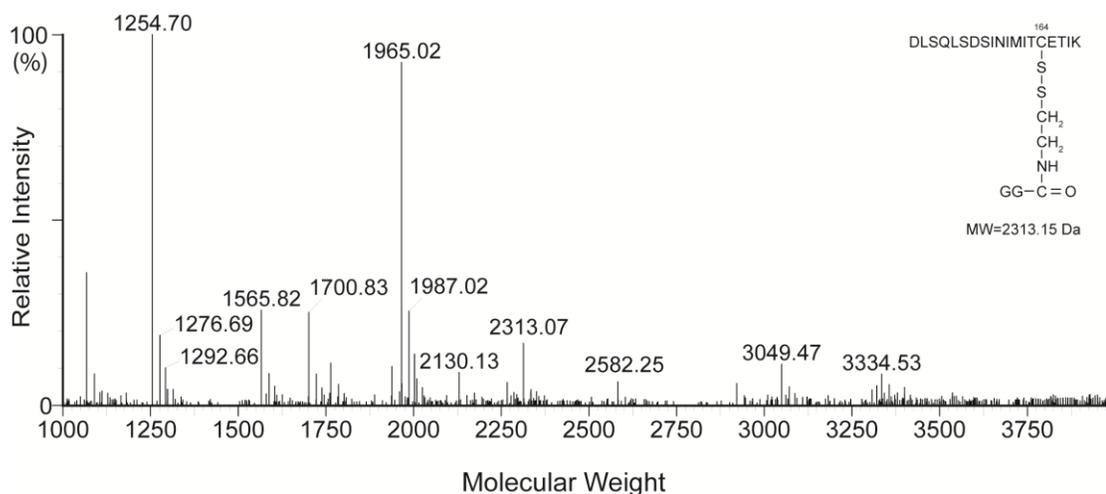
**Figure 2.8: Denaturing SDS-PAGE analysis of the chemically ubiquitinated PCNA (K164C PCNA-Ub) treated with 5 mM DTT at different time points (up to one hour) or 100 mM DTT for 5 min.**



**Figure 2.9:** A control chemical ubiquitination reaction using a cysteine-free PCNA mutant (C22S/C30S/C62S/C81S PCNA) that retains Lys164 PCNA (designated as K164 PCNA).

### 2.2.3 Confirm position-specific monoubiquitination by tryptic digestion and MS/MS

We also subjected the chemically ubiquitinated PCNA (K164C PCNA-Ub) to trypsin digestion and analyzed the tryptic peptides by ESI-MS and MS/MS. From ESI-MS (Figure 2.10), the peptide of 2313.07 Da corresponds to the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys164 (DLSQLSDSINIMITCETIK) was observed and it agrees with a theoretical molecular weight of 2313.15 Da. A second peptide of 2582.25 Da corresponds to a miscleavage product, in which the ubiquitin C-terminal tetrapeptide (LRGG) is conjugated to the same PCNA peptide.

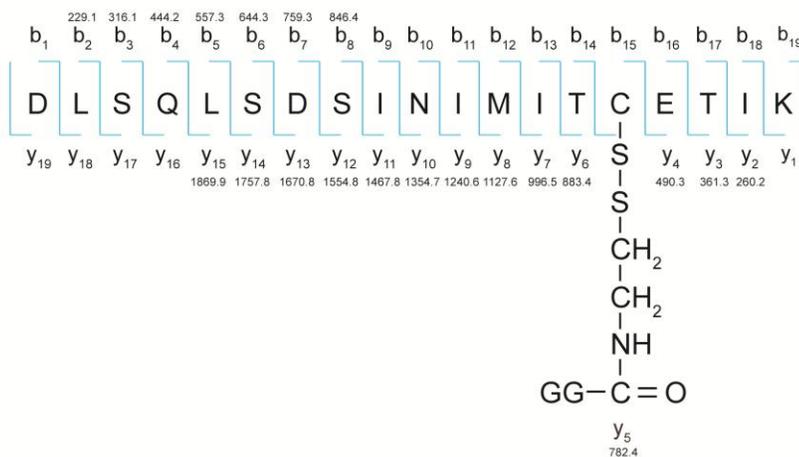


**Figure 2.10: Deconvoluted electrospray ionization mass spectra of the tryptic peptides of K164C PCNA-Ub.**

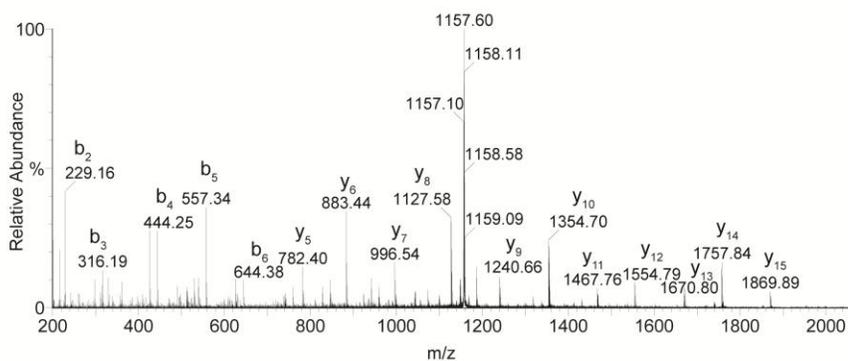
The peptide that contains the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys164 (DLSQLSDSINIMITCETIK) (Figure 2.11a) was also subjected to MS/MS. The MS/MS spectra of the ion at  $m/z$  of 1157.60 that corresponds to the precursor peptide from K164C PCNA-Ub trypsin digestion that contains the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys164 (DLSQLSDSINIMITCETIK) with a molecular weight of 2313.07 Da ( $m/z$  1157.60 for the doubly charged species). The MS/MS data (Figure 2.11b and 2.11c) confirms the peptide sequence.

a

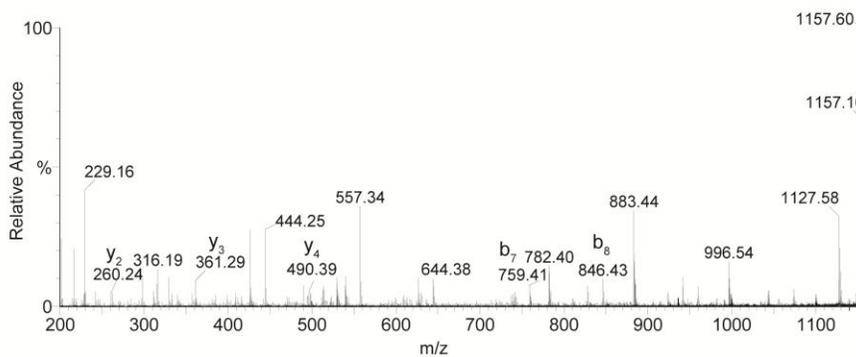
MS/MS m/z: 1157.60



b



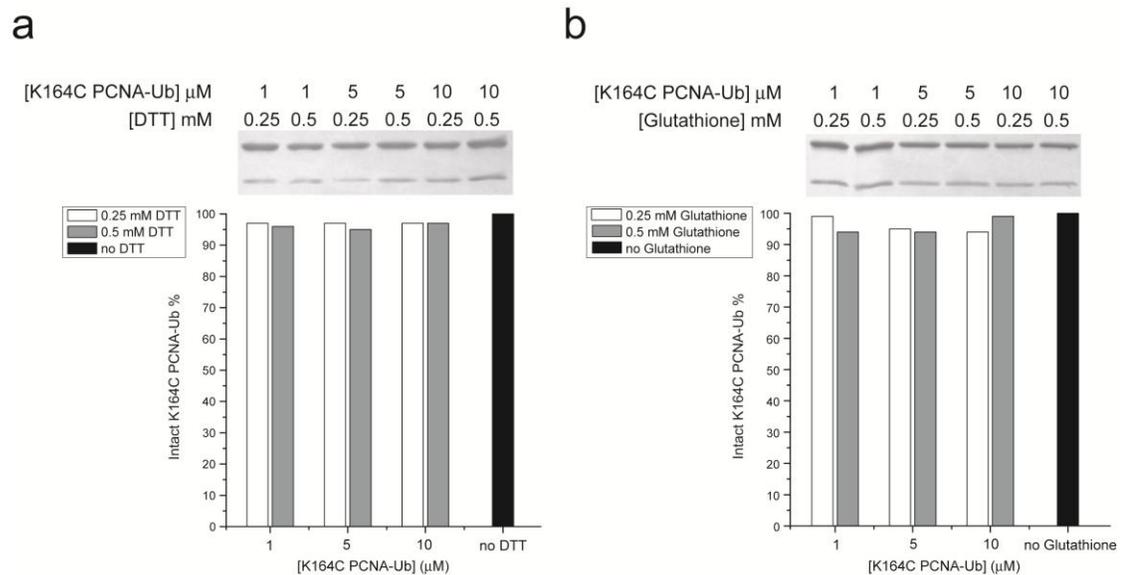
c



**Figure 2.11:** The MS/MS spectra of the ion at m/z of 1157.60 that corresponds to the precursor peptide from K164C PCNA-Ub trypsin digestion that contains the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys164 (DLSQLSDSINIMITCETIK) with a molecular weight of 2313.07 Da (m/z 1157.60 for the doubly charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The MS/MS spectrum with the measured mass and the fragments labeled. (c) The zoomed-in MS/MS spectrum between m/z 200 and 1160.

#### 2.2.4 Assess the chemical stability of the disulfide linkage

K164C PCNA-Ub was largely stable in the presence of moderate concentration of reducing agents. Over 95% of K164C PCNA-Ub stays intact even under up to 0.5 mM DTT or glutathione (Figure 2.12).

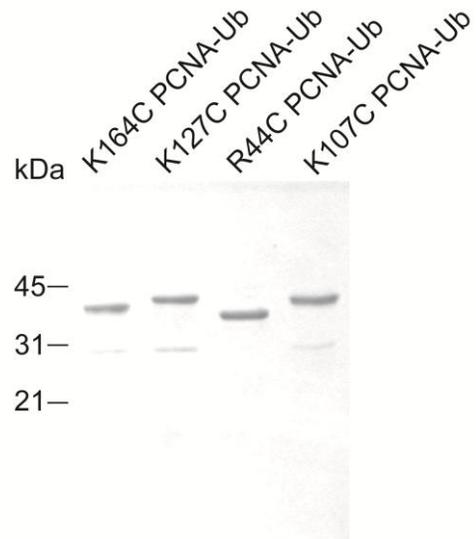


**Figure 2.12: Denaturing SDS-PAGE analysis of the chemically ubiquitinated PCNA (K164C PCNA-Ub) treated with (a) 0.25, 0.5 mM DTT (left) and (b) 0.25, 0.5 mM glutathione (right) respectively for 5 minutes at 37 °C. Top, SDS-PAGE gel of the treated K164C PCNA-Ub (comparable amount of protein was loaded); Bottom, the normalized percentage of intact K164C PCNA-Ub after the treatment. The amount of K164C PCNA-Ub loaded onto the SDS-PAGE gel was normalized for the different incubation reactions to give a comparable intensity.**

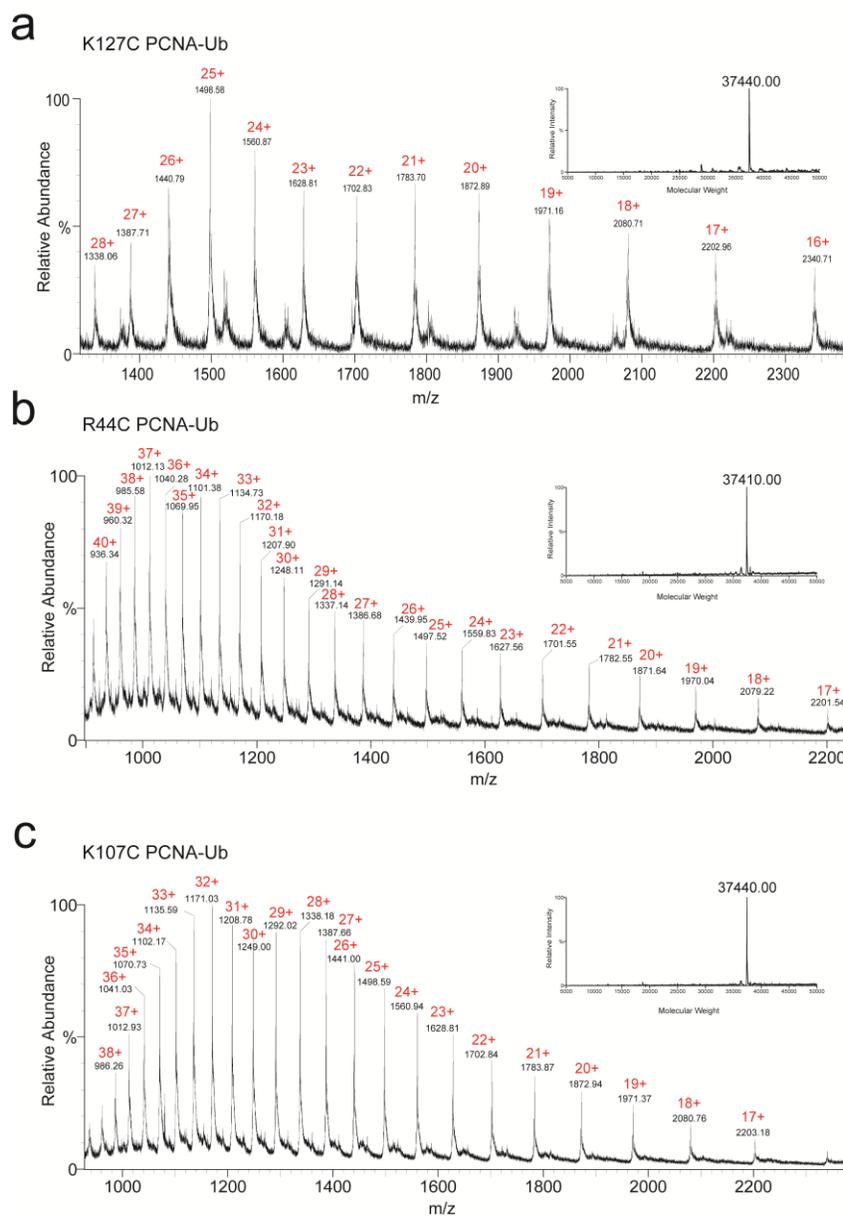
Dr. Yongxing Ai also tested the stability of K164C PCNA-Ub when incubated with yeast cell lysate at 37°C for varied time periods. The majority of the chemically ubiquitinated PCNA stayed intact after three hours of incubation. Thus, the chemically ubiquitinated PCNA can be applied not only to the assays reconstituted from highly purified proteins, but also to the cell lysate-based assays.

### **2.2.5 Chemical ubiquitination on PCNA position 127, 44 and 107**

PCNA monoubiquitinated at each of the positions was prepared following the same protocol as described in section 2.4.4 by using K127C PCNA, R44C PCNA and K107C PCNA mutants (Figure 2.13). Varied gel mobility was observed for the PCNA ubiquitinated at different positions. ESI-MS analyses of the K127C PCNA-Ub, R44C PCNA-Ub and K107C PCNA-Ub (Figure 2.14) and MS/MS analysis (Figure 2.15 – Figure 2.17) of the tryptic peptide confirmed the site-specific ubiquitination of the PCNA mutants.



**Figure 2.13: Denaturing SDS-PAGE analysis of the chemically ubiquitinated PCNA at residue 164, 127, 44 and 107 respectively.**



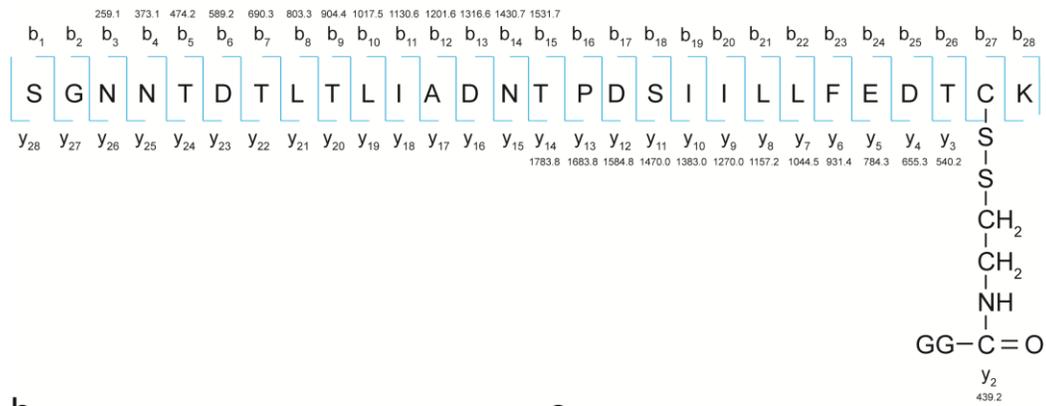
**Figure 2.14: The multiple charge states of the ubiquitinated PCNA subunit (K127C PCNA-Ub, R44C PCNA-Ub and K107C PCNA-Ub) determined by electrospray ionization mass spectrometry. The inset shows the deconvoluted peak of a molecular weight of 37,440 Da for K127C PCNA-Ub, 37,410 Da for R44C PCNA-Ub and 37,440 Da for K107C PCNA-Ub.**



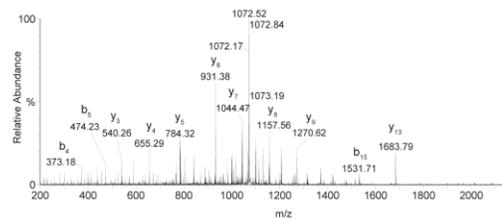


a

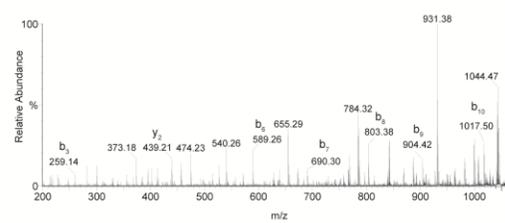
MS/MS m/z: 1072.52



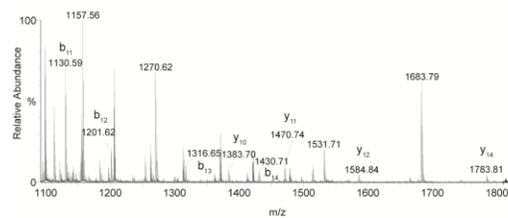
b



c



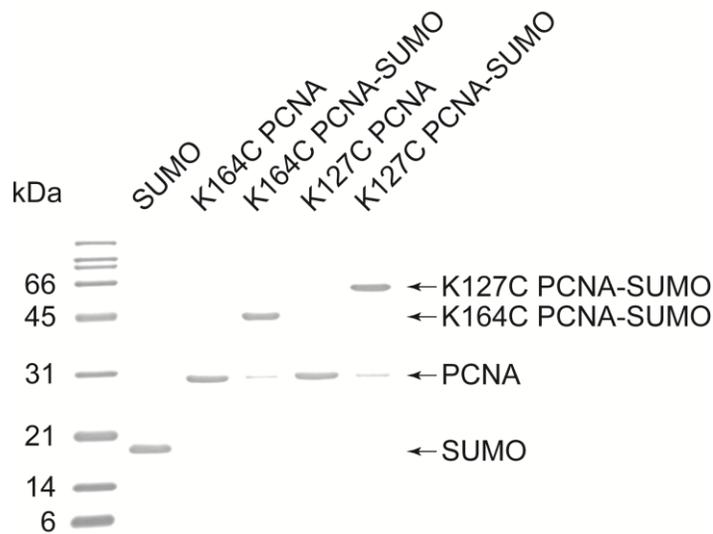
d



**Figure 2.17: The MS/MS spectra of the precursor peptide from K107C PCNA-Ub trypsin digestion that contains the ubiquitin C-terminal dipeptide (GG) conjugated to the PCNA peptide flanking Cys107 (SGNNTDTLTLIADNTPDSIILLFEDTCK) with a molecular weight of 3213.57 Da (m/z 1072.52 for the triply charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The full MS/MS spectrum of the ion at m/z 1,072.52. (c) & (d) The zoomed-in regions of the MS/MS spectra with the measured mass and the fragments labeled.**

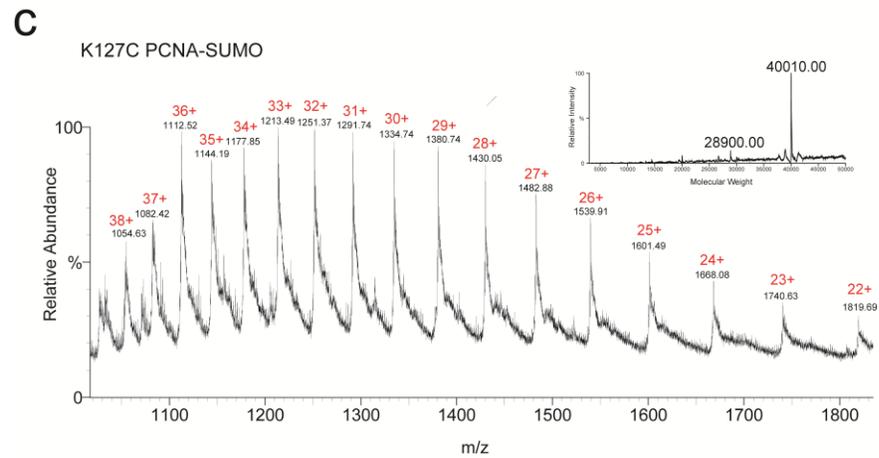
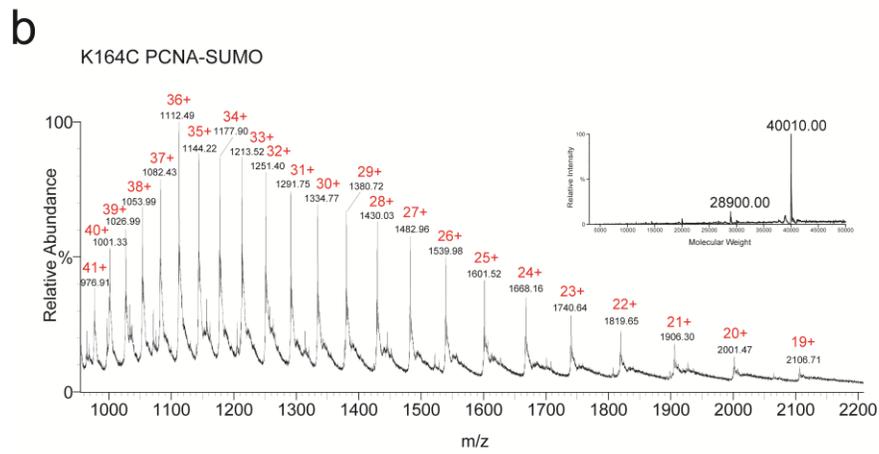
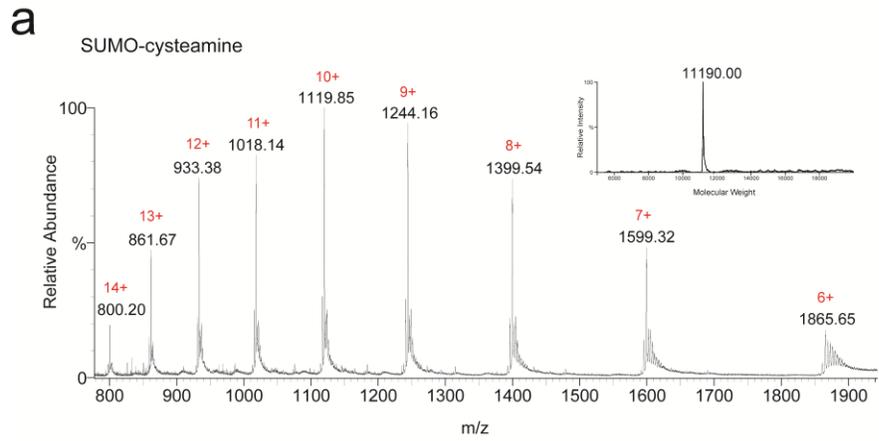
### 2.2.6 Chemical SUMOylation on PCNA position 164 and 127

Besides ubiquitination at Lys164, PCNA also undergoes SUMOylation at both Lys164 and Lys127. SUMO shares low sequence similarity to ubiquitin despite a similar fold. Following a similar protocol, we were able to chemically SUMOylate PCNA at residue 127 and 164 respectively (Figure 2.18). The percentages of SUMOylation were 95% and 90% for K164C PCNA-SUMO and K127C PCNA-SUMO respectively. Different gel electrophoretic behavior was observed for the two forms of SUMOylated PCNA. A similar observation was also made in the denaturing SDS-PAGE analysis of the enzymatic SUMOylation product of PCNA at K164 and K127.<sup>48</sup> Because the SUMOylated PCNA exists as a branched polymer under denaturing condition, it is likely that the actual branching point (determined by the modification site) affects the hydrodynamic property of the SUMOylated PCNA.



**Figure 2.18: Denaturing SDS-PAGE gel analysis of the chemically SUMOylated PCNA at residue 164 (K164C PCNA-SUMO) and 127 (K127C PCNA-SUMO).**

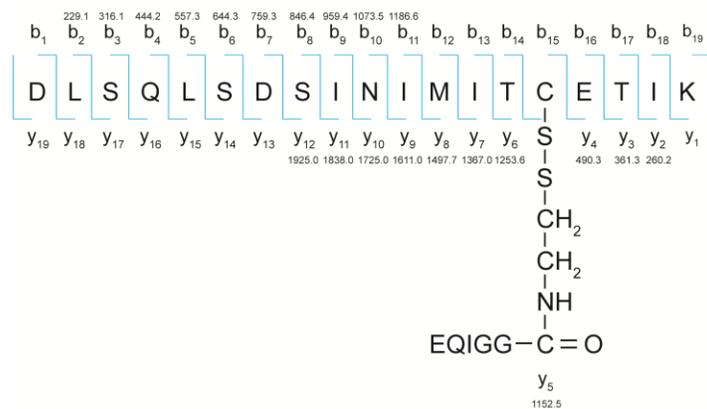
The SUMOylated PCNAs were characterized by ESI mass spectrometry (Figure 2.19). For K164C PCNA-SUMO and K127C PCNA-SUMO, the measured molecular weight of 40,010.00 Da agrees well with the theoretical molecular weight (40,014.3 Da) (Figure 2.19b&c). Further MS/MS analysis of the tryptic peptide confirmed the site-specific SUMOylation of both PCNA mutants (Figure 2.20 and 2.21).



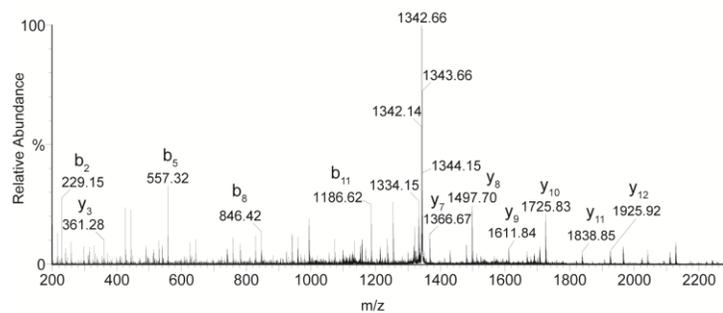
**Figure 2.19: The multiple charge states of (a) SUMO-cysteamine (SUMO-SH), (b) K164C PCNA-SUMO, (c) K127C PCNA-SUMO determined by electrospray ionization mass spectrometry. The inset in (a) shows the deconvoluted peak with a molecular weight of 11,190.0 Da for SUMO-SH, which indicates the N-terminal methionine of the recombinant SUMO is removed by methionine aminopeptidase. The theoretical molecular weight of SUMO-SH (amino acid 2-98, without N-ter Met) is 11,189.5 Da. The measured molecular weight of K164C PCNA-SUMO (40,010 Da) and K127C PCNA-SUMO (40,010 Da) agrees with the calculated molecular weight (40,014.3 Da) for PCNA-SUMO (without the N-ter Met in SUMO).**

a

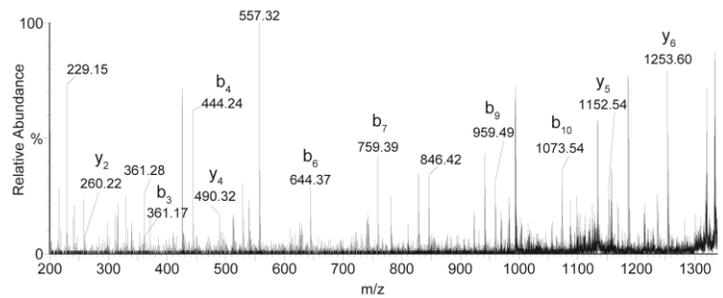
MS/MS m/z: 1342.66



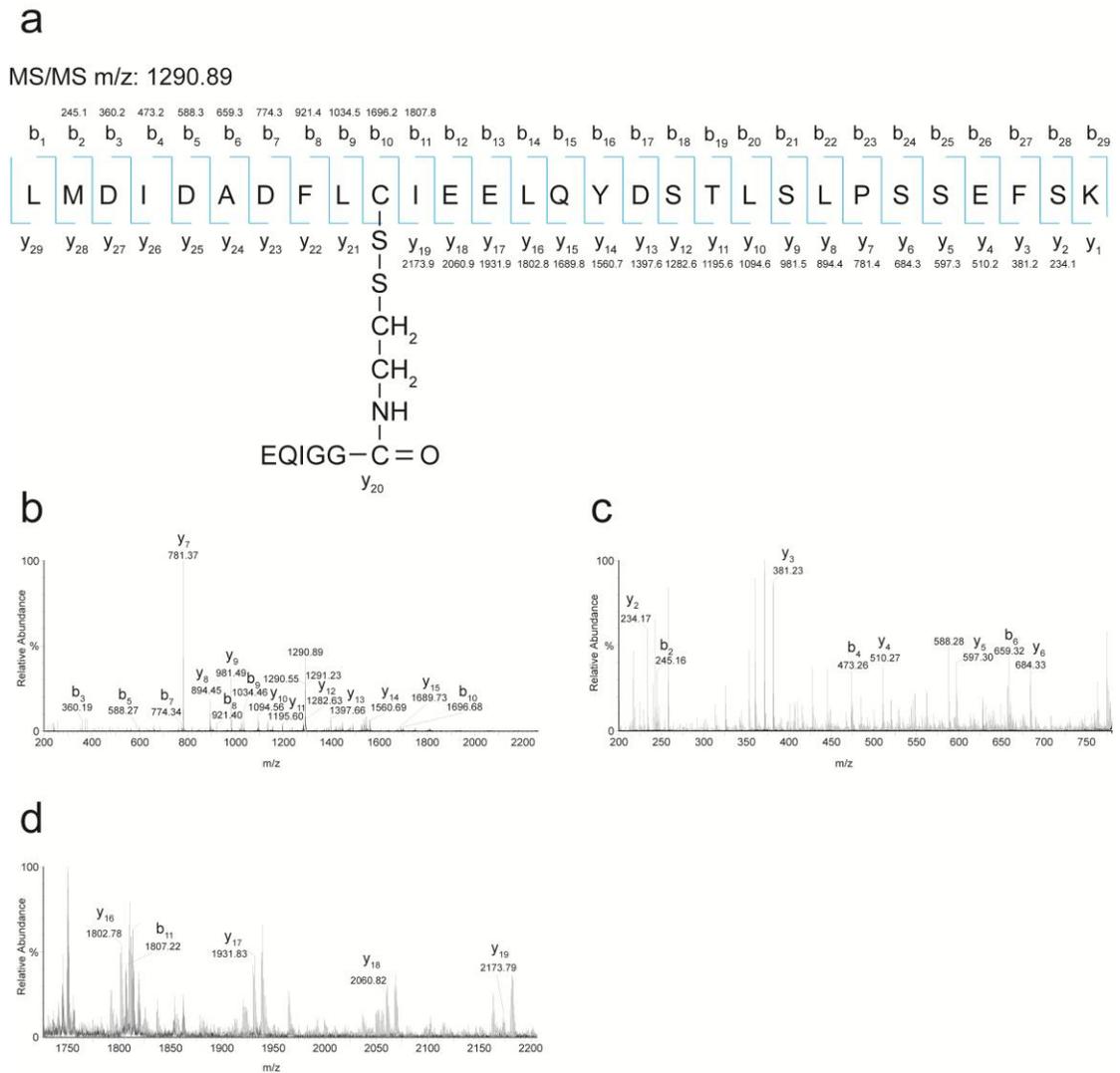
b



c



**Figure 2.20: The MS/MS spectra of the precursor peptide from K164C PCNA-SUMO trypsin digestion that contains the SUMO C-terminal pentapeptide (EQIGG) conjugated to the PCNA peptide flanking Cys164 (DLSQLSDSINIMITCETIK) with a molecular weight of 2683.65 Da ( $m/z$  1342.66 for the doubly charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The full MS/MS spectrum with the measured mass and the fragments labeled. (c) The zoomed-in MS/MS spectrum between  $m/z$  200 and 1340.**



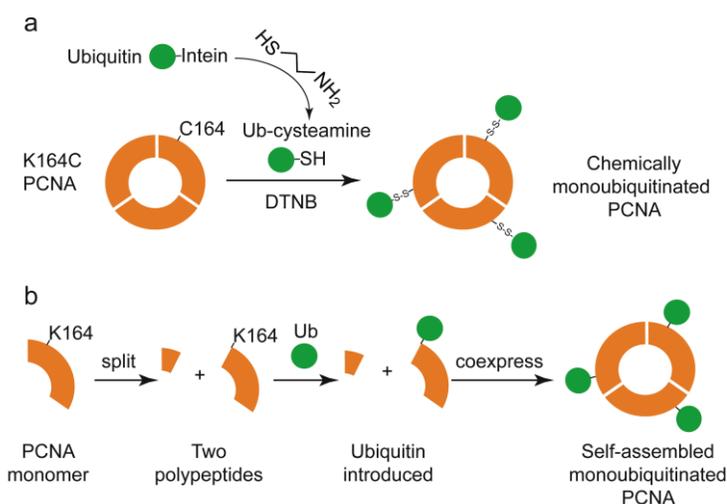
**Figure 2.21: The MS/MS spectra of the precursor peptide from K127C PCNA-SUMO trypsin digestion that contains the ubiquitin C-terminal pentapeptide (EQIGG) conjugated to the PCNA peptide flanking Cys127 (LMDIDADFLCIEELQYDSTLSLPSEFSK) with a molecular weight of 3869.14 Da (m/z 1290.89 for the triply charged species). (a) The PCNA peptide sequence and the theoretical molecular weight of the different fragments. (b) The full MS/MS spectrum with the measured mass and the fragments labeled. (c) & (d) The zoomed-in MS/MS spectra between m/z 200-780 and 1725-2205 respectively.**

### 2.3 Discussion

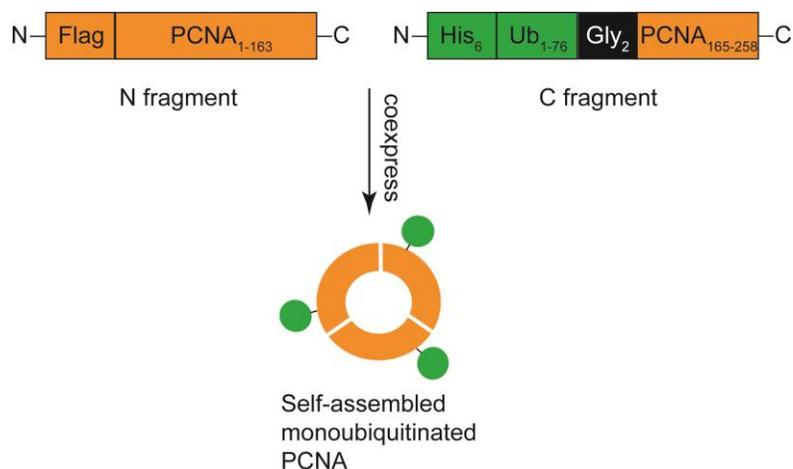
We have developed a non-enzymatic approach to conjugate a single ubiquitin moiety to PCNA at residue 164 through disulfide exchange and intein chemistry (Figure 2.22a). A yield of ca. 4.5 milligram of ubiquitinated PCNA in a typical preparation represents substantial improvement (>100 fold) over the enzymatic approach, which typically yield 5~10 micrograms of monoubiquitinated PCNA per preparation (Haracska, L. personal communication). Following a similar protocol we also achieved efficient site-specific ubiquitination of PCNA at position 127, 44, 107 and SUMOylation of PCNA at position 164 and 127 respectively. Therefore the chemical approach can be readily generalized for ubiquitination, SUMOylation, and potentially modification by other Ubls.

Washington group developed a split PCNA method (Figure 2.22b) to generate large amounts of yeast monoubiquitinated PCNA.<sup>49</sup> First, PCNA was split into two polypeptides at the site of ubiquitination. Then ubiquitin was inserted N-terminal to the C-terminal half of PCNA. The two polypeptides were co-expressed *in vivo* and self assembled into a functional subunit. Using this ubiquitinated PCNA, the first high resolution X-ray crystal structure of ubiquitinated PCNA was solved. The

crystal structure of this monoubiquitinated PCNA reveals that ubiquitin resides on the back side of the PCNA, where ubiquitin can also interact with Pol $\eta$ , a non-classical DNA polymerase. The ubiquitin modification does not change the conformation of PCNA. In the Washington group's splitting method (Figure 2.23), the linkage generated between ubiquitin and PCNA deviates from the native isopeptide linkage. The C fragment contained the entire ubiquitin sequence (residues 1–76) was fused via a short linker to residues 165 to 258 of PCNA. In comparison our chemical ubiquitination method generates the linkage close to the native form. No splitting and *in vivo* expression steps were required in our ligation method. Also, our method is more flexible in manipulating ubiquitination or sumoylation on different sites of PCNA.



**Figure 2.22: New methods for monoubiquitination of PCNA. (a) Chemical ubiquitination. (b) Splitting PCNA method by Washington group.**

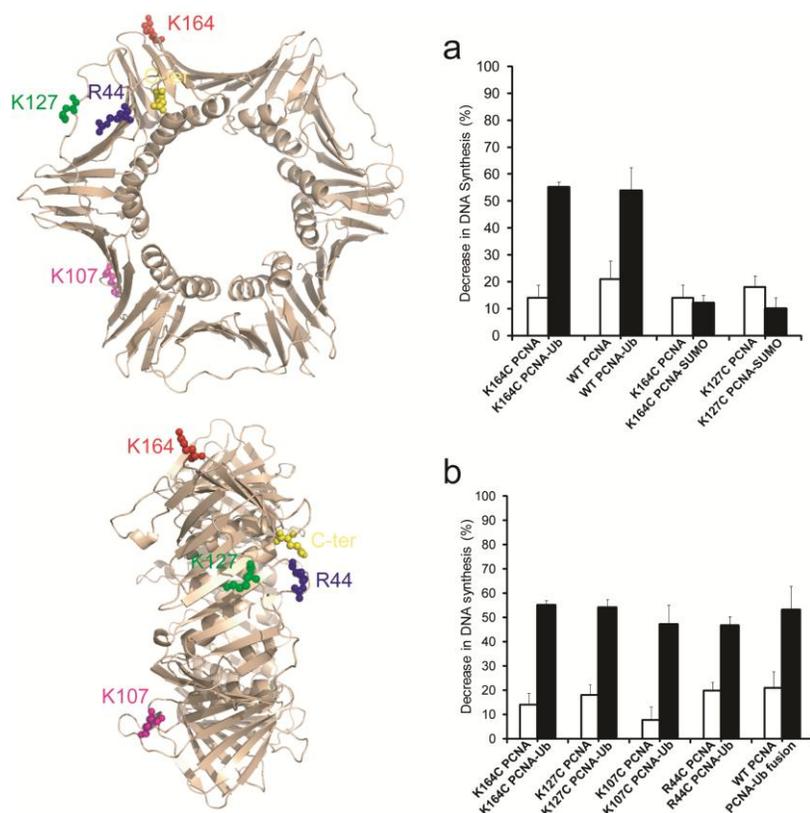


**Figure 2.23: Diagram of the two polypeptides used to generate monoubiquitinated PCNA by Washington group.<sup>49</sup> The N fragment contained amino acid residues 1 to 163 of PCNA and was N-terminally FLAG-tagged. The C fragment contained the entire ubiquitin sequence (residues 1–76) fused via a short linker to residues 165 to 258 of PCNA and was N-terminally His<sub>6</sub>-tagged. The short linker consisted of two glycine residues because this is nearly isosteric with the side chain of Lys 164 and the isopeptide bond to the C-terminus of ubiquitin.**

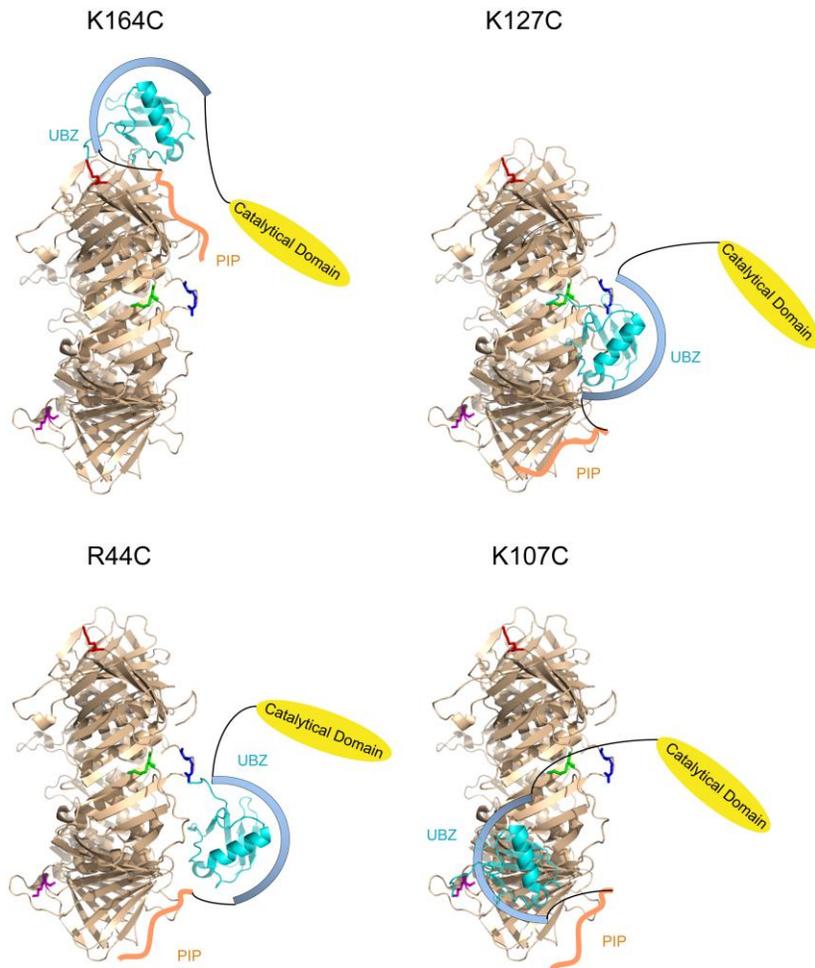
Next, the K164C PCNA-Ub was tested in the polymerase exchange assay by Dr. Yongxing Ai. Dr. Ai found the chemically monoubiquitinated PCNA (K164C PCNA-Ub) is functionally equivalent to the native monoubiquitinated PCNA as demonstrated by the polymerase exchange assay. Enzymatic ubiquitination of PCNA by Rad6/Rad18 occurs specifically at Lys164 on PCNA. Several studies have utilized the PCNA-ubiquitin chimeras, in which ubiquitin was fused to either N- or C-terminus of PCNA, in substitution of the native Lys164 monoubiquitinated PCNA.<sup>11, 13</sup> Interestingly, despite the differences in the nature of linkage and the position of modification, the PCNA-Ub fusion can, to a large extent, substitute for the loss of the

native monoubiquitinated PCNA when expressed in cells under genotoxic stress.<sup>13</sup> These unusual observations prompted us to investigate whether the normal function of ubiquitination is dictated by the position of modification. Therefore, we chemically ubiquitinated PCNA at several different positions, including Lys164, Lys127, Lys107, and Arg44. Lys127 undergoes SUMOylation, but not ubiquitination.<sup>22, 26</sup> Arg44 is close to the C-terminus of PCNA, which was used to fuse ubiquitin in the PCNA-Ub chimera. Lys107 is located on the back side of PCNA, but in a loop important in TLS.<sup>50</sup> A comparison of the chemically ubiquitinated PCNAs that differ only in the position of the modification allowed us to explicitly probe the effect of the site of ubiquitination. We found that monoubiquitination of PCNA at the various positions in PCNA did not significantly affect the efficiency of polymerase exchange (Figure 2.24). These observations suggest that the ubiquitin moiety is likely conformationally flexible on PCNA. Given the long linkage formed by the C-terminal peptide of ubiquitin and the side chain of Lys in PCNA, the covalently conjugated ubiquitin is likely capable of adopting different conformations that can lead to a functional ternary complex consisting of PCNA, ubiquitin and Pol $\eta$  (Figure 2.25). Modification by ubiquitin or UBL is likely unique among the various forms of post-translational modifications in that it can mediate a long-range protein-protein interaction and facilitate the formation of a dynamic protein complex. Two SUMOylated PCNAs (K164C PCNA-SUMO and K127C PCNA-SUMO) were tested in the polymerase exchange assay. Both of them had little effect on the DNA synthesis by the Pol $\delta$ -PCNA holoenzyme in the presence of Pol $\eta$  (Figure 2.24). This result corroborates the notion that the inhibitory effect of Pol $\eta$  on the Pol $\delta$  DNA synthesis is specific and strictly dependent on monoubiquitination of PCNA.

The same method was also applied to H2B ubiquitination by Muir group.<sup>51</sup> The disulfide-linked ubiquitinated H2B (uH2B) on position 120 stimulated the H3K79 methyltransferase activity of hDot1L (responsible for methylating H3 residues Lys4 and Lys79) as effective as the native uH2B. Positions 108, 116 and 125 on H2B and position 22 on H2A were selected for ubiquitination in addition to position 120 on H2B. Ubiquitin-like proteins Smt3, Nedd8 and Hub1, which are structurally close to ubiquitin, were conjugated to H2B K120C by following the same method. The stimulation assay results depicted a possible ubiquitin surface required for hDot1L activity stimulation and indicated that the stimulatory effect is not strictly dependent on ubiquitin conjugation sites. Also, it was the first time to report that the function of H2BK125 ubiquitination might be stimulating hDot1L activity by the “site specific” advantage of using the disulfide linkage method.



**Figure 2.24:** (Left) The positions of chemical ubiquitination mapped to the homotrimeric PCNA, shown as front and side views. Lys164 (red), Lys127 (green), Lys107 (magenta), Arg44 (blue) and the C-terminal residue Glu258 (yellow) are shown in one PCNA subunit. (Right) (a) The extent of polymerase exchange reported by the decrease in DNA synthesis for ubiquitinated and SUMOylated PCNA. The different forms of modified PCNA include chemically ubiquitinated PCNA at residue 164 (K164C PCNA-Ub), enzymatically ubiquitinated PCNA (WT PCNA-Ub), chemically SUMOylated PCNA at residue 164 (K164C PCNA-SUMO) and chemically SUMOylated PCNA at residue 127 (K127C PCNA-SUMO). (b) Comparing the function of PCNA ubiquitinated at different positions in the polymerase exchange assay. The positions of chemical ubiquitination include residues 164, 127, 107 and 44. PCNA-Ub fusion refers to a chimeric protein with ubiquitin fused to the C terminus of PCNA. The ubiquitinated PCNAs and the unmodified wild-type and mutant PCNAs are compared. Note: Bar graph a&b were generated by Dr. Yongxing Ai.



**Figure 2.25: Cartoon illustration of possible interactions among PCNA, ubiquitin and Pol $\eta$  when ubiquitin is conjugated to various positions (164, 127, 44 and 107) on PCNA. Pol $\eta$  is composed of catalytical domain, UBZ (ubiquitin binding) domain and PIP (PCNA interacting peptide). Pol $\eta$  interacts with the front side (the right side in the cartoon) of PCNA. The long linkage formed by the C-terminal peptide of ubiquitin and the lysine side chain of PCNA allows ubiquitin to adopt various conformations that can lead to a functional ternary complex with Pol $\eta$ .**

## 2.4 Materials and Methods

### 2.4.1 Preparation of ubiquitin-cysteamine (Ub-SH)

Ubiquitin gene containing residues 1-76 was amplified from *S. cerevisiae* genomic DNA by PCR. After digestion by *NdeI* and *SapI* restriction enzymes, ubiquitin gene was ligated into a similarly digested pTYB1 vector (New England Biolabs). The resulting plasmid was confirmed by DNA sequencing. For protein expression, the plasmid was transformed into BL21(DE3) cells. Cells were cultivated at 37 °C in LB medium (1L) containing 100 µg/ml ampicillin. The cell culture was induced with 0.5 mM IPTG at OD<sub>600</sub> 0.6 ~ 0.8, and grown for an additional 18 hrs at 15 °C. Cells were harvested by centrifugation at 8,000 g for 30 mins and resuspended in lysis buffer (20 mM Tris, 200 mM NaCl, 1 mM EDTA, 5% glycerol, pH 7.5). Cells were lysed by sonication and the resulting lysate was cleared by centrifugation at 38,000 g for 30 mins. The supernatant was incubated with 4 ml chitin resin (New England Biolabs) at 4 °C for 2 hr. The resin was then washed with 250 ml of high salt wash buffer (20 mM Tris, 1 M NaCl, 1 mM EDTA, 5% glycerol, pH 7.5) and 250 ml of low salt wash buffer (50 mM Tris, 100 mM NaCl, pH 8.5). The resin was then incubated with 10 ml cleavage buffer (50 mM Tris, 100 mM NaCl, 100 mM cysteamine, pH 8.5) for 12 hrs at room temperature. The column was eluted with 10 ml cleavage buffer. To prepare the N-terminally His-tagged ubiquitin, the ubiquitin gene with an engineered N-terminal 6xHis-tag was cloned into pTYB1 vector using *NdeI* and *SapI* restriction sites. The 6xHis Ub was expressed and purified as described for ubiquitin.

#### **2.4.2 Preparation of SUMO-cysteamine**

The gene encoding SUMO (SMT3) containing residues 1-98 was amplified from the genomic DNA of *S. cerevisiae*. After digestion by *NdeI* and *SapI* restriction enzymes, the gene was ligated into a similarly digested pTYB1 vector (New England Biolabs). The resulting plasmid was confirmed by DNA sequencing. For protein expression, the plasmid was transformed into BL21(DE3) cells. Cells were cultivated at 37 °C in LB medium (1L) containing 100 µg/ml ampicillin. The cell culture was induced with 0.5 mM IPTG at OD<sub>600</sub> 0.6 ~ 0.8, and grown for an additional 18 hrs at 15 °C. Cells were harvested by centrifugation at 8,000 g for 30 mins and resuspended in lysis buffer (20 mM Tris, 200 mM NaCl, 1 mM EDTA, 5% glycerol, pH 7.5). Cells were lysed by sonication and the resulting lysate was cleared by centrifugation at 38,000 g for 30 mins. The supernatant was incubated with 4 ml chitin resin (New England Biolabs) at 4 °C for 2 hrs. The resin was then washed with 250 ml of high salt wash buffer (20 mM Tris, 1 M NaCl, 1 mM EDTA, 5% glycerol, pH 7.5) and 250 ml of low salt wash buffer (50 mM Tris, 100 mM NaCl, pH 8.5). The resin was then incubated with 10 ml cleavage buffer (50 mM Tris, 100 mM NaCl, 100 mM cysteamine, pH 8.5) for 12 hrs at room temperature. The column was eluted with 10 ml cleavage buffer.

#### **2.4.3 Preparation of PCNA mutants**

K164C/C22S/C30S/C62S/C81S PCNA (K164C PCNA) was prepared using Quikchange mutagenesis kit (Stratagene) by using the yeast *S. cerevisiae* gene PCNA in pET-22b plasmid as template. For expression, the plasmid was transformed into *E. coli* BL21(DE3) and cells were cultivated at 37 °C in LB medium (1L) containing 100 µg/ml ampicillin. The expression was induced with 0.5 mM IPTG at OD<sub>600</sub> of 0.6 ~ 0.8 and incubated for an additional 6 hrs at 37 °C. Cells were harvested

by centrifugation at 8,000 g for 30 mins and resuspended in lysis buffer (50 mM Tris, 1 mM EDTA, 5% glycerol, pH 7.5). Cells were lysed by sonication and the resulting lysate was cleared by centrifugation at 38,000 g for 30 mins. Streptomycin sulfate was added to the supernatant to 0.5% (w/v) to precipitate nucleic acids. After centrifugation, ammonium sulfate to 50% (w/v) was added to precipitate PCNA. The protein pellet was then resuspended in 20 ml buffer A (20 mM Tris, 20 mM NaCl, 1 mM EDTA, 5% glycerol, pH 7.5) and dialyzed against 2 × 1L buffer A with 1 mM DTT overnight. The dialyzed protein was loaded onto a DEAE cellulose (Sigma-Aldrich) column and eluted using a NaCl gradient up to 700 mM NaCl in buffer A. The purified fractions were combined and concentrated to 20 ml, followed by dialysis against 2 × 1L buffer A with 1 mM DTT overnight. PCNA was further purified using a HiTrap Q FF column using a salt gradient (20 mM ~700 mM NaCl) in buffer A. Pure PCNA fractions were combined and concentrated. Following a buffer exchange step the final PCNA was stored in buffer A with 100 mM NaCl. Other PCNA mutants (K127C, R44C and K107C) described in this work were purified following the same protocol.

#### **2.4.4 Chemical ubiquitination of PCNA**

In a typical reaction, Ub-SH (5.2 mg, 0.6 μmol) in 500 μl cleavage buffer (50 mM Tris, 100 mM NaCl, 100 mM cysteamine, pH 8.5) was passed through a PD-10 column (GE) equilibrated in ligation buffer (20 mM Tris, 50 mM NaCl, 1 mM EDTA, pH 7.0). In some reactions, a buffer exchange step was included to facilitate the removal of cysteamine. The fractions containing Ub-SH were combined and concentrated to 400 μl. To this solution, 5, 5'-dithiobis-(2-nitrobenzoic acid) (DTNB) (4.8 mg, 12 μmol) in 50 mM sodium phosphate (pH 7.5) was added. The mixture was

incubated at room temperature for 30 mins, then passed through a PD-10 column to remove excess DTNB. The ubiquitin protein fractions were combined (2.5 ml) and concentrated to 250  $\mu$ l.

A solution of K164C PCNA (454  $\mu$ l, 0.15  $\mu$ mol monomer) was passed through a PD-10 column to remove residual DTT. The solution was mixed with the DTNB treated Ub-SH solution (250  $\mu$ l, 0.6  $\mu$ mol) and incubated at room temperature. Aliquots were removed at various time points and quenched by flash freezing with liquid N<sub>2</sub>. The reaction products at various time points were analyzed in a 15% denaturing SDS-PAGE gel. To remove the excess amount of ubiquitin, the reaction product was loaded onto a HiTrap Q FF column and eluted with a NaCl gradient of 20 mM ~700 mM NaCl in buffer B (20 mM Tris, 700 mM NaCl, 1 mM EDTA, 5% glycerol, pH 7.5). While ubiquitin alone did not bind to the column, the ubiquitinated PCNA was eluted with 500 mM NaCl. The same protocol was followed to prepare other ubiquitinated PCNA mutants (K127C, K107C and Arg44C). The percentage of conversion is 82-89%, 90-99%, and 82-90% for K127C, R44C and K107C respectively. An identical protocol was followed to chemically ubiquitinate PCNA with N-ter His-tagged ubiquitin except that an affinity purification step with the Co<sup>2+</sup> column was included.

#### **2.4.5 Determine the molecular weight of ubiquitinated PCNA by ESI-MS**

Samples were separated on a C5 reversed-phase column (1 mm i.d., 5 cm length; Supelco). Mobile phase A (99.9% H<sub>2</sub>O and 0.1% TFA) and mobile phase B (100% acetonitrile) were delivered at a flow rate of 50  $\mu$ L/min. A two-step wash of 6 min (100% phase A) followed by 7 min (100% phase B) was used. Proteins were analyzed with a Q-TOF Ultima API-US quadrupole time-of-flight mass spectrometer

(Micromass, Manchester, U.K.). The sample eluted from the RPLC column was sent into a nano flow API-ESI source. Mass spectra were acquired from 400 to 1800  $m/z$  with a scan time of 0.5 s and an interscan time of 0.01 s. Data processing was performed with MaxEnt1, which converted the raw spectrum of multiple charged ions into a deconvoluted display of singly charged ions. MaxEnt1 was performed over the range 5000 – 50,000  $m/z$  for all protein data.

#### **2.4.6 Tryptic digestion of the chemically ubiquitinated PCNA**

A solutions of 50  $\mu$ l (150  $\mu$ g) K164C PCNA-Ub was incubated for 5 min at 100 °C to denature the protein. The protein precipitate was pelleted by centrifugation and suspended in a solution (150  $\mu$ L) containing 2  $\mu$ g trypsin (0.013 mg/mL) in 50 mM ammonium bicarbonate buffer, pH 8.0. The digestion mixture was incubated at 37 °C for 8 hours. The tryptic peptides were concentrated by loading the solution onto a ZipTip SCX pipetter tip (Millipore, Billerica, MA). The resulting sample was analyzed using a Q-TOF Ultima API-US quadrupole time-of-flight mass spectrometer (Micromass, Manchester, U.K.). The sample mixture was injected into a nano flow API-ESI source. Mass spectra were acquired from 400 to 1800  $m/z$  with a scan time of 0.5 s and an interscan time of 0.01 s. Data processing was performed with MaxEnt3, which converted the raw spectrum of multiple charged ions into a deconvoluted display of singly charged ions. MaxEnt3 was performed over the range 1000 – 4000  $m/z$  for all protein data.

#### **2.4.7 Complete cleavage of the disulfide linkage with high concentration of DTT**

To a reaction mixture containing 16  $\mu$ M K164C PCNA-Ub in 20 mM Tris buffer (pH 7.0) with 50 mM NaCl, 1 mM EDTA, 5 or 100 mM DTT was added and

incubated at 25 °C. Under 5 mM DTT condition, the reaction products at various time points were taken out and flash frozen by liquid nitrogen. The sample incubated with 100 mM DTT was analyzed right after 5-min incubation. All samples were analyzed by a 15% SDS-PAGE gel.

#### **2.4.8 Sensitivity of chemically ubiquitinated PCNA to reducing reagents**

To a reaction mixture containing 1 µM to 10 µM K164C PCNA-Ub in 20 mM Tris buffer (pH 7.0) with 50 mM NaCl, 1 mM EDTA, 0.25 mM or 0.5 mM DTT or glutathione was added and incubated at 37 °C for 5 mins. The reaction was stopped by passing through a G-25 spin column. The sample was analyzed by a 15% SDS-PAGE gel. Over 94% of K164C PCNA-Ub stayed intact following the treatment.

#### **2.4.9 Chemical SUMOylation of PCNA**

In a typical reaction, SUMO-SH (2.3 mg, 0.2 µmol) in 400 µl cleavage buffer (50 mM Tris, 100 mM NaCl, 100 mM cysteamine, pH 8.5) was passed through a PD-10 column (GE) equilibrated in ligation buffer (20 mM Tris, 50 mM NaCl, 1 mM EDTA, pH 7.0). A buffer exchange step was included to remove the residual cysteamine. The fractions containing SUMO-SH were combined and concentrated to 400 µl. To this solution, 5, 5'-dithiobis-(2-nitrobenzoic acid) (DTNB) (1.6 mg, 4 µmol) in 50 mM sodium phosphate (pH 7.5) with 1 mM EDTA was added. The mixture was incubated at room temperature for 30 min, then passed through a PD-10 column to remove excess DTNB. The SUMO protein fractions were combined (3 ml) and concentrated to 400 µl. A solution of K164C PCNA (100 µl, 0.05 µmol monomer) was passed through a PD-10 column to remove residual DTT. The solution was mixed with the DTNB-treated SUMO-SH solution (400 µl, 0.2 µmol) and incubated at room temperature for overnight. To remove the excess amount of SUMO, reaction product

was loaded onto a HiTrap Q FF column and eluted with a NaCl gradient of 20 ~ 700 mM NaCl in buffer B (20 mM Tris, 700 mM NaCl, 1 mM EDTA, 5% glycerol, pH 7.5). SUMO alone did not bind to the column; the SUMOylated PCNA was eluted with 500 mM NaCl. The PCNA-SUMO fractions were pooled together, followed by buffer exchange (20 mM Tris, 20 mM NaCl, 1 mM EDTA, 5% glycerol, pH 7.5) and concentrated to 400  $\mu$ l. The percentage of conversion of the PCNA-SUMO was approximately 95%. The SUMOylation of K127C PCNA was carried out following the same protocol as described for K164C PCNA. A final conversion of 90% was achieved.

#### **2.4.10 Determine the content of free thiol in unmodified and ubiquitinated PCNA**

K164C PCNA, K164C PCNA-Ub or K164C PCNA-SUMO at 38  $\mu$ M were incubated with 1.2 mM 5, 5'-dithiobis-(2-nitrobenzoic acid) (DTNB) at 25  $^{\circ}$ C for 5 mins in a buffer containing 20 mM Tris (pH 7.5), 20 mM NaCl, and 1 mM EDTA. The absorbance was recorded at 412 nm ( $\epsilon = 14,150 \text{ M}^{-1} \text{ cm}^{-1}$ ) and corrected for background. For K164C PCNA, 1.18 DTNB-reactive thiol per PCNA monomer was determined by the assay. In contrast, 0.01 DTNB-reactive thiol per K164C PCNA-Ub molecule and 0.1 DTNB-reactive thiol per K164C PCNA-SUMO molecule was determined.

## REFERENCES

1. Hershko, A.; Ciechanover, A., The ubiquitin system. *Annu Rev Biochem* **1998**, *67*, 425-79.
2. Kirkin, V.; Dikic, I., Role of ubiquitin- and Ubl-binding proteins in cell signaling. *Curr Opin Cell Biol* **2007**, *19* (2), 199-205.
3. Kerscher, O.; Felberbaum, R.; Hochstrasser, M., Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **2006**, *22*, 159-80.
4. Welchman, R. L.; Gordon, C.; Mayer, R. J., Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat Rev Mol Cell Biol* **2005**, *6* (8), 599-609.
5. Peng, J.; Schwartz, D.; Elias, J. E.; Thoreen, C. C.; Cheng, D.; Marsischky, G.; Roelofs, J.; Finley, D.; Gygi, S. P., A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* **2003**, *21* (8), 921-6.
6. Pickart, C. M., Ubiquitin in chains. *Trends Biochem Sci* **2000**, *25* (11), 544-8.
7. Nijman, S. M.; Luna-Vargas, M. P.; Velds, A.; Brummelkamp, T. R.; Dirac, A. M.; Sixma, T. K.; Bernards, R., A genomic and functional inventory of deubiquitinating enzymes. *Cell* **2005**, *123* (5), 773-86.
8. Zhuang, Z.; Ai, Y., Processivity factor of DNA polymerase and its expanding role in normal and translesion DNA synthesis. *Biochim Biophys Acta* **2010**, *1804* (5), 1081-93.
9. Kannouche, P. L.; Wing, J.; Lehmann, A. R., Interaction of human DNA polymerase eta with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage. *Mol Cell* **2004**, *14* (4), 491-500.
10. Watanabe, K.; Tateishi, S.; Kawasuji, M.; Tsurimoto, T.; Inoue, H.; Yamaizumi, M., Rad18 guides poleta to replication stalling sites through physical interaction and PCNA monoubiquitination. *EMBO J* **2004**, *23* (19), 3886-96.

11. Bienko, M.; Green, C. M.; Crosetto, N.; Rudolf, F.; Zapart, G.; Coull, B.; Kannouche, P.; Wider, G.; Peter, M.; Lehmann, A. R.; Hofmann, K.; Dikic, I., Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science* **2005**, *310* (5755), 1821-4.
12. Wood, A.; Garg, P.; Burgers, P. M., A ubiquitin-binding motif in the translesion DNA polymerase Rev1 mediates its essential functional interaction with ubiquitinated proliferating cell nuclear antigen in response to DNA damage. *J Biol Chem* **2007**, *282* (28), 20256-63.
13. Parker, J. L.; Bielen, A. B.; Dikic, I.; Ulrich, H. D., Contributions of ubiquitin- and PCNA-binding domains to the activity of Polymerase eta in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **2007**, *35* (3), 881-9.
14. Broomfield, S.; Chow, B. L.; Xiao, W., MMS2, encoding a ubiquitin-conjugating-enzyme-like protein, is a member of the yeast error-free postreplication repair pathway. *Proc Natl Acad Sci U S A* **1998**, *95* (10), 5678-83.
15. Brusky, J.; Zhu, Y.; Xiao, W., UBC13, a DNA-damage-inducible gene, is a member of the error-free postreplication repair pathway in *Saccharomyces cerevisiae*. *Curr Genet* **2000**, *37* (3), 168-74.
16. Branzei, D.; Seki, M.; Enomoto, T., Rad18/Rad5/Mms2-mediated polyubiquitination of PCNA is implicated in replication completion during replication stress. *Genes Cells* **2004**, *9* (11), 1031-42.
17. Unk, I.; Hajdu, I.; Fatyol, K.; Szakal, B.; Blastyak, A.; Bermudez, V.; Hurwitz, J.; Prakash, L.; Prakash, S.; Haracska, L., Human SHPRH is a ubiquitin ligase for Mms2-Ubc13-dependent polyubiquitylation of proliferating cell nuclear antigen. *Proc Natl Acad Sci U S A* **2006**, *103* (48), 18107-12.
18. Chiu, R. K.; Brun, J.; Ramaekers, C.; Theys, J.; Weng, L.; Lambin, P.; Gray, D. A.; Wouters, B. G., Lysine 63-polyubiquitination guards against translesion synthesis-induced mutations. *PLoS Genet* **2006**, *2* (7), e116.
19. Motegi, A.; Sood, R.; Moinova, H.; Markowitz, S. D.; Liu, P. P.; Myung, K., Human SHPRH suppresses genomic instability through proliferating cell nuclear antigen polyubiquitination. *J Cell Biol* **2006**, *175* (5), 703-8.

20. Motegi, A.; Liaw, H. J.; Lee, K. Y.; Roest, H. P.; Maas, A.; Wu, X.; Moinova, H.; Markowitz, S. D.; Ding, H.; Hoeijmakers, J. H.; Myung, K., Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks. *Proc Natl Acad Sci U S A* **2008**, *105* (34), 12411-6.
21. Unk, I.; Hajdu, I.; Fatyol, K.; Hurwitz, J.; Yoon, J. H.; Prakash, L.; Prakash, S.; Haracska, L., Human HLTF functions as a ubiquitin ligase for proliferating cell nuclear antigen polyubiquitination. *Proc Natl Acad Sci U S A* **2008**, *105* (10), 3768-73.
22. Hoege, C.; Pfander, B.; Moldovan, G. L.; Pyrowolakis, G.; Jentsch, S., RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **2002**, *419* (6903), 135-41.
23. Papouli, E.; Chen, S.; Davies, A. A.; Huttner, D.; Krejci, L.; Sung, P.; Ulrich, H. D., Crosstalk between SUMO and ubiquitin on PCNA is mediated by recruitment of the helicase Srs2p. *Mol Cell* **2005**, *19* (1), 123-33.
24. Pfander, B.; Moldovan, G. L.; Sacher, M.; Hoege, C.; Jentsch, S., SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. *Nature* **2005**, *436* (7049), 428-33.
25. Moldovan, G. L.; Pfander, B.; Jentsch, S., PCNA controls establishment of sister chromatid cohesion during S phase. *Mol Cell* **2006**, *23* (5), 723-32.
26. Stelter, P.; Ulrich, H. D., Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* **2003**, *425* (6954), 188-91.
27. Zhang, S.; Chea, J.; Meng, X.; Zhou, Y.; Lee, E. Y.; Lee, M. Y., PCNA is ubiquitinated by RNF8. *Cell Cycle* **2008**, *7* (21), 3399-404.
28. Terai, K.; Abbas, T.; Jazaeri, A. A.; Dutta, A., CRL4(Cdt2) E3 ubiquitin ligase monoubiquitinates PCNA to promote translesion DNA synthesis. *Mol Cell* **2010**, *37* (1), 143-9.
29. Das-Bradoo, S.; Nguyen, H. D.; Wood, J. L.; Ricke, R. M.; Haworth, J. C.; Bielinsky, A. K., Defects in DNA ligase I trigger PCNA ubiquitylation at Lys 107. *Nat Cell Biol* **2010**, *12* (1), 74-9; sup pp 1-20.

30. Yao, N. Y.; O'Donnell, M., Replisome structure and conformational dynamics underlie fork progression past obstacles. *Curr Opin Cell Biol* **2009**, *21* (3), 336-43.
31. Kunkel, T. A.; Burgers, P. M., Dividing the workload at a eukaryotic replication fork. *Trends Cell Biol* **2008**, *18* (11), 521-7.
32. Benkovic, S. J.; Valentine, A. M.; Salinas, F., Replisome-mediated DNA replication. *Annu Rev Biochem* **2001**, *70*, 181-208.
33. Ohmori, H.; Friedberg, E. C.; Fuchs, R. P.; Goodman, M. F.; Hanaoka, F.; Hinkle, D.; Kunkel, T. A.; Lawrence, C. W.; Livneh, Z.; Nohmi, T.; Prakash, L.; Prakash, S.; Todo, T.; Walker, G. C.; Wang, Z.; Woodgate, R., The Y-family of DNA polymerases. *Mol Cell* **2001**, *8* (1), 7-8.
34. Prakash, S.; Johnson, R. E.; Prakash, L., Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu Rev Biochem* **2005**, *74*, 317-53.
35. Washington, M. T.; Carlson, K. D.; Freudenthal, B. D.; Pryor, J. M., Variations on a theme: eukaryotic Y-family DNA polymerases. *Biochim Biophys Acta* **2010**, *1804* (5), 1113-23.
36. Virdee, S.; Ye, Y.; Nguyen, D. P.; Komander, D.; Chin, J. W., Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol* **2010**, *6* (10), 750-7.
37. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B., Synthesis of proteins by native chemical ligation. *Science* **1994**, *266* (5186), 776-9.
38. Dawson, P. E.; Kent, S. B., Synthesis of native proteins by chemical ligation. *Annu Rev Biochem* **2000**, *69*, 923-60.
39. Yang, R.; Pasunooti, K. K.; Li, F.; Liu, X. W.; Liu, C. F., Synthesis of K48-linked diubiquitin using dual native chemical ligation at lysine. *Chem Commun (Camb)* **2010**, *46* (38), 7199-201.
40. Kumar, K. S.; Spasser, L.; Erlich, L. A.; Bavikar, S. N.; Brik, A., Total chemical synthesis of di-ubiquitin chains. *Angew Chem Int Ed Engl* **2010**, *49* (48), 9126-31.

41. McGinty, R. K.; Kim, J.; Chatterjee, C.; Roeder, R. G.; Muir, T. W., Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **2008**, *453* (7196), 812-6.
42. Krishna, T. S.; Kong, X. P.; Gary, S.; Burgers, P. M.; Kuriyan, J., Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **1994**, *79* (7), 1233-43.
43. Haracska, L.; Unk, I.; Prakash, L.; Prakash, S., Ubiquitylation of yeast proliferating cell nuclear antigen and its implications for translesion DNA synthesis. *Proc Natl Acad Sci U S A* **2006**, *103* (17), 6477-82.
44. Garg, P.; Burgers, P. M., Ubiquitinated proliferating cell nuclear antigen activates translesion DNA polymerases eta and REV1. *Proc Natl Acad Sci U S A* **2005**, *102* (51), 18361-6.
45. Chen, S.; Levin, M. K.; Sakato, M.; Zhou, Y.; Hingorani, M. M., Mechanism of ATP-driven PCNA clamp loading by *S. cerevisiae* RFC. *J Mol Biol* **2009**, *388* (3), 431-42.
46. Gomes, X. V.; Burgers, P. M., ATP utilization by yeast replication factor C. I. ATP-mediated interaction with DNA and with proliferating cell nuclear antigen. *J Biol Chem* **2001**, *276* (37), 34768-75.
47. Chen, J.; Ai, Y.; Wang, J.; Haracska, L.; Zhuang, Z., Chemically ubiquitylated PCNA as a probe for eukaryotic translesion DNA synthesis. *Nat Chem Biol* **2010**, *6* (4), 270-2.
48. Parker, J. L.; Bucceri, A.; Davies, A. A.; Heidrich, K.; Windecker, H.; Ulrich, H. D., SUMO modification of PCNA is controlled by DNA. *EMBO J* **2008**, *27* (18), 2422-31.
49. Freudenthal, B. D.; Gakhar, L.; Ramaswamy, S.; Washington, M. T., Structure of monoubiquitinated PCNA and implications for translesion synthesis and DNA polymerase exchange. *Nat Struct Mol Biol* **2010**, *17* (4), 479-84.
50. Freudenthal, B. D.; Ramaswamy, S.; Hingorani, M. M.; Washington, M. T., Structure of a mutant form of proliferating cell nuclear antigen that blocks translesion DNA synthesis. *Biochemistry* **2008**, *47* (50), 13354-61.
51. Chatterjee, C.; McGinty, R. K.; Fierz, B.; Muir, T. W., Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation. *Nat Chem Biol* **2010**, *6* (4), 267-9.

## Chapter 3

### A NOVEL REGULATORY MECHANISM OF UBIQUITIN SPECIFIC PROTEIN 1 BY REACTIVE OXYGEN SPECIES AND THE REQUIREMENT OF THE UAF1 SUBUNIT

#### 3.1 Introduction

##### 3.1.1 Deubiquitinase

Ubiquitin modification, like other essential forms of post-translational modification, is a dynamic, reversible process. The reverse process of ubiquitination, is called deubiquitination. A protein family, the deubiquitinating enzymes (DUBs), are responsible for cleaving the isopeptidic bond between ubiquitin and a substrate protein, or between ubiquitin molecules within a polyubiquitin chain. The important cellular roles of DUBs are manifested by the large number of DUBs identified in eukaryotes. There are approximately 100 putative DUBs in the human proteome.<sup>1</sup> The human DUBs are highly divergent in their modular structures and primary sequences, especially outside the enzyme's catalytic domain. To date little is known about the physiological function of the DUBs, and the specific substrate of most DUBs remains elusive. Collectively DUBs belong to the protease superfamily and most DUBs are cysteine proteases. A small portion of DUBs are metalloproteases.<sup>1</sup> Based on the sequence and fold of the catalytic domain, the DUBs are classified into five subclasses that are ubiquitin-specific protease (USP), ubiquitin C-terminal hydrolase (UCH), Otubain protease (OTU), Machado-Joseph disease protease (MJD) and JAB1/MPN/Mov34 metalloenzyme (JAMM). UCHs were the first reported DUBs but

their physiological functions are still under investigation. Some previous studies proposed that the UCH-L1, a UCH specifically expressed in neurons, is likely implicated in Parkinson's disease.<sup>2,3</sup> USPs consist the largest DUB family and play a variety of important cellular functions. For example, USP2a (ubiquitin-specific protease-2a) plays an important role in fatty acid synthase (FAS) metabolism. Ataxin-3 and Ataxin-3 like proteins belong to MJDs family of DUBs and deubiquitinate polyubiquitylated proteins.<sup>4</sup> The sequence similarity of the catalytic domain between Ataxin-3 and other DUBs is low. However, the NMR structure analysis showed the conservation of a catalytic triad composed of Cys, His and Asn.<sup>1</sup> Otubian-1 and Otubian-2 are the first two members reported in the OTU DUB family.<sup>5</sup> It has been found that otubian-1 can bind to the anergy factor GRAIL and serve as epistatic regulators to destabilize or stabilize GRAIL expression and resultant function.<sup>6</sup> The biological function of Otubian-2 remains unknown. DUBs that share JAB1/MPN/Mov34 metalloenzyme (JAMM) motif are classified into the JAMM family.

USP1, being a member of USP family, is an essential enzyme in both the TLS (translesion synthesis) and FA (fanconi anemia) pathways. It represents an attractive target for the simultaneous modulation of both pathways. To date only human USP1 was reported to deubiquitinate PCNA.<sup>7</sup> It has been shown that monoubiquitination of PCNA is required for the successful translesion DNA synthesis.<sup>8-10</sup> Monoubiquitination of PCNA recruits the TLS polymerase to the stalled DNA replication fork for translesion DNA synthesis. Once Pol $\eta$ -PCNA complex synthesizes past the lesion, a reverse polymerase exchange step is required to restore the normal high-fidelity DNA polymerase  $\delta$ , which is likely signaled by the

deubiquitination of PCNA.<sup>8</sup> Human USP1 is also a component of the Fanconi Anemia (FA) pathway, in which it deubiquitinates FANCD2.<sup>11</sup> Genetic and biochemical evidence demonstrated that both PCNA and FANCD2 are bona fide protein substrates of USP1.<sup>7, 11</sup> A recent *in vivo* study in a chicken DT40 cell line suggested that USP1 is indispensable for the cell's resistance to DNA damaging agents like cisplatin and mitomycin C.<sup>12</sup> Therefore disruption of USP1 function would likely sensitize cancer cells to DNA crosslinking agents.

### **3.1.2 Crosstalk between translesion synthesis and Fanconi anemia pathway**

There are close relationship between translesion synthesis and Fanconi anemia pathway. Fanconi anemia (FA) is a genome instability syndrome that is caused by the disruption of any one of the following thirteen genes: *FANCA*, *-B*, *-C*, *-DI*, *-D2*, *-E*, *-F*, *-G*, *-I*, *-J*, *-L*, *-M*, *-N*.<sup>13</sup> FA patients develop various types of malignancies in their later childhood, including acute myeloid leukemia, squamous cell carcinoma of the head and neck, and liver and brain tumors. They are also hypersensitive to DNA crosslinking agents, such as cisplatin, diepoxybutane and mitomycin C.<sup>14</sup> Evidence suggests that FA proteins could function as both signal transducers and DNA-processing molecules in a complex DNA damage detection/repair network. One important aspect of the FA pathway is the involvement of ubiquitination of FA proteins. A central protein of the FA complex, FANCD2, undergoes reversible monoubiquitination catalyzed by the FA core complex. This ubiquitination is induced by DNA damage caused by DNA crosslinker. The monoubiquitination of FANCD2 occurs specifically at Lys561 in humans and Lys563 in chickens. Ubiquitinated FANCD2 was first identified in DNA repair foci after exposure to DNA damaging agent.<sup>15, 16</sup> The colocalization of ubiquitinated FANCD2 and other DNA damage

tolerance/repair proteins, such as PCNA, REV1, RAD51, BRCA1 and BRCA2, suggests possible interplay among the FA pathway, translesion synthesis and homologous recombination.<sup>17-20</sup>

Another important FA complex protein, FANCI, is also subject to monoubiquitination. Its ubiquitination is independent of FANCD2. Besides ubiquitination, FANCI also undergoes phosphorylation. Phosphorylation at multiple sites of on FANCI was reported to facilitate the ubiquitination of FANCD2 and help to activate DNA repair pathway.<sup>21</sup> The detailed mechanism of how phosphorylated FANCI stimulate the ubiquitination of FANCD2 is still under investigation. In the absence of DNA damage, FANCD2 and FANCI are present as independent and unmodified species. Upon DNA damage, FANCD2 and FANCI separately undergo ubiquitination and phosphorylation. These modifications greatly enhance the affinity of the two proteins to the chromatin. It has been proposed that special receptors recognize the ubiquitinated and phosphorylated residues and recruit the FANCD2/FANCI complex to a specific location in the nuclear foci to carry out DNA repair.

A recent report has demonstrated crosstalk between human translesion synthesis and the Fanconi anemia pathway.<sup>22</sup> The authors showed that FANCD2 monoubiquitination requires the Rad6/Rad18-dependent PCNA monoubiquitination at K164. Specifically, K164 PCNA monoubiquitination is required to recruit FANCL to the chromatin structure and stimulate FANCL-catalyzed ubiquitination of FANCD2 and FANCI. Based on these new findings, a model was proposed to account for the molecular events that occur when a replication fork encounters a cisplatin-induced DNA lesion. Upon stalling of the replication fork, PCNA is monoubiquitinated at

K164, which in turn stimulates the recruitment of FANCL and the corresponding monoubiquitination of FANCD2.

### **3.1.3 Molecular mechanism of USP1 regulation by UAF1**

The close connection between dysregulation of USP and human diseases underscores the importance of the regulation of USP activities. The newly discovered role of UAF1 in up-regulating USP1's activity raised a new regulatory mechanism of USP by WD40-repeat protein.<sup>23</sup> The activation of USP1 activity by UAF1 can be achieved either directly through an interaction between the USP1 active site sequence and UAF1, or indirectly through an allosteric mechanism, in which UAF1 binds to a region distant to the USP1 active site. Recently, the structural determination of a yeast ubiquitin specific protease Ubp8 that is responsible for the deubiquitination of histone H2B, in complex with Sgf11 was reported.<sup>24</sup> Even Sgf11 is not related to WD40-repeat protein, the result suggested that a direct contact between the Ubp8 active site loop and the Sgf11 zinc finger motif could organize the Ubp8 active site residues into a catalytically competent conformation.<sup>24</sup> Although further study is needed to validate this structure-based prediction, this observation implies a possible mechanism of activating DUB by reorganizing the active site through direct contact with the partner protein. At present, little biochemical and structural information is available for the USP•WD40-repeat protein complex.

Recent reports on protein tyrosine phosphatase (PTP), which also contains a cysteine catalytic triad, suggested that its active site cysteine residue is prone to reversible oxidation by reactive oxygen species produced when cell is under stress.<sup>25</sup> This reversible modification of PTP active site was proposed to be a novel regulatory mechanism that could turn on or off the PTP activity in response to cellular stress. In

2005, Pagano *et al* proposed FA may be caused by oxidative stress and the ability of the mammalian cell to respond effectively to such damage.<sup>26</sup> In view of the involvement of oxidative stress in the FA phenotype I am interested in exploring the possibility that USP1 activity may be affected through reversible or irreversible oxidation of active site cysteine residue. In this chapter, we will focus on the regulation of USP1 catalytic activities by its partner protein UAF1 and by reactive oxygen species.

## 3.2 Results

### 3.2.1 Kinetic constants for USP1/UAF1 and USP1

The enzymatic activity of the USP1 alone and the USP1/UAF1 complex, was assayed using a fluorogenic ubiquitin substrate, Ub-AMC. Cleavage of the AMC group from ubiquitin results in the increase of fluorescence signal at 435 nm upon excitation at 350 nm. With this assay, we obtained steady-state enzyme kinetic rate constants for USP1 and USP1/UAF1 at pH 7.8. USP1 alone showed a low level of activity ( $k_{\text{cat}} = 0.004 \pm 0.0001 \text{ s}^{-1}$ ,  $K_{\text{m}} = 567 \pm 57.4 \text{ nM}$ ,  $k_{\text{cat}}/K_{\text{m}} = 7.1 \pm 0.7 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ ). The heterodimeric complex USP1/UAF1 demonstrated a 35-fold increase in catalytic efficiency ( $k_{\text{cat}} = 0.096 \pm 0.002 \text{ s}^{-1}$ ,  $K_{\text{m}} = 383.6 \pm 27.7 \text{ nM}$ ,  $k_{\text{cat}}/K_{\text{m}} = 2.5 \pm 0.2 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ ). The  $K_{\text{m}}$  of USP1 was only slightly decreased compared to that of the USP1/UAF1 complex, suggesting that the affinity of USP1/UAF1 complex binds Ub-AMC with a similar affinity as USP1 alone. In contrast, the catalytic turnover was greatly stimulated by UAF1 in the USP1/UAF1 complex, from  $0.004 \text{ s}^{-1}$  to  $0.096 \text{ s}^{-1}$ . As a result, the  $k_{\text{cat}}/K_{\text{m}}$  is 35-fold higher for the USP1/UAF1 complex than USP1 alone.

**Table 3.1 Kinetic parameters for USP1 and USP1/UAF1 with Ub-AMC**

<i>Enzyme</i>	$k_{cat}$	$K_m$ (nM)	$k_{cat}/K_m (\times 10^3 M^{-1}s^{-1})$	<i>Fold increase in catalytic efficiency</i>
USP1	$0.004 \pm 0.0001$	$567 \pm 57.4$	$7.1 \pm 0.7$	1
USP1/UAF1	$0.096 \pm 0.002$	$383.6 \pm 27.7$	$250 \pm 18$	35

### 3.2.2 Reversibility of inactivation by H<sub>2</sub>O<sub>2</sub>

#### 3.2.2.1 Inactivation of USP1 and USP1/UAF1 by H<sub>2</sub>O<sub>2</sub>

H<sub>2</sub>O<sub>2</sub>, a reactive oxygen species, has been shown to oxidize the sulfhydryl group in enzyme active sites to sulfenic acid reversibly and to sulfinic and sulfonic acids irreversibly (Figure 3.1). Because USP1 contains a critical catalytic cysteine residue, we expect that the enzyme will be irreversibly inactivated if the active site cysteine is oxidized into sulfinic acid and sulfonic acid (Figure 3.3). In contrast, if the USP1 active cysteine is oxidized to sulfenic acid, its activity can be regained upon reduction by DTT.

Treating USP1 with 0.25 to 1 mM H<sub>2</sub>O<sub>2</sub> resulted in 40~70% decrease in USP1's activity (Figure 3.4c). In comparison, treating USP1/UAF1 with significantly lower concentrations of H<sub>2</sub>O<sub>2</sub> (10 to 250  $\mu$ M) led to 30~70% decrease in activity (Figure 3.4a). We also found that inactivation of USP1 by hydrogen peroxide is largely reversible; above 80% activity for USP1 was regained following treatment with catalase and DTT (Figure 3.4c). In comparison, the inactivation of USP1/UAF1 is largely irreversible. Little recovery of the USP1/UAF1 activity was observed even at a lower concentration of H<sub>2</sub>O<sub>2</sub>. In the assay where glutathione was used as reducing reagents (Figure 3.4b&d), relatively lower activity was retained compared to the

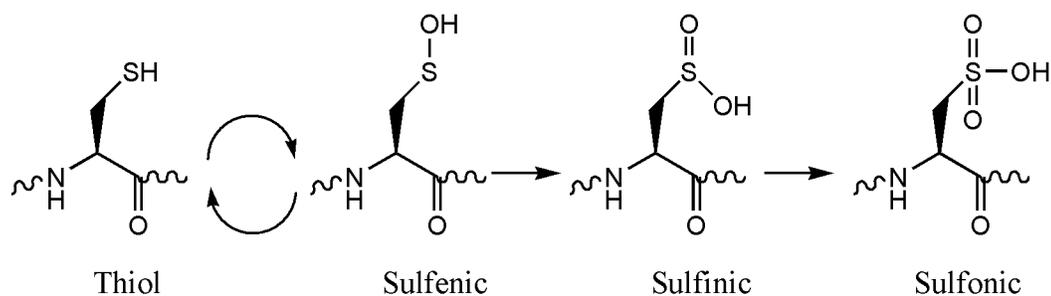
results obtained in the presence of DTT. However, a similar trend was observed, i.e. USP1/UAF1 is more sensitive than USP1 to inactivation by H<sub>2</sub>O<sub>2</sub>.

Given that inactivation of USP1 by H<sub>2</sub>O<sub>2</sub> is largely reversible, we looked into the possibility that sulfenic acid is the oxidation product of the USP1 active site cysteine. To this end, dimedone was used to trap the possible sulfenic acid intermediate in the USP1 active site as shown in Figure 3.2. Dimedone modification of sulfenic acid intermediate is irreversible, thus leading to inactivation of the enzyme. However, if the enzyme was first treated with H<sub>2</sub>O<sub>2</sub>, then followed by addition of catalase, the sulfenic acid intermediate will be reduced back to a thiol group in the presence of reducing agent. This will lead to the reactivation of the enzyme. The experimental design for the inactivation of DUBs by H<sub>2</sub>O<sub>2</sub> is shown in Figure 3.3. Control reactions were done in parallel by incubating the DUBs with only either catalase or dimedone to ensure that both reagents have minimum effect on USP1's activity. Catalase did not affect USP1's activity, while dimedone has moderate effect (10% - 20% activity decrease on fully active DUBs, data not shown).

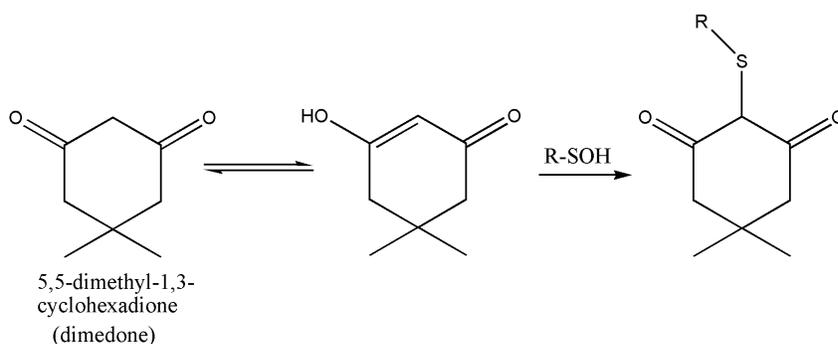
Treating USP1 with 0.25 to 1 mM H<sub>2</sub>O<sub>2</sub> followed by 6.5 mM dimedone resulted in 10~20% decrease in USP1's activity compared with only treated with H<sub>2</sub>O<sub>2</sub> (Figure 3.4c). In comparison, treating USP1/UAF1 with 10 to 250 μM of H<sub>2</sub>O<sub>2</sub> followed by dimedone led to 20~40% decrease in activity compared to USP1/UAF1 only treated with H<sub>2</sub>O<sub>2</sub> (Figure 3.4a).

USP7, another member from USP family, was studied to compare with USP1. Treating USP7 with 0.25 mM to 1 mM H<sub>2</sub>O<sub>2</sub> resulted in 20~30% decrease in USP7's activity (Figure 3.4e). Inactivation of USP7 by hydrogen peroxide is largely reversible; above 60% activity for USP7 was regained following treatment with

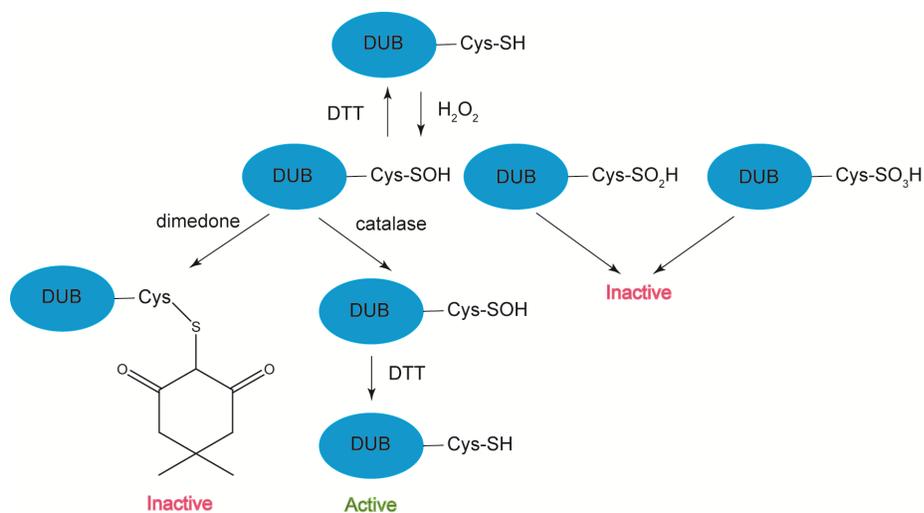
catalase and DTT. Also, a ~20% decrease was observed in USP7's activity by being treated with H<sub>2</sub>O<sub>2</sub> followed by dimedone compared with only treated with H<sub>2</sub>O<sub>2</sub>.



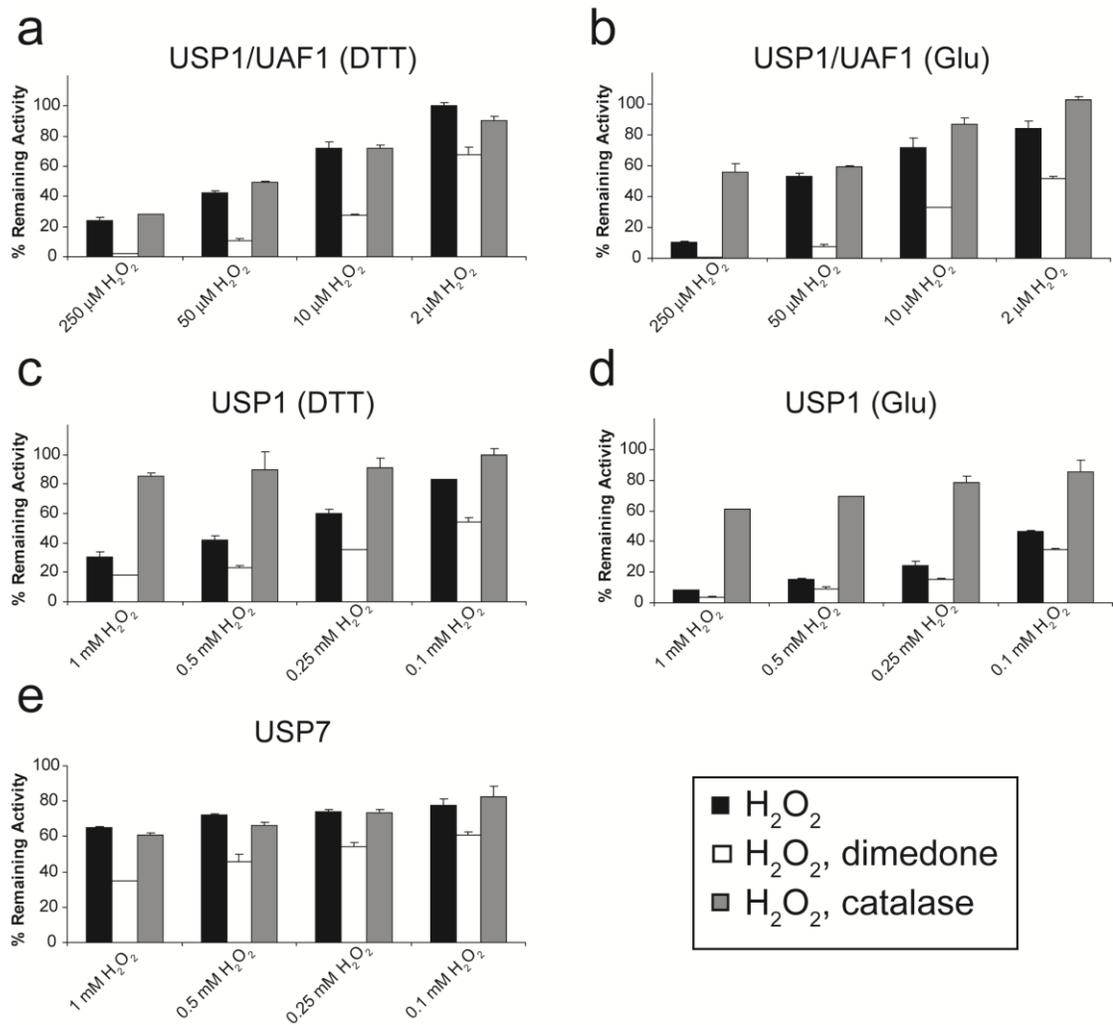
**Figure 3.1: Oxidation states of protein cysteines that are implicated in biological function. Thiol can be oxidized into sulfenic, sulfinic and sulfonic acid.**



**Figure 3.2: Selective reaction of dimedone with a sulfenic acid affords a new thioether bond.**



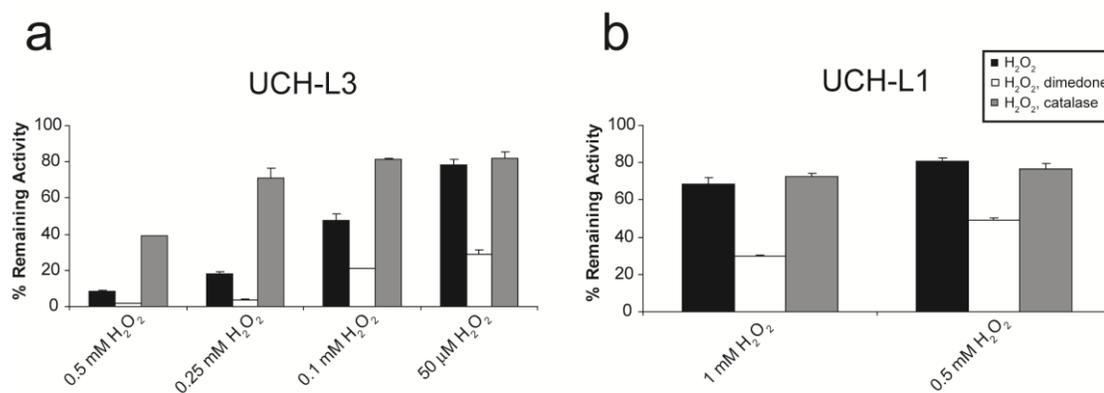
**Figure 3.3:** A flow chart shows the experimental design of the inactivation of DUBs by  $\text{H}_2\text{O}_2$ . The DUB was inactivated when it was oxidized into sulfenic acid and sulfonic acid because the oxidation is not reversible. Sulfenic acid form can be reduced back to  $-\text{SH}$  in the presence of reducing agent, like DTT. To verify the existence of sulfenic acid intermediate, dimedone was used to trap the sulfenic acid intermediate. This modification is irreversible and thus inactivating the DUB. However, if the DUB was first treated with  $\text{H}_2\text{O}_2$  then followed by catalase instead of dimedone, the sulfenic acid intermediate will be reduced back to  $-\text{SH}$  in the presence of DTT. This will lead to the reactivation of the DUB.



**Figure 3.4:** Inactivation and reactivation of various DUBs. DUBs inactivated by 2 μM – 1 mM H<sub>2</sub>O<sub>2</sub> treatment (black bar) is diluted into either DTT or glutathione and monitored for recovery of original activity. Dimedone was used to trap sulfenic acid form after H<sub>2</sub>O<sub>2</sub> treatment (white bar) or catalase was used to degrade H<sub>2</sub>O<sub>2</sub> following H<sub>2</sub>O<sub>2</sub> treatment (grey bar). The plotted data represents mean value ± s.d. (a) USP1/UAF1 in DTT (b) USP1/UAF1 in glutathione (c) USP1 in DTT (d) USP1 in glutathione (e) USP7 in DTT.

### 3.2.2.2 Inactivation of UCHs by H<sub>2</sub>O<sub>2</sub>

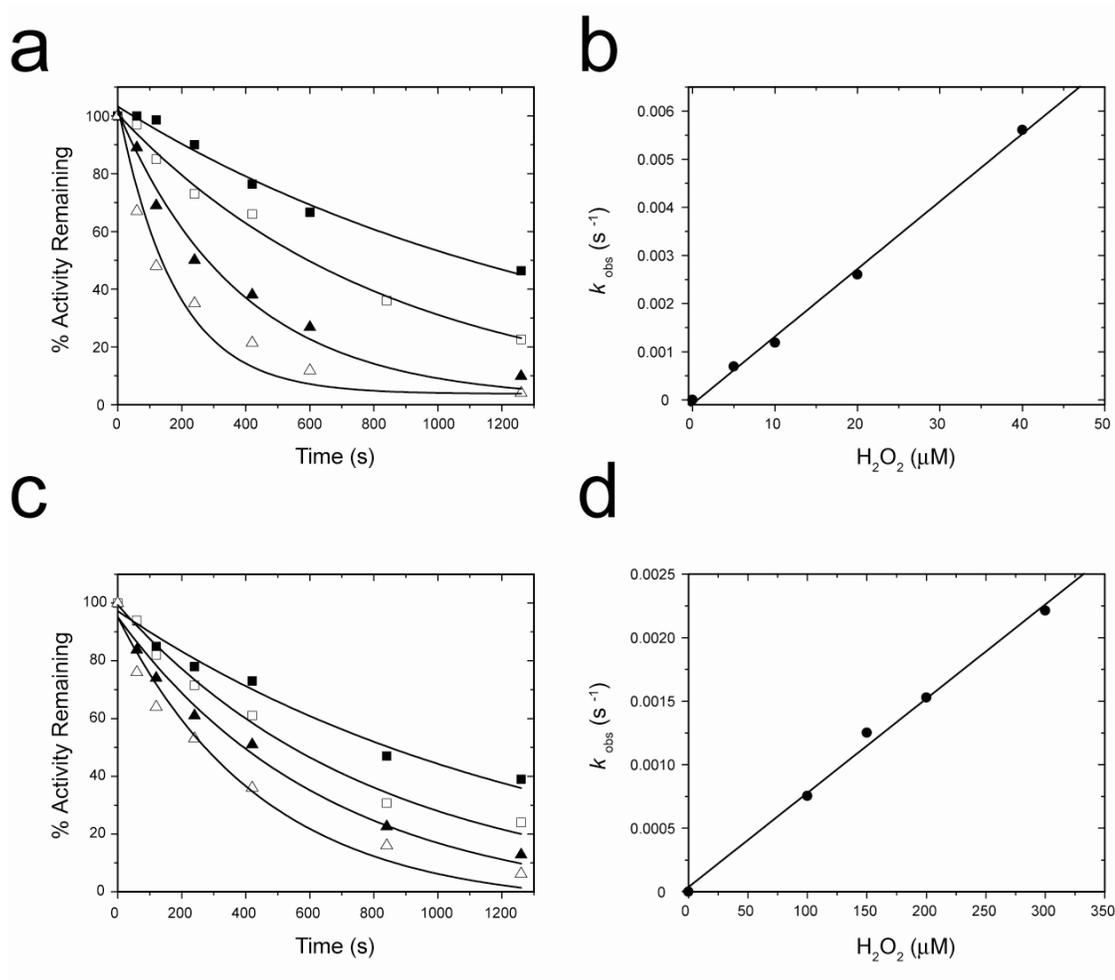
The UCH family DUBs adopt a different fold from the USP family DUBs. Nonetheless, the UCH family DUB also utilize a catalytic triad for catalysis. UCH-L1 and UCH-L3 are two well-studied DUBs. Treating UCH-L1 with 0.5 mM to 1 mM H<sub>2</sub>O<sub>2</sub> resulted in 20~30% decrease in UCH-L1's activity (Figure 3.5b). In comparison, treating UCH-L3 with lower concentrations of H<sub>2</sub>O<sub>2</sub> (0.1 mM to 0.5 mM) led to 50~90% decrease in activity (Figure 3.5a). We also found that inactivation of UCH-L1 and UCH-L3 by hydrogen peroxide is largely reversible; above 60% activity for UCH-L1 and above 40% activity for UCH-L3 was regained following treatment with catalase and DTT.



**Figure 3.5: Inactivation and reactivation of (a) UCH-L3 and (b) UCH-L1. UCHs inactivated by 50 μM – 1 mM H<sub>2</sub>O<sub>2</sub> treatment (black bar) is diluted into DTT and monitored for recovery of original activity. Dimedone was used to trap sulfenic acid form after H<sub>2</sub>O<sub>2</sub> treatment (white bar) or catalase was used to degrade H<sub>2</sub>O<sub>2</sub> following H<sub>2</sub>O<sub>2</sub> treatment (grey bar). The plotted data represents mean value ± s.d.**

### 3.2.3 Time- and concentration-dependent inhibition of USP1/UAF1 by H<sub>2</sub>O<sub>2</sub>

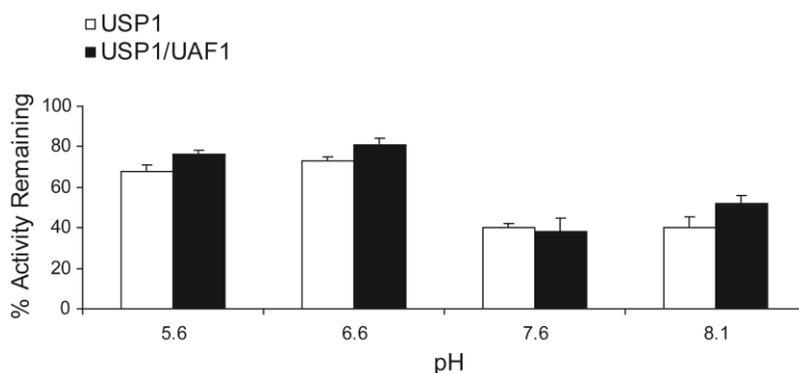
To obtain kinetic information of the inactivation of USP1/UAF1 by H<sub>2</sub>O<sub>2</sub>, time-dependent inactivation of USP1/UAF1 and USP1 by different concentrations of H<sub>2</sub>O<sub>2</sub> was assayed as described in Materials and Methods section. The data were fit to Equation 1 (see Materials and Methods) to obtain the pseudo first-order inactivation rates,  $k_{\text{obs}}$ . Secondary plots of these rates against the concentration of H<sub>2</sub>O<sub>2</sub> gave a linear plot, indicating a second-order rate constant of 140.33 M<sup>-1</sup> s<sup>-1</sup> for USP1/UAF1 (Figure 3.6a&b) and 7.41 M<sup>-1</sup> s<sup>-1</sup> for USP1 (Figure 3.6c&d). The inactivation of USP1/UAF1 occurs 19 fold faster than that of USP1 alone. Compared with other Cys dependent enzymes known to undergo redox regulation, the inactivation rate for USP1/UAF1 is similar to Cdc25B (164 M<sup>-1</sup> s<sup>-1</sup>).<sup>27</sup>



**Figure 3.6:** Time- and concentration-dependent inhibition of USP1/UAf1 and USP1 by H<sub>2</sub>O<sub>2</sub>. (a) Time-dependent loss in USP1/UAf1 activity is observed after pre-incubation with 5 (■), 10 (□), 20 (▲), 40 (△) μM H<sub>2</sub>O<sub>2</sub>, with observed inactivation rates of 0.69, 1.19, 2.60 and 5.61 × 10<sup>-3</sup> s<sup>-1</sup>, respectively. Fits are obtained using Equation 1. (b) Observed pseudo first-order inactivation rates for USP1/UAf1 vary linearly with H<sub>2</sub>O<sub>2</sub> concentration and are fit using a second order rate constant of 140.33 M<sup>-1</sup> s<sup>-1</sup>. (c) Time-dependent loss in USP1 activity is observed after pre-incubation with 100 (■), 150 (□), 200 (▲), 300 (△) μM H<sub>2</sub>O<sub>2</sub>, with observed inactivation rates of 0.75, 1.25, 1.53 and 2.22 × 10<sup>-3</sup> s<sup>-1</sup>, respectively. Fits are obtained using Equation 1. (d) Observed pseudo first-order inactivation rates for USP1 vary linearly with H<sub>2</sub>O<sub>2</sub> concentration and are fit using a second order rate constant of 7.41 M<sup>-1</sup> s<sup>-1</sup>.

### 3.2.4 Inactivation of USP1/UAF1 and USP1 by iodoacetamide

As discussed in section 3.2.2-3.3.3, USP1/UAF1 is more sensitive to inactivation by H<sub>2</sub>O<sub>2</sub> compared to USP1 alone. Next, we want to explore whether they show similar behavior towards iodoacetamide, another commonly used cysteine-reactive molecule. Figure 3.7 clearly shows that treatment of USP1 and USP1/UAF1 with iodoacetamide led to inactivation of both species. However, the difference in their reactivity to iodoacetamide is small from pH 5.6 to 8.1.

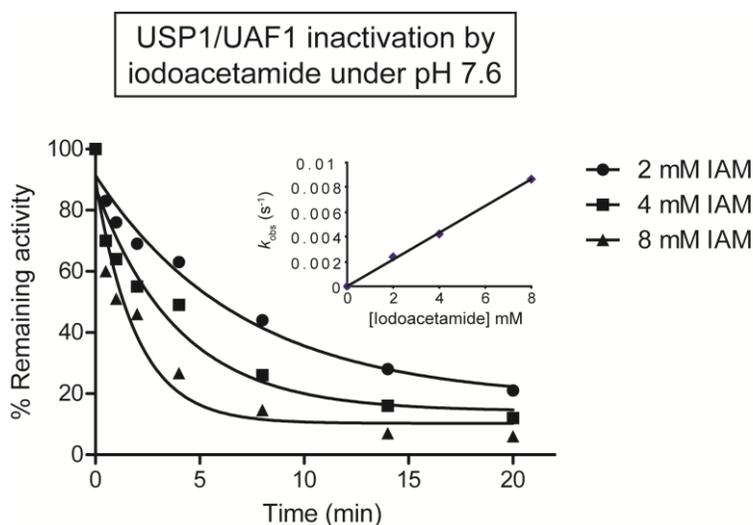


**Figure 3.7:** Inactivation of USP1 (white bar) and USP1/UAF1 (black bar) by iodoacetamide under different pHs. Both species were inactivated by 4 mM iodoacetamide for 2 min under room temperature. The remaining activity was monitored by diluting the treated DUBs into an assay buffer containing Ub-AMC and DTT. The plotted data represents mean value  $\pm$  s.d.

Since this result is determined under a single concentration of iodoacetamide at a single time point, a rigorous determination of the inactivation kinetics at several different iodoacetamide concentrations, different time points and several different pHs will likely reveal difference between USP1 and USP1/UAF1.

### 3.2.5 pH dependence of inactivation of USP1/UAF1 by iodoacetamide

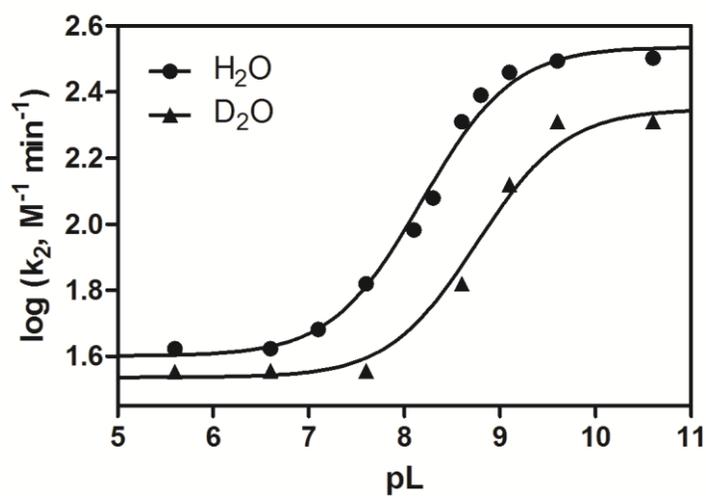
The reactivity of the thiol group of USP1/UAF1 was investigated by measuring the second order rate constant of its inactivation by alkylating reagent iodoacetamide. A typical data set of inactivating USP1/UAF1 at a given pH with varying concentration of iodoacetamide was shown in Figure 3.8.



**Figure 3.8:** Inactivation of USP1/UAF1 by iodoacetamide under pH 7.6. Time-dependent inactivation of USP1/UAF1 activity was observed after pre-incubation with (A) 2 (●), 4 (■), 8 (▲) mM iodoacetamide. The pseudo first-order inactivation rate constants are plotted against iodoacetamide concentration as shown in the inset. The inactivation rate constant ( $k_2$ ) obtained is  $1.1 \text{ M}^{-1} \text{ s}^{-1}$  ( $66 \text{ M}^{-1} \text{ min}^{-1}$ ).

Figure 3.9 shows the pH dependence of USP1/UAF1 inactivation at  $25 \text{ }^\circ\text{C}$  by iodoacetamide over the pH range of 5.6-10.6, which led to the determination of a thiol  $pK_a$  value of  $8.6 \pm 0.1$  by using equation 2 (see Materials and Methods). The inactivation of the enzyme by iodoacetamide was also measured in  $\text{D}_2\text{O}$  at several pL values between 5.6 and 10.6. The fitted  $pK_a$  was 9.2, representing a shift of 0.6 pL unit

compared to that determined in H<sub>2</sub>O. A normal kinetic solvent isotope effect of 1.5 was observed (Table 3.2).



**Figure 3.9:** Plot of  $\log(k_2)$  measured for the inactivation of USP1/UAF1 with iodoacetamide vs reaction solution pH fitted to equation 2 to define the  $pK_a$ . The points marked with (●) signs were determined in H<sub>2</sub>O and with (▲) were determined in 98% deuterium oxide.

**Table 3.2 Inactivation rate constants ( $M^{-1} \text{ min}^{-1}$ ) measured in  $H_2O$  or  $D_2O$  under different pLs (H or D). ND indicated not determined.**

pL (H or D)	$k_2^H$ ( $M^{-1} \text{ min}^{-1}$ )	$k_2^D$ ( $M^{-1} \text{ min}^{-1}$ )	$k_2^H / k_2^D$
5.6	41.69	35.48	1.17
6.6	41.69	36.31	1.17
7.1	47.86	ND	ND
7.6	66.07	36.31	1.83
8.1	95.50	ND	ND
8.3	120.22	ND	ND
8.6	204.17	66.07	3.09
8.8	251.19	ND	ND
9.1	288.40	131.83	2.18
9.6	309.03	204.17	1.53
10.6	316.23	204.17	1.56

### 3.2.6 Probe the sulfenic acid species in USP1/UAF1 upon oxidation by DAz1-Biotin

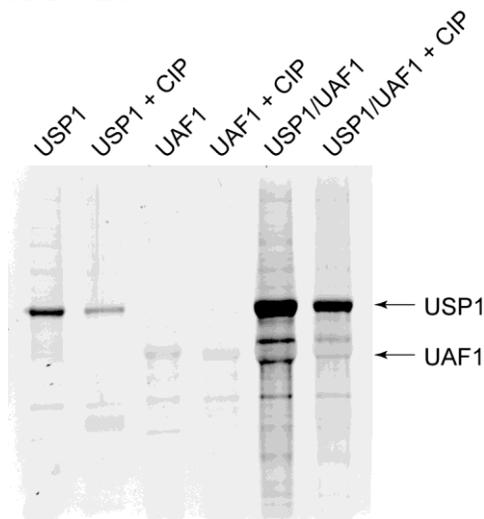
DAz1-Biotin (Figure 3.10a) is a dimedone derivative that was shown to specifically modify sulfenic acid by Kate Carroll and coworkers.<sup>28</sup> The sulfenic acid trapped by DAz1-Biotin can be detected by HRP-streptavidin in western blot (Figure 3.10b). We obtained the DAz1-Biotin from the Carroll lab at Scripps Florida. First, USP1/UAF1 was treated with TCEP to reduce the enzyme. P-6 column was used to remove TCEP followed by the addition of moderate concentration of  $H_2O_2$  (20  $\mu M$  and 40  $\mu M$ ) and large excess of DAz1-Biotin. However, weak labeling of USP1/UAF1



**Figure 3.11: Western blot of DAZ1-Biotin labeled USP1/UAF1. USP1/UAF1 was treated with or without H<sub>2</sub>O<sub>2</sub> (20 or 40 μM). It was supposed to see increased USP1 band intensity with the increased H<sub>2</sub>O<sub>2</sub> concentration. P-6 column was used to remove reducing reagent TCEP here. The reaction treated without P-6 column was used as a negative control in which no difference was observed between with or without H<sub>2</sub>O<sub>2</sub>.**

### 3.2.7 Phosphorylation of USP1/UAF1

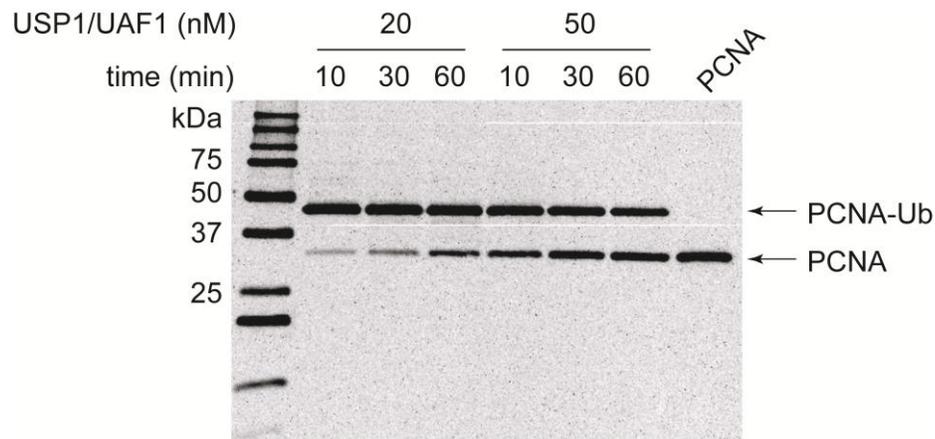
The large difference in the catalytic activities of USP1 and USP1/UAF1 complex prompted us to investigate the role of UAF1 in regulating USP1's activity. In order to explore whether the post-translational modification of either USP1 or UAF1 affects the complex formation, or the activity of USP1, we examined the phosphorylation of USP1 alone, UAF1 alone and USP1/UAF1 complex. Using a phosphoprotein-specific dye, Pro-Q diamond stain (Sigma-Aldrich), we showed that the USP1 was phosphorylated when expressed either alone or as a complex with UAF1, while UAF1 was not phosphorylated (Figure 3.12). Treating USP1 and USP1/UAF1 with calf intestinal phosphatase resulted in decrease in the intensity of the phosphorylated USP1 band.



**Figure 3.12: Phosphorylation of the USP1 subunit.** A polyacrylamide gel containing USP1, UAF1 and USP1/UAF1 with or without calf intestinal alkaline phosphatase (CIP) were stained with Pro-Q Diamond phosphoprotein stain.

### 3.2.8 *In vitro* deubiquitination of monoubiquitinated human PCNA (hPCNA-Ub) by USP1/UAF1

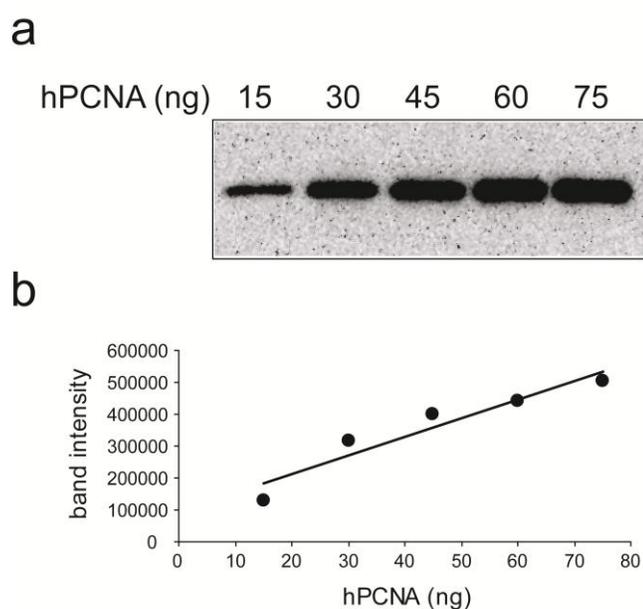
*In vitro* deubiquitination of monoubiquitinated human PCNA (hPCNA-Ub) by USP1/UAF1 has not been demonstrated. Here we showed that USP1/UAF1 indeed deubiquitinates native hPCNA-Ub *in vitro* (Figure 3.13), agreeing with the *in vivo* finding.<sup>29</sup> The hPCNA-Ub was obtained from the Haracska lab at Institute of Genetics, Hungarian Academy of Sciences.



**Figure 3.13: Time-dependent deubiquitination of native hPCNA-Ub by USP1/UAF1.** 200 ng native hPCNA-Ub were incubated with 20 nM (36 ng) or 50 nM (90 ng) USP1/UAF1 for 10, 30, and 60 mins. hPCNA-Ub and hPCNA were detected by anti-PCNA antibody.

To quantify deubiquitination of monoubiquitinated hPCNA by USP1/UAF1, we used denaturing SDS-PAGE and Western blotting (with anti-hPCNA

antibody) to quantify the reaction product. We first demonstrated the correlation between the Western blotting signal and known amounts of hPCNA. A linear response was observed (Figure 3.14) between 15 and 75 ng hPCNA. This allows us to quantify the deubiquitination reaction. We measured the specific activity of USP1/UAF1 deubiquitinating hPCNA-Ub at 20 nM USP1/UAF1 and 520 nM hPCNA-Ub. The estimated specific activity was approximately 0.002 to 0.006 s<sup>-1</sup>.



**Figure 3.14: The correlation between the Western blotting signal and known amounts of hPCNA. (a) 15 – 75 ng hPCNA was loaded on SDS-PAGE and detected by Western blotting (with anti-PCNA antibody). (b) A linear response was observed between 15 and 75 ng hPCNA by plotting the known amount of hPCNA against the Western blotting signal.**

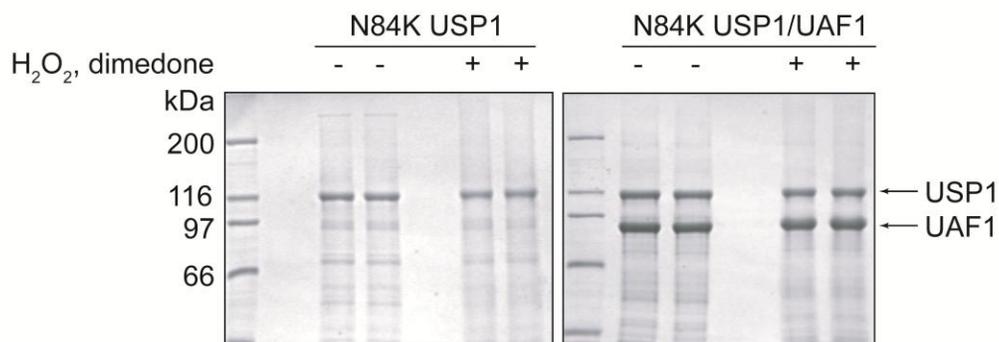
### 3.2.9 Mass spectrometry analysis of USP1/UAF1 following in-gel digestion

To obtain direct evidence that the USP1 active site cysteine is oxidized to sulfenic acid by  $\text{H}_2\text{O}_2$ , we carried out mass spectrometric analysis of the tryptic fragment of USP1 or USP1/UAF1 treated sequentially with hydrogen peroxide and dimedone. In this study, an N84K USP1 mutant was subjected to tryptic digestion. To facilitate mass spectrometry analysis, Lysine was introduced at position 84 in USP1 to generate a smaller tryptic fragment that contains active site cysteine 90. N84K USP1 alone showed a low level of activity ( $k_{\text{cat}} = 0.00523 \pm 0.0004 \text{ s}^{-1}$ ,  $K_{\text{m}} = 683.6 \pm 151.2 \text{ nM}$ ,  $k_{\text{cat}}/K_{\text{m}} = 7.7 \pm 0.18 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ ). The N84K USP1/UAF1 complex demonstrated a 13-fold increase in catalytic efficiency ( $k_{\text{cat}} = 0.0838 \pm 0.0034 \text{ s}^{-1}$ ,  $K_{\text{m}} = 862.1 \pm 85.75 \text{ nM}$ ,  $k_{\text{cat}}/K_{\text{m}} = 9.7 \pm 1.0 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ ) than N84K USP1. Based on the kinetic parameters, the N84K USP1 has identical catalytic efficiency as wild-type USP1 based on the  $k_{\text{cat}}/K_{\text{m}}$ . However, N84K USP1/UAF1 has 2.6-fold lower catalytic efficiency than wild-type USP1/UAF1 due to decreased binding affinity ( $K_{\text{m}}$ ). N84K USP1 or N84K USP1/UAF1 complex was treated sequentially with  $\text{H}_2\text{O}_2$  and dimedone. The treated protein sample was separated on an 8% denaturing SDS-PAGE (Figure 3.15a). The SDS-PAGE samples were subjected to in-gel digestion followed by MALDI-MS by analysis in Kelvin Lee's lab. The results were summarized in Table 3.3. Compared to the theoretical  $m/z$  value listed in Figure 3.15b, we found both cysteines in the tryptic fragment of USP1 undergo iodoacetamide modification. The  $m/z$  of 2621.29, 2621.28 were detected in untreated and treated samples (treated with  $\text{H}_2\text{O}_2$  followed by dimedone), respectively. The measured molecular weights of both fragments are almost identical to the theoretical  $m/z$  value of 2621.31. Further, both fragments were subjected to MS/MS (Figure 3.16) to confirm the peptide sequence. For N84K USP1/UAF1 sample, the peak with a  $m/z$  of 2621 (Table 3.3) was detected in both

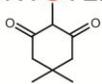
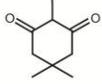
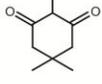
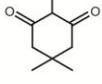
untreated and treated samples. Iodoacetamide was introduced during the mass spectrometry analysis to prevent the cysteines from forming disulfide bond, which complicated the interpretation of the mass spectrometry results. No dimeric modified species was detected. This result indicates that either the active site cysteine was not modified by dimeric or the dimeric trapped species was not abundant enough to be detected by MALDI-MS.

To simplify the interpretation of the mass spectrometry data, iodoacetanilide (Figure 3.17b) was used to modify the proteins before mass spec analysis, during which iodoacetamide was introduced. Since iodoacetanilide has a larger mass than iodoacetamide, a 133 mass shift should be observed for iodoacetanilide modification. A 57 mass shift will be observed for iodoacetamide modification. Also, C102 was replaced by serine in N84K USP1 (Figure 3.17a) to simplify the analysis since two cysteines in one fragment generates many modification combinations as in shown in Figure 3.16b. From the results in Table 3.3, we found C90 in the tryptic fragment of USP1 undergoes iodoacetamide modification. The  $m/z$  of 2548 were detected from untreated and treated samples (treated with  $H_2O_2$  followed by dimeric or treated with iodoacetanilide). All three numbers are almost identical with the theoretical  $m/z$  value of 2548.30 (Figure 3.17b). Further, all three fragments were subjected to MS/MS (data not shown here) and it confirms the peptide sequence. No iodoacetanilide modification was observed in the iodoacetanilide treated sample.

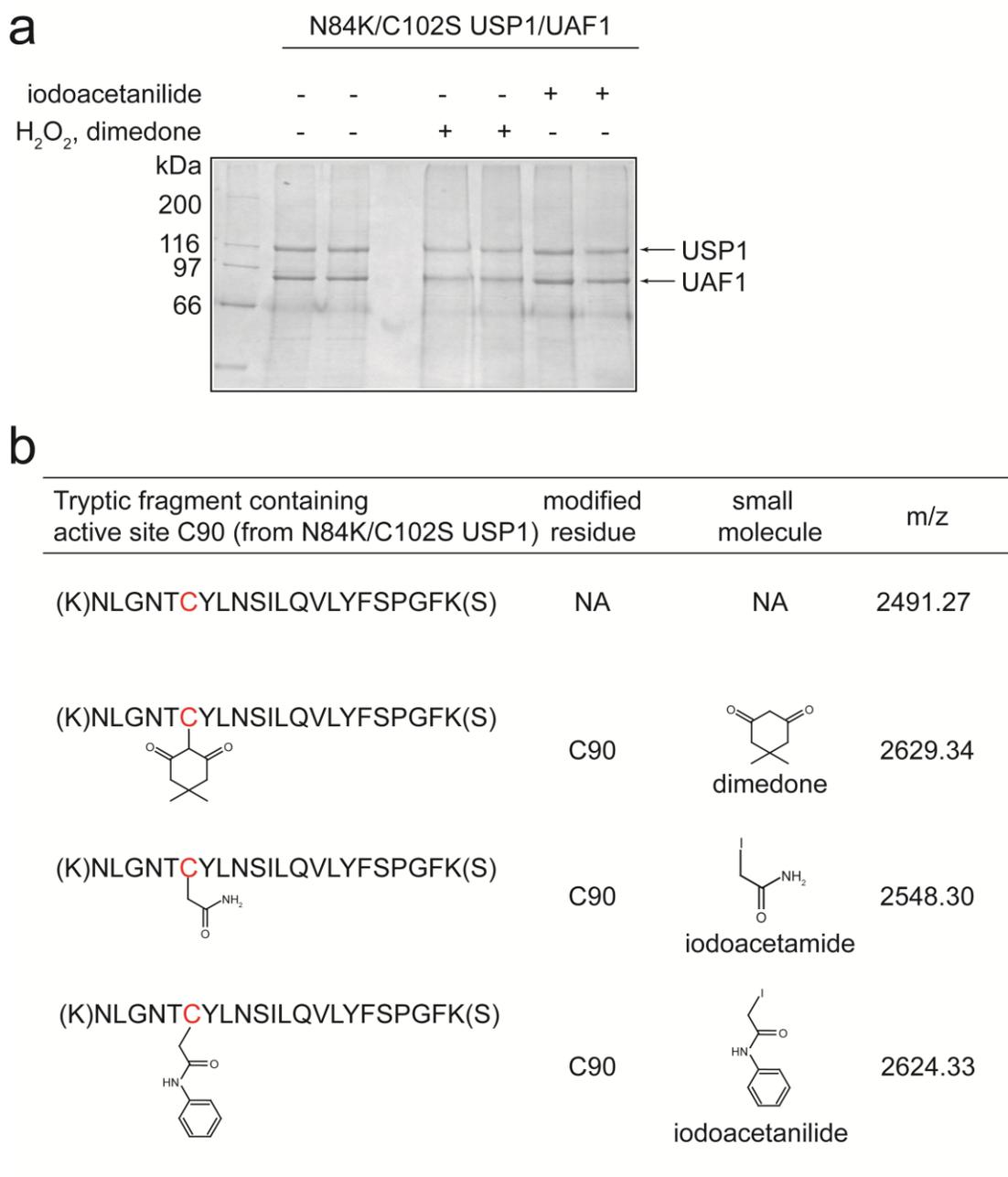
**a**



**b**

Tryptic fragment containing active site C90 (from N84K USP1)	modified residue	small molecule	m/z
(K)NLGNTC <sup>+</sup> YLNSILQVLYFCPGFK(S)	NA	NA	2507.25
(K)NLGNTC <sup>+</sup> YLNSILQVLYFCPGFK(S) 	C90	dimedone	2645.43
(K)NLGNTC <sup>+</sup> YLNSILQVLYFCPGFK(S) 	C102	dimedone	2645.43
(K)NLGNTC <sup>+</sup> YLNSILQVLYFCPGFK(S)  	C90, C102	dimedone	2783.60
(K)NLGNTC <sup>+</sup> YLNSILQVLYFCPGFK(S)  	C90, C102	iodoacetamide	2621.31





**Figure 3.17: (a) SDS-PAGE of N84K/C102S USP1/UAF1 complex treated without, with H<sub>2</sub>O<sub>2</sub> followed by dimedone, or with iodoacetanilide. (b) The trypsin digested fragment containing active site cysteine 90 (in red) can be modified by dimedone, iodoacetamide and iodoacetanilide. Only C90 in this fragment can be modified and the corresponding m/z value was listed.**

**Table 3.3 Mass spectrometry results of various USP1 samples subjected to different modification.**

sample	treatment	peak of interest detected (m/z)	MS/MS confirmed
N84K USP1	no	2621.29	yes
N84K USP1	H <sub>2</sub> O <sub>2</sub> , dimedone	2621.28	yes
N84K USP1/UAF1	no	2621.42	yes
N84K USP1/UAF1	H <sub>2</sub> O <sub>2</sub> , dimedone	2621.45	yes
N84K/C102S USP1/UAF1	no	2548.10	yes
N84K/C102S USP1/UAF1	H <sub>2</sub> O <sub>2</sub> , dimedone	2548.11	yes
N84K/C102S USP1/UAF1	iodoacetanilide	2548.10	yes

### 3.3 Discussion

#### 3.3.1 UAF1 acts as an activator of USP1

UAF1 is a WD40 repeat containing protein and was reported to interact with and stimulate the activity of USP1, USP12 and USP46.<sup>23, 30</sup> The kinetic data (Table 3.1) suggests that UAF1 does not affect the binding affinity of USP1 to its substrate, ubiquitin. However, the catalytic turnover was greatly increased by UAF1 (the  $k_{cat}/K_m$  value for the the USP1/UAF1 complex is 35-fold higher compared to USP1 alone). Thus, UAF1 does not increase the affinity of USP1 for ubiquitin. Instead, USP1 may undergo a conformational change upon association with UAF1. It should be noted that  $K_m$  value was used to estimate the ubiquitin binding affinity. This

needs to be further validated by binding studies that can directly measure the affinity between USP1 and Ub.

### **3.3.2 USP1/UAF1 complex is more sensitive to H<sub>2</sub>O<sub>2</sub>**

We found that USP1/UAF1 is more sensitive to inactivation by H<sub>2</sub>O<sub>2</sub> compared to USP1 alone at pH 7.8. The inactivation of USP1 by H<sub>2</sub>O<sub>2</sub> is largely reversible; above 80% activity of USP1 was regained following treatment with catalase and DTT (Figure 3.3). In comparison, the inactivation of USP1/UAF1 is largely irreversible. Little recovery of the USP1/UAF1 activity was observed even at a lower concentration of H<sub>2</sub>O<sub>2</sub>. Also, a second-order rate constant of 140.33 M<sup>-1</sup> s<sup>-1</sup> was obtained for USP1/UAF1 and 7.41 M<sup>-1</sup> s<sup>-1</sup> for USP1 alone by H<sub>2</sub>O<sub>2</sub>. Next, we compared USP1's inactivation rate by H<sub>2</sub>O<sub>2</sub> to other enzymes that contain active site cysteines (Table 3.4). Caspase-3 is a cysteine protease. Cdc25B and PTP1B are tyrosine phosphatases. DDAH-1 is dimethylarginine dimethylaminohydrolase. The rate of inactivation of USP1/UAF1 by H<sub>2</sub>O<sub>2</sub> is comparable to that of Cdc25B. In contrast, USP1 is between PTP1B and DDAH-1.

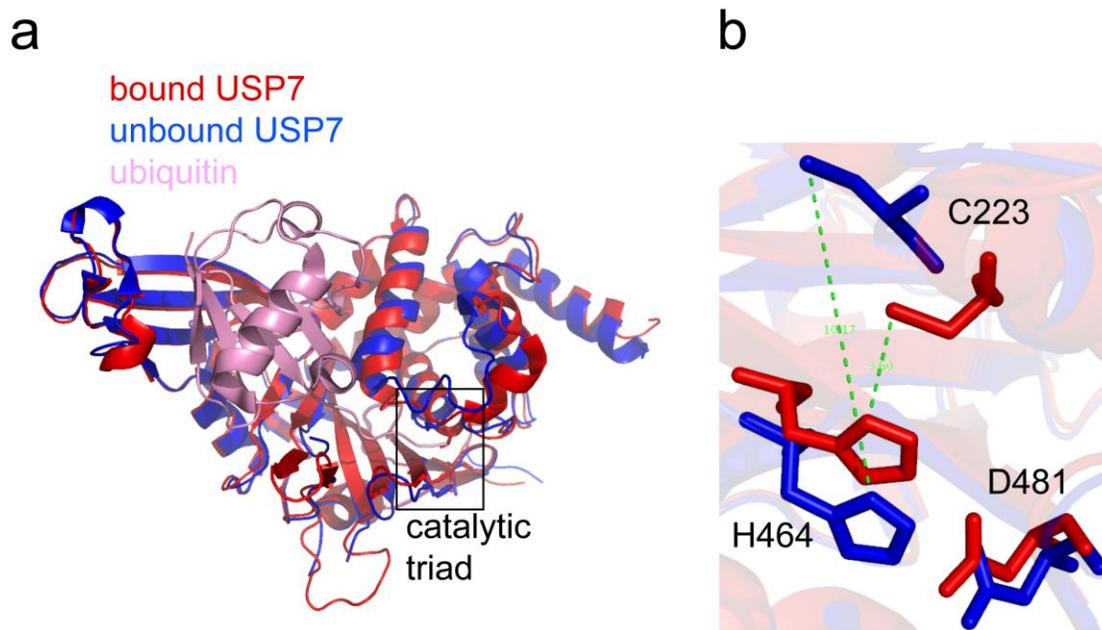
**Table 3.4 Inactivation rates by H<sub>2</sub>O<sub>2</sub> of various proteins containing active site cysteines**

<b>Enzyme</b>	<b>Inactivation rate constant (M<sup>-1</sup> s<sup>-1</sup>)</b>
Caspase-3 <sup>31</sup>	750
Cdc25B <sup>27</sup>	164
PTP1B <sup>32</sup>	43
DDAH-1 <sup>33</sup>	0.088
USP1	7.41
USP1/UAF1	140.33

These observations suggest that a catalytically important cysteine thiol in USP1 and USP1/UAF1 possesses different reactivity towards H<sub>2</sub>O<sub>2</sub>. To better explain why USP1/UAF1 shows a higher sensitivity than USP1 alone, we propose a model based on the USP7 structure (no USP1 structure has been reported so far). USP1 and USP7 belong to the same USP family and share 17% sequence identity. The ubiquitin bound USP7 aligns very well with the unbound one with an RMSD value of 1.21 Å (Figure 3.18a). Notably, the active site in the unbound form (blue) of USP7 is misaligned (Figure 3.18b). . The distance from the N<sup>δ1</sup> of His 464 to the sulfur atom of Cys 223 is 10.17 Å, too far for hydrogen binding. Upon ubiquitin binding, structural elements surrounding the catalytic cleft undergo dramatic changes that realign the active site residues for productive catalysis (Figure 3.18b, red). The distance from the N<sup>δ1</sup> of His 464 to the sulfur atom of Cys 223 changes to 3.60 Å.

Hence, we hypothesize that for USP1 alone may have a misaligned active site which is less reactive with H<sub>2</sub>O<sub>2</sub>. Repositioning of the catalytic triad by the binding of UAF1 may facilitate the deprotonation of the active site cysteine which

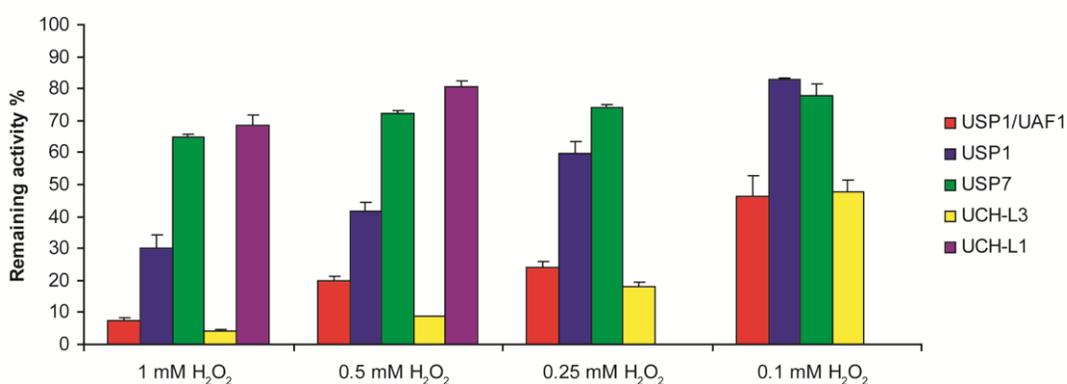
dramatically increases the catalytic efficiency of USP1 through re-aligning the active site. Since thiolate has a higher reactivity than thiol group with  $H_2O_2$ , the USP1/UAF1 active site is more sensitive to  $H_2O_2$ .



**Figure 3.18:** (a) Aligned crystal structure of USP7 bound (red, PDB: 1NBF) with ubiquitin (pink) and unbound USP7 (blue, PDB: 1NB8). Catalytic triad area is boxed. (b) Zoomed-in catalytic triad of bound and unbound USP7. The distance from the  $N^{\delta 1}$  of His 464 to the sulfur atom of Cys 223 is 10.17 Å in unbound form while it is 3.60 Å in the bound form.

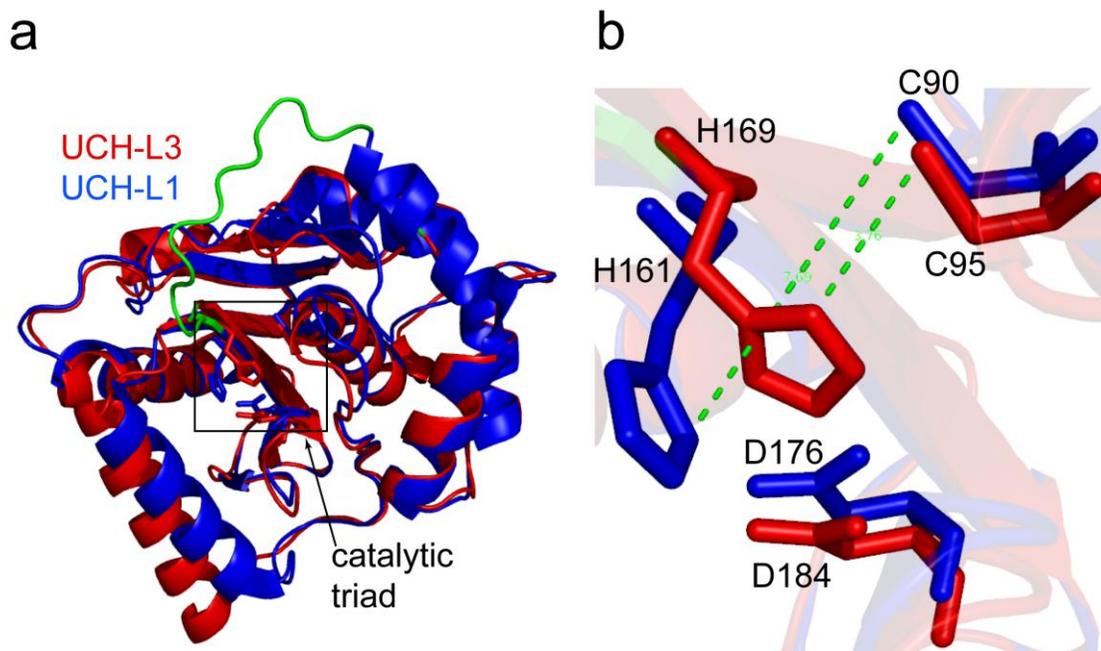
To answer the question of whether other DUBs are also sensitive to  $H_2O_2$ , we tested USP7 from the USP family and UCH-L3, UCH-L1 from the UCH family. We found that USP7 is less sensitive to  $H_2O_2$  compared to USP1 alone. More than 60% activity was observed under 1 mM  $H_2O_2$  treatment. However, we found that UCH-L3 is very sensitive to  $H_2O_2$ . Substantial activity loss was observed even under

50  $\mu\text{M}$   $\text{H}_2\text{O}_2$ . In contrast, UCH-L1 is largely insensitive to  $\text{H}_2\text{O}_2$ . The relative sensitivity to  $\text{H}_2\text{O}_2$  was summarized in Figure 3.19. Among the DUBs tested, USP1/UAF1 and UCH-L3 are most sensitive to oxidation under all the  $\text{H}_2\text{O}_2$  concentrations tested. In contrast, USP7 and UCH-L1 are largely insensitive to  $\text{H}_2\text{O}_2$  treatment. USP1 alone is in between of USP7 and UCH-L1 in terms of  $\text{H}_2\text{O}_2$  sensitivity.



**Figure 3.19: Comparison of sensitivity to  $\text{H}_2\text{O}_2$  of various DUBs. Each DUB's remaining activity was plotted against different concentrations of  $\text{H}_2\text{O}_2$ .**

UCH-L1 and UCH-L3 both belong to UCH family and they share a sequence identity of 51%. Their crystal structures aligned with each other very well with an RMSD value of 1.3  $\text{\AA}$  (Figure 3.20a).



**Figure 3.20:** (a) Aligned crystal structure of UCH-L3 (red, PDB: 1UCH) with UCH-L1 (blue, PDB: 2ETL). Catalytic triad area is boxed. The loop in green from UCH-L1 distinguishes itself from UCH-L3 in which no such loop was found. (b) Zoomed-in catalytic triad of UCH-L3 and UCH-L1. The distance from the  $N^{\delta 1}$  of His 169 to the sulfur atom of Cys 95 is 3.76 Å in UCH-L3 while it is 7.69 Å in UCH-L1 from the N atom of His 161 to the S atom of Cys 90.

However, a close inspection of the active site (Figure 3.20b) revealed interesting differences in the active sites of UCH-L1 and UCH-L3. In UCH-L1, the distance between the  $N^{\delta 1}$  of His 161 and the sulfur atom of Cys 90 is 7.69 Å, compared to a distance of 3.76 Å measured between the  $N^{\delta 1}$  of His 169 and the sulfur atom of Cys 95 in UCH-L3. This suggests that the UCH-L3 active site likely contains a thiolate-imidazolium ion pair. Given that thiolate is more susceptible to oxidation by  $H_2O_2$  than sulfhydryl group, it is understandable that UCH-L3 is highly sensitive to inactivation by  $H_2O_2$ . In comparison, the large distance between the triad Cys and His

residues in unbound UCH-L1 prevents the formation of a thioate-imidazolium ion pair. This agrees with the lower sensitivity of UCH-L3 inactivation by H<sub>2</sub>O<sub>2</sub>.

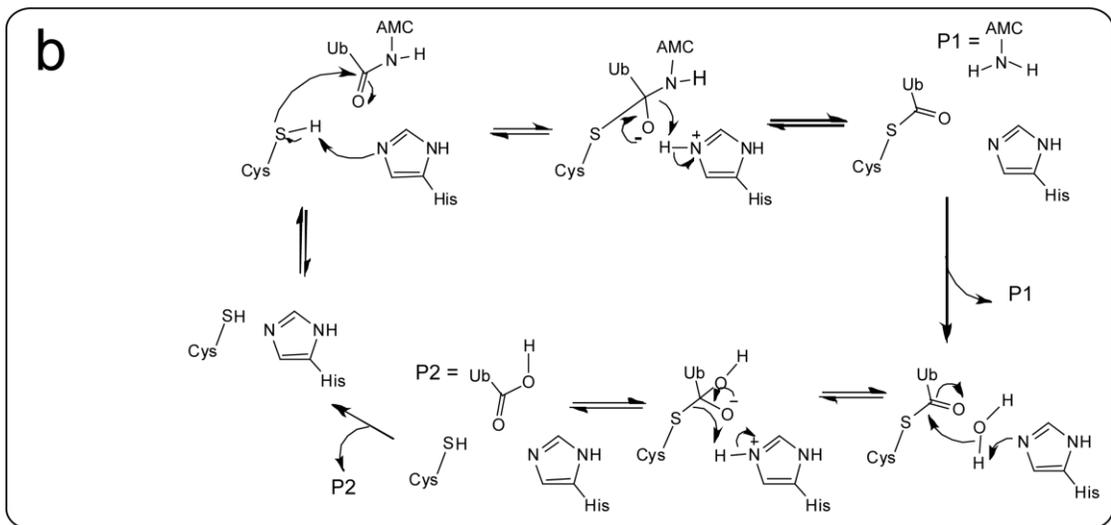
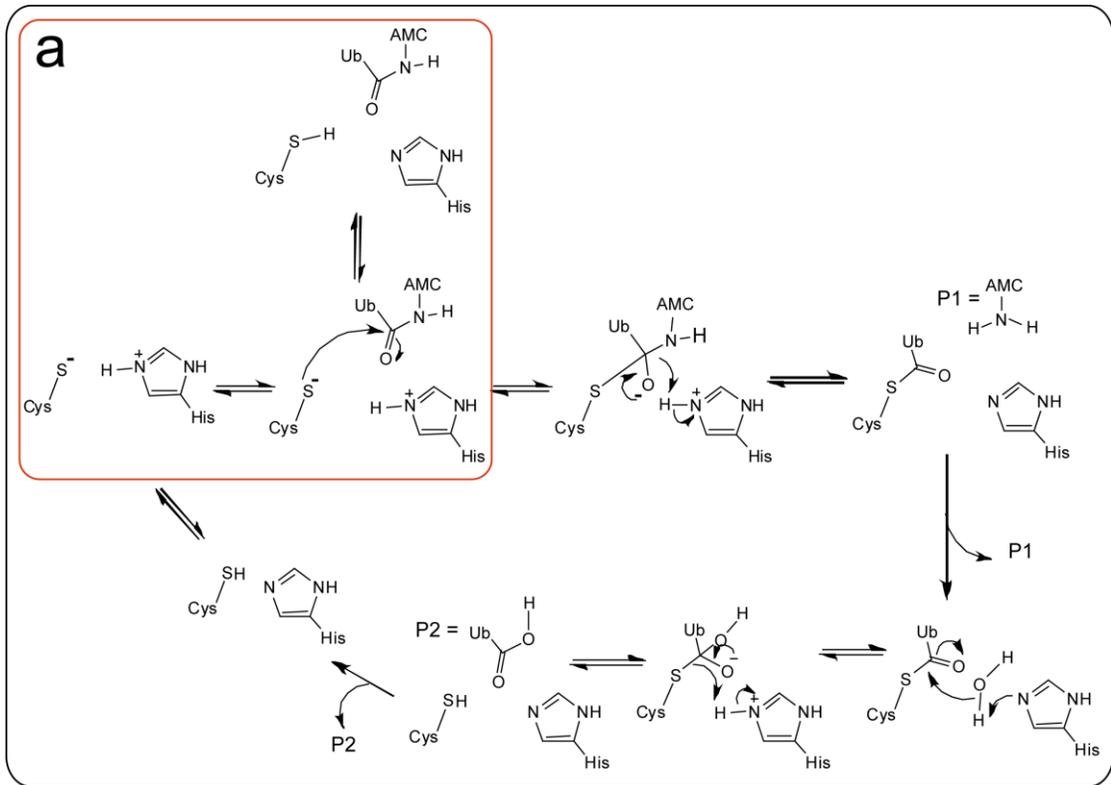
Case *et al.* reported detailed mechanistic studies of UCH-L1.<sup>34</sup> Solvent isotope effect studies revealed that the active site cysteine in the free UCH-L1 exists as the thiol. After bound with ubiquitin, the diad in the resultant Michaelis complex tautomerized to ion pair Cys-S<sup>-</sup>/His-ImH<sup>+</sup>. This provides further evidence to support the low reactivity to H<sub>2</sub>O<sub>2</sub> for unbound UCH-L1. No mechanistic study has specifically reported the ionization state of UCH-L3 so far.

### 3.3.3 Modification of USP1/UAF1 by iodoacetamide

USP1, like other cysteine proteases, contains a catalytic triad. Two mechanisms, i.e. ion-pair and general base catalysis, have been proposed for cysteine proteases. Cysteine proteases, represented by papain, form a thioate-imidazolium ion pair that is required for catalysis.<sup>35</sup> In this mechanism, Asp helps to orientate the His to become a proton acceptor and the Cys is the proton donor. The Cys/His residues exist mainly as Cys-S<sup>-</sup>/His-ImH<sup>+</sup> ion pair (Figure 3.21a). The thiolate of active site Cys acts as a nucleophilic to attack the substrate's carbonyl group, which leads to a tetrahedral intermediate and subsequently the acyl-enzyme intermediate. In the deacylation step, a lytic water molecule serves as a nucleophile to attack the carbonyl group on the acyl-enzyme intermediate. This step results in the formation of the second tetrahedral intermediate, which then collapses to produce the product and also regenerates the free enzyme.<sup>36</sup>

General base catalysis does not involve Cys-S<sup>-</sup>/His-ImH<sup>+</sup> ion-pair. Instead, the triad His residue acts as a general base to activate the nucleophilic cysteine

sulfhydryl group (Figure 3.21b). The rest of the steps are the same as in thioate-imidazolium ion pair mechanism as described above. This type of mechanism is commonly seen in serine proteases.<sup>37</sup>



**Figure 3.21: Two proposed mechanisms for cysteine proteases using Ub-AMC as the substrate. (a) Thiolate-imidazolium ion pair mechanism. The boxed area indicates the uncharged pair of Cys-SH/His-Im and the ion pair of Cys-S<sup>-</sup>/His-ImH<sup>+</sup>. (b) General base mechanism. Figure b was generated by Mark Villamil. Figure a was modified by Junjun Chen based on Figure b.**

The reactivity of the essential thiol group of USP1/UAF1 has been explored by measuring the rate constant of alkylation by iodoacetamide. Only the deprotonated form of the thiol group reacts with iodoacetamide with an appreciable rate. The pH dependence of the alkylation reaction provided useful information about the chemical reactivity of the thiol group. It is apparent from Figure 3.9 that the pH dependence curve has a sigmoid shape. The  $pK_a$  value obtained from this curve is 8.6, which is tentatively assigned to the active site cysteine 90. It is likely that USP1/UAF1 follows a general base mechanism instead of ion pair mechanism based on the high  $pK_a$  value obtained here.

In the following section, we will discuss several viral proteases that were proposed to use general base catalysis. Poliovirus proteinase 3C (PV3C) belongs to picornaviruses family proteases,<sup>38</sup> which are structurally related to the trypsin family of serine peptidases.<sup>38</sup> Sarkany *et al* reported that PV3C follows general base catalysis as found with serine peptidases.<sup>39</sup> In their study, a  $pK_a$  value of 9.11 for Cys was observed when PV3C was treated with iodoacetamide while a  $pK_a$  value of 8.55 was observed under the treatment with iodoacetate. A  $pK_a$  value of 8.52 was obtained by measuring the pH dependence of  $k_{cat}/K_m$ . Also, they utilized a PV3C mutant which contains a single cysteine to measure the absorbance change upon alkylation with iodoacetate at 250 nm, where the nondissociated thiol group has a negligible absorbance relative to the ionized form. Alkylation of thiolate abolishes the

absorbance at 250 nm. The pH dependence of the absorbance at 250 nm suggested that PV3C contains an active site thiol group with an unperturbed  $pK_a$ , which excludes the thiolate-imidazolium ion-pair mechanism.

Another example of cysteine proteases proceeding through a general base mechanism is SARS 3C-like proteinase (SARS 3CL), a key enzyme for drug design against severe acute respiratory syndrome (SARS).<sup>40</sup> Huang *et al* reported the  $pK_a$  for the active site Cys is 8.34 in SARS 3CL as determined from the pH dependency of  $k_{cat}/K_m$ .<sup>40</sup> Also, a normal isotope effect around 1.3 was determined for  $k_{cat}/K_m$ . This indicates the catalysis proceeds slower in heavy water. In the case of general base catalysis, the reaction is expected to proceed in buffered H<sub>2</sub>O 2-3 times faster than in buffered D<sub>2</sub>O.

Papain, the most thoroughly investigated member of the cysteine proteinase superfamily, shows a  $pK_a$  value for the active site cysteine of 4.0 by measuring pH-dependence of inactivation of papain by chloroacetamide.<sup>41</sup> Almost identical second order rate constants for alkylation of papain with chloroacetamide were observed in buffered H<sub>2</sub>O and buffered D<sub>2</sub>O. This indicates no isotope effect on the alkylation of Papain active site cysteine.<sup>41</sup> Also, the absorbance change upon alkylation of active site cysteine with iodoacetamide clearly reveals an ionized cysteine with a  $pK_a$  value 3.7.<sup>35</sup> Based on the above evidence, it was widely accepted that the papain catalysis follows a thiolate-imidazolium ion pair mechanism.

Further support for the notion that USP1/UAF1 follows a general base mechanism in catalysis was obtained from SKIE determination. We observed a normal SKIE of 1.5 for inactivation of USP1/UAF1 by iodoacetamide. This is again consistent with the general base-catalyzed processes, in which a normal SKIE is expected. A

normal SKIE of 2.9 was also observed for  $k_{cat}/K_m$ , which is close to that determined in the cysteine alkylation experiment (Villamil *et al*, unpublished results).

In view of the catalytic mechanism, USP1/UAF1 is more closely related to the viral proteases that utilize a general base mechanism, rather than the papain type of proteases that utilize ion-pair mechanism. Notably, the SARS-CoV PLpro has been shown to be also a deubiquitinating enzyme.<sup>42, 43</sup> PLpro may have critical roles not only in proteolytic processing of the viral replicase complex but also in subverting cellular ubiquitination machinery to facilitate viral replication.

To support the active site cysteine  $pK_a$  assignment, the  $pK_a$  of active site cysteine can also be determined by measuring the pH-dependent difference spectrum of the wild-type and active site cysteine mutant (C90S USP1 and C90S USP1/UAF1) at 250 nm as discussed for PV3C and papain above (experiments have not been done yet).<sup>44, 45</sup> At this wavelength (250 nm), the absorbance of the thiol form is negligible. The molar extinction coefficient of the thiolate ion can be calculated from the absorbance difference measured before and after alkylation of the thiol group. The formation of a covalent bond during alkylation abolishes the absorbance of the ionized form. If USP1/UAF1 follows a general base mechanism as we proposed, no deviation of the thiolate ion should be observed from the theoretical curve at low pH, where the thiolate-imidazolium ion-pair was expected to occur.

### 3.3.4 Phosphorylation of USP1

To explore whether post-translational modification of either USP1 or UAF1 affects the complex formation and the activity of USP1, we investigated the phosphorylation state of USP1, UAF1 and USP1/UAF1 complex by using a phosphoprotein-specific dye. In large-scale proteomic analyses, USP1 was reported to

be phosphorylated at several serine residues (Ser13, Ser42, Ser67, Ser313 and Ser 475),<sup>46-48</sup> in which Ser42 is involved in response to DNA damage.<sup>47</sup> Here we showed that the USP1 was phosphorylated when expressed either alone or as a complex with UAF1, while UAF1 was not phosphorylated. Treating USP1 and USP1/UAF1 with calf intestinal phosphatase resulted in decrease in the intensity of the phosphorylated USP1 band. This provides important evidence for future phosphorylation studies. We can identify the phosphorylation sites on USP1 by mass spectrometry. With that information, we will generate USP1 mutants by mutating the residue(s) that undergoes phosphorylation and assess the mutant's ability in forming complex with UAF1.

### **3.3.5 *In vitro* deubiquitination of native monoubiquitinated human PCNA-Ub**

I am the first to show USP1/UAF1 indeed deubiquitinates native human PCNA-Ub *in vitro*. Since we have developed the chemical ubiquitination method described in chapter 2, the same method can be applied onto human protein to get human ubiquitinated PCNA. Because the ubiquitin moiety in the chemically ubiquitinated PCNA contains an intact ubiquitin C-terminus and a normal C-terminal isopeptide bond, the small difference in the linkage length is likely tolerated by the USP1/UAF1 complex. First we can test whether the chemically ubiquitinated PCNA can act as a substrate of USP1/UAF1 or not. If the chemically ubiquitinated PCNA can be cleaved by USP1/UAF1, the steady-state kinetic constants can be obtained. A comparison between enzymatically and chemically monoubiquitinated hPCNA will provide important information on the structural determinant in the isopeptide linkage for the recognition of ubiquitinated protein substrate.

### 3.4 Materials and Methods

#### 3.4.1 Determine the steady state kinetic parameters of USP1 and USP1/UAF1

USP1 and USP1/UAF1 activity assays were run with Ub-AMC as substrates. The reaction mixture was 50 mM HEPES buffer, pH 7.8, containing 0.1 mg/ml BSA, 0.5 mM EDTA, and Ub-AMC (100 – 3000 nM). Reactions were initiated by the addition of 20 nM USP1 or 2 nM USP1/UAF1. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence at an excitation wavelength of 350 nm and an emission wavelength of 435 nm.

#### 3.4.2 Time- and concentration-dependent inhibition of USP1/UAF1 and USP1 by H<sub>2</sub>O<sub>2</sub>

USP1/UAF1 or USP1 (130 nM) was incubated with various concentrations of H<sub>2</sub>O<sub>2</sub> (5 – 40 μM for USP1/UAF1, 100 – 300 μM for USP1) in a buffer containing 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 0.4 mg/ml BSA. Before each experiment, H<sub>2</sub>O<sub>2</sub> was freshly diluted from a concentrated stock (8.8 M, 30%) and quantified using  $\epsilon_{240} = 43.6 \text{ M}^{-1} \text{ cm}^{-1}$ . At various time points between 0 and 21 min, aliquots (2 μl for USP1/UAF1, 20 μl for USP1) of the pre-incubation mix were diluted into a reaction mixture containing 100 nM Ub-AMC (Enzo Life Sciences International, Inc., Plymouth Meeting, PA) in 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 1 mM DTT and 0.1 mg/ml BSA. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence at an excitation wavelength of 350 nm and an emission wavelength of 435 nm. At each H<sub>2</sub>O<sub>2</sub> concentration, remaining USP1/UAF1 or USP1 activity was fitted to equation 1.

$$\text{Remaining activity (\%)} = 100 \times \exp(-k_{\text{obs}} \times t) + C \quad (\text{Eq. 1})^{33}$$

The observed pseudo first-order inactivation rates,  $k_{\text{obs}}$ , were then plotted against the concentration of  $\text{H}_2\text{O}_2$ , and the second-order rate constant was obtained from the slope of the resulting linear plot.

### **3.4.3 Reversibility of inactivation by $\text{H}_2\text{O}_2$ of USP1/UAF1 or USP1**

#### **3.4.3.1 Inactivation of USP1/UAF1 or USP1 by $\text{H}_2\text{O}_2$**

USP1/UAF1 or USP1 (130 nM) was incubated with varying concentrations of freshly prepared  $\text{H}_2\text{O}_2$  (2  $\mu\text{M}$  – 1 mM) in a buffer containing 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 0.4 mg/ml BSA at room temperature for 5 min. In the parallel experiments, 1, 3-cyclohexanedione (dimedone) (Fluka analytical, St. Louis, MO) (6.5 mM) was added to the inactivated USP1/UAF1 or USP1 and incubated for 30 min at room temperature. Aliquots (2  $\mu\text{l}$  for USP1/UAF1, 20  $\mu\text{l}$  for USP1) of the dimedone-treated USP1/UAF1 or USP1 mix were diluted into a reaction mixture containing 100 nM Ub-AMC in 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 1 mM DTT or glutathione and 0.1 mg/ml BSA. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence at an excitation wavelength of 350 nm and an emission wavelength of 435 nm.

#### **3.4.3.2 Reactivation by DTT or glutathione**

USP1/UAF1 or USP1 (130 nM) was incubated with varying concentrations of freshly prepared  $\text{H}_2\text{O}_2$  (2  $\mu\text{M}$  – 1 mM) in a buffer containing 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 0.4 mg/ml BSA at room temperature for 5 min. Bovin liver catalase (300 units) was added to the above deactivated USP1/UAF1 or USP1 to degrade all remaining  $\text{H}_2\text{O}_2$  and incubated for 10 min at room temperature. Aliquots (2  $\mu\text{l}$  for USP1/UAF1, 20  $\mu\text{l}$  for USP1) of the reactivated USP1/UAF1 or USP1 mix were diluted into a reaction mixture containing 100 nM Ub-AMC (Biomol)

in 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 1 mM DTT or glutathione and 0.1 mg/ml BSA. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence at an excitation wavelength of 350 nm and an emission wavelength of 435 nm.

#### **3.4.4 Reversibility of inactivation by H<sub>2</sub>O<sub>2</sub> of USP7, UCH-L3 and UCH-L1 in DTT**

##### 3.4.4.1 Inactivation of DUBs by H<sub>2</sub>O<sub>2</sub>

USP7 (6.5 nM), UCH-L3 (325 pM) and UCH-L1 (32.5 nM) were incubated with varying concentrations of freshly prepared H<sub>2</sub>O<sub>2</sub> (0.05 – 1 mM) in a buffer containing 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 0.4 mg/ml BSA at room temperature for 5 min. In the parallel experiments, 1, 3-cyclohexanedione (6.5 mM) was added to the inactivated DUBs and incubated for 30 min at room temperature. Aliquots of the dimedone-treated reaction mix were diluted into a reaction mixture containing 100 nM Ub-AMC in 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 1 mM DTT and 0.1 mg/ml BSA. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence at an excitation wavelength of 350 nm and an emission wavelength of 435 nm.

##### 3.4.4.2 Reactivation of treated DUBs by DTT

USP7 (6.5 nM), UCH-L3 (325 pM) and UCH-L1 (32.5 nM) were incubated with varying concentrations of freshly prepared H<sub>2</sub>O<sub>2</sub> (0.05 – 1 mM) in a buffer containing 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 0.4 mg/ml BSA at room temperature for 5 min. Bovin liver catalase (300 units) was added to the above deactivated DUBs to degrade all remaining H<sub>2</sub>O<sub>2</sub> and incubated for 10 min at room temperature. Aliquots of the reactivated DUBs were diluted into a reaction mixture

containing 100 nM Ub-AMC (Biomol) in 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 1 mM DTT and 0.1 mg/ml BSA. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence at an excitation wavelength of 350 nm and an emission wavelength of 435 nm.

#### **3.4.5 Inactivation of USP1/UAF1 and USP1 by iodoacetamide under different pHs**

USP1/UAF1 or USP1 (200 nM) was incubated with various 4 mM iodoacetamide in a buffer containing 5 mM NaOAc, 5 mM MES, 5 mM glycine, 15 mM Tris, pH varying from 5.6 – 8.1, 0.4 mg/ml BSA for 2 min. Aliquots (2  $\mu$ l for USP1/UAF1, 20  $\mu$ l for USP1) of the pre-incubation mix were diluted into a reaction mixture containing 100 nM Ub-AMC (Enzo Life Sciences International, Inc., Plymouth Meeting, PA) in 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 1 mM DTT and 0.1 mg/ml BSA. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence at an excitation wavelength of 350 nm and an emission wavelength of 435 nm. At each pH, remaining USP1/UAF1 or USP1 activity was plotted against their remaining activity.

#### **3.4.6 pH dependence of the inactivation rate of USP1/UAF1 by iodoacetamide**

USP1/UAF1 (208 nM) was incubated at 25 °C with various concentrations of iodoacetamide (0.5 – 8 mM) in a buffer containing 5 mM NaOAc, 5 mM MES, 5 mM glycine, 15 mM Tris, pH varying from 5.6 – 10.6, 0.4 mg/ml BSA. At various time points between 0 and 20 min, aliquots (13  $\mu$ l) of the pre-incubation mix (16  $\mu$ l) were diluted into a reaction mixture containing 200 nM Ub-AMC in 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 1 mM DTT and 0.1 mg/ml BSA. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence at an excitation

wavelength of 350 nm and an emission wavelength of 435 nm. For each incubation pH, at each H<sub>2</sub>O<sub>2</sub> concentration, remaining USP1/UAF1 activity was fitted to equation 1.

The observed pseudo first-order inactivation rates,  $k_{obs}$ , were then plotted against the concentration of iodoacetamide, and the second-order rate constant was obtained from the slope of the resulting linear plot. The pH profile data [ $\log(k_2)$  vs pH] were fitted to equation 2 by using GraphPad Prism 5 (GraphPad Software, Inc. La Jolla, CA).

$$\log Y = \log[Y_L + Y_H(K / [H]) / (1 + K / [H])] \quad (\text{Eq. 2})^{49}$$

where  $K$  is the acid dissociation constant,  $Y$  is the  $k_2$  at the hydrogen ion concentration  $[H]$ , and  $Y_L$  and  $Y_H$  are the minimum and maximum values of  $k_2$ , respectively.

Kinetic isotope effect experiments were carried out exactly the same as above except incubation buffer contains 98% deuterium oxide. The p<sup>2</sup>H of deuterium oxide solution can be obtained from pH meter readings following the equation p<sup>2</sup>H = pH(meter reading) + 0.4.<sup>50</sup>

#### **3.4.7 Modifications of USP1/UAF1 by DAz1-Biotin probe**

75  $\mu$ l USP1/UAF1 (2  $\mu$ M) was pre-incubated with 1 mM TCEP (tris(2-carboxyethyl)phosphine) on ice for 30 min. TCEP-treated USP1/UAF1 was loaded onto a Micro Bio-Spin 6 Column (P-6 column) (Biorad, Hercules, CA) twice. 22  $\mu$ l reduced USP1/UAF1 was taken out and incubated with 0.5 mM DAz1-Biotin (a generous gift from Dr. Kate S. Carroll, University of Michigan, Ann Arbor, MI) and 20  $\mu$ M H<sub>2</sub>O<sub>2</sub> in phosphate-buffered saline (PBS) for 30 min at RT with a gentle rocking. Biotinylated proteins were separated by 8% SDS-PAGE and transferred to a

Hybond ECL nitrocellulose membrane (GE healthcare). After transfer, the nitrocellulose membrane was blocked with 5% milk in phosphate-buffered saline Tween-20 (PBST) for 1 h at RT. The membrane was washed with PBST (1 × 15 min, 2 × 10 min) and then incubated with 1 : 50,000 HRP-streptavidin for 1 h at RT. Nitrocellulose membrane was washed with PBST (1 × 15 min, 2 × 10 min) and then developed with ECL western blotting substrate (Fisher Scientific).

#### **3.4.8 Detect phosphorylation of USP1/UAF1**

50 µl of each protein solution (2.4 µg USP1, 5.2 µg UAF1, 83 µg USP1/UAF1) was treated without or with 10 units (1 µl) of alkaline phosphatase (New England Biolabs, Ipswich, MA) under 37 °C for 30 mins. For each sample, 200 µl methanol was added and mixed by vortexing. Next, 50 µl chloroform was added and mixed by vortexing. Last, 150 µl ddH<sub>2</sub>O was added and mixed by vortexing. Centrifuge at 12,000 rpm (Eppendorf centrifuge 5424) for 5 mins. Supernatant was removed and the white precipitate disc that forms between the upper and lower phases was kept. 150 µl methanol was added to the white precipitate and mixed by vortexing. Centrifuge at 12,000 rpm (Eppendorf centrifuge 5424) for 5 mins. Supernatant was discarded and the pellet was kept. Let the pellet air dry and store in -20 °C.

Take out the dried pellet just before use and add 20 µl ddH<sub>2</sub>O to get it resuspended by vortexing. 4 µl 6x Laemmli dye was added to each sample and separated by an 8% SDS-PAGE. Stain the acrylamide gel by Pro-Q® Diamond phosphoprotein gel stain (Invitrogen Corporation, Carlsbad, CA) by a standard protocol provided by the manufacturer. Stained gel was visualized using excitation at 532 – 560 nm by AlphaImager HP located in center for translational cancer research core facility in University of Delaware.

### **3.4.9 *In vitro* deubiquitination assay of monoubiquitinated human PCNA by USP1/UAF1**

20 nM or 50 nM USP1/UAF1 was incubated with 520 nM hPCNA-Ub (assay volume is 18  $\mu$ l) in a buffer containing 50 mM HEPES, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA and 1 mM DTT at 37 °C. At 10 min, 30 min and 60 mins time points, 5  $\mu$ l aliquot was taken out and quenched by the addition of Laemmli sample buffer. The reaction product was separated on a 15% denaturing SDS-PAGE gel. After SDS-PAGE, the proteins were transferred to a Hybond ECL nitrocellulose membrane (GE healthcare). After transfer, the nitrocellulose membrane was blocked with 5% milk in phosphate-buffered saline Tween-20 (PBST) for 1 h at RT. The membrane was washed with PBST (1  $\times$  15 min, 2  $\times$  10 min) and then incubated with 1 : 10,000 MAbs (primary antibody for human PCNA) (Santa Cruz Biotechnology, Inc., Santa Cruz, CA) for 1 h at RT. The membrane was washed with PBST (1  $\times$  15 min, 2  $\times$  10 min) and then incubated with 1 : 80,000 anti-mouse antibody (Sigma-Aldrich, St. Louis, MO) and 1 : 50,000 peroxidase-conjugated streptavidin (Fisher Scientific) for 1 h at RT. Nitrocellulose membrane was washed with PBST (1  $\times$  15 min, 2  $\times$  10 min) and then developed with ECL western blotting substrate (Fisher Scientific).

### **3.4.10 Sample preparation for in-gel digestion**

30  $\mu$ l USP1 or USP1/UAF1 (2  $\mu$ M) was incubated with 2 mM of freshly prepared H<sub>2</sub>O<sub>2</sub> in a buffer containing 50 mM HEPES/NaOH, pH 7.8, 0.5 mM EDTA, 0.4 mg/ml BSA at room temperature for 5 min. Dimedone (1, 3-cyclohexanedione) (final concentration is 6.5 mM) was added to the inactivated DUBs and incubated for 30 min at room temperature. Or 30  $\mu$ l USP1 or USP1/UAF1 (2  $\mu$ M) was incubated with 2 mM iodoacetanilide at room temperature for 5 min. 6x laemmli dye was added to each sample and separated by an 8% SDS-PAGE.

## REFERENCES

1. Nijman, S. M.; Luna-Vargas, M. P.; Velds, A.; Brummelkamp, T. R.; Dirac, A. M.; Sixma, T. K.; Bernards, R., A genomic and functional inventory of deubiquitinating enzymes. *Cell* **2005**, *123* (5), 773-86.
2. Leroy, E.; Boyer, R.; Auburger, G.; Leube, B.; Ulm, G.; Mezey, E.; Harta, G.; Brownstein, M. J.; Jonnalagada, S.; Chernova, T.; Dehejia, A.; Lavedan, C.; Gasser, T.; Steinbach, P. J.; Wilkinson, K. D.; Polymeropoulos, M. H., The ubiquitin pathway in Parkinson's disease. *Nature* **1998**, *395* (6701), 451-2.
3. Liu, Y.; Fallon, L.; Lashuel, H. A.; Liu, Z.; Lansbury, P. T., Jr., The UCH-L1 gene encodes two opposing enzymatic activities that affect alpha-synuclein degradation and Parkinson's disease susceptibility. *Cell* **2002**, *111* (2), 209-18.
4. Burnett, B.; Li, F.; Pittman, R. N., The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity. *Hum Mol Genet* **2003**, *12* (23), 3195-205.
5. Balakirev, M. Y.; Tcherniuk, S. O.; Jaquinod, M.; Chroboczek, J., Otubains: a new family of cysteine proteases in the ubiquitin pathway. *EMBO Rep* **2003**, *4* (5), 517-22.
6. Soares, L.; Seroogy, C.; Skrenta, H.; Anandasabapathy, N.; Lovelace, P.; Chung, C. D.; Engleman, E.; Fathman, C. G., Two isoforms of otubain 1 regulate T cell anergy via GRAIL. *Nat Immunol* **2004**, *5* (1), 45-54.
7. Huang, T. T.; Nijman, S. M.; Mirchandani, K. D.; Galardy, P. J.; Cohn, M. A.; Haas, W.; Gygi, S. P.; Ploegh, H. L.; Bernards, R.; D'Andrea, A. D., Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat Cell Biol* **2006**, *8* (4), 339-47.
8. Zhuang, Z.; Johnson, R. E.; Haracska, L.; Prakash, L.; Prakash, S.; Benkovic, S. J., Regulation of polymerase exchange between Pol $\epsilon$  and Pol $\delta$  by monoubiquitination of PCNA and the movement of DNA polymerase holoenzyme. *Proc Natl Acad Sci U S A* **2008**, *105* (14), 5361-6.

9. Hoege, C.; Pfander, B.; Moldovan, G. L.; Pyrowolakis, G.; Jentsch, S., RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **2002**, *419* (6903), 135-41.
10. Stelter, P.; Ulrich, H. D., Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* **2003**, *425* (6954), 188-91.
11. Nijman, S. M.; Huang, T. T.; Dirac, A. M.; Brummelkamp, T. R.; Kerkhoven, R. M.; D'Andrea, A. D.; Bernards, R., The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. *Molecular cell* **2005**, *17* (3), 331-9.
12. Oestergaard, V. H.; Langevin, F.; Kuiken, H. J.; Pace, P.; Niedzwiedz, W.; Simpson, L. J.; Ohzeki, M.; Takata, M.; Sale, J. E.; Patel, K. J., Deubiquitination of FANCD2 is required for DNA crosslink repair. *Molecular cell* **2007**, *28* (5), 798-809.
13. Alpi, A. F.; Patel, K. J., Monoubiquitylation in the Fanconi anemia DNA damage response pathway. *DNA Repair (Amst)* **2009**, *8* (4), 430-5.
14. Kook, H., Fanconi anemia: current management. *Hematology* **2005**, *10 Suppl 1*, 108-10.
15. Garcia-Higuera, I.; Taniguchi, T.; Ganesan, S.; Meyn, M. S.; Timmers, C.; Hejna, J.; Grompe, M.; D'Andrea, A. D., Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Molecular cell* **2001**, *7* (2), 249-62.
16. Montes de Oca, R.; Andreassen, P. R.; Margossian, S. P.; Gregory, R. C.; Taniguchi, T.; Wang, X.; Houghtaling, S.; Grompe, M.; D'Andrea, A. D., Regulated interaction of the Fanconi anemia protein, FANCD2, with chromatin. *Blood* **2005**, *105* (3), 1003-9.
17. Taniguchi, T.; Garcia-Higuera, I.; Andreassen, P. R.; Gregory, R. C.; Grompe, M.; D'Andrea, A. D., S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. *Blood* **2002**, *100* (7), 2414-20.
18. Wang, X.; Andreassen, P. R.; D'Andrea, A. D., Functional interaction of monoubiquitinated FANCD2 and BRCA2/FANCD1 in chromatin. *Mol Cell Biol* **2004**, *24* (13), 5850-62.

19. Hussain, S.; Wilson, J. B.; Medhurst, A. L.; Hejna, J.; Witt, E.; Ananth, S.; Davies, A.; Masson, J. Y.; Moses, R.; West, S. C.; de Winter, J. P.; Ashworth, A.; Jones, N. J.; Mathew, C. G., Direct interaction of FANCD2 with BRCA2 in DNA damage response pathways. *Hum Mol Genet* **2004**, *13* (12), 1241-8.
20. Niedzwiedz, W.; Mosedale, G.; Johnson, M.; Ong, C. Y.; Pace, P.; Patel, K. J., The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Molecular cell* **2004**, *15* (4), 607-20.
21. Ishiai, M.; Kitao, H.; Smogorzewska, A.; Tomida, J.; Kinomura, A.; Uchida, E.; Saberi, A.; Kinoshita, E.; Kinoshita-Kikuta, E.; Koike, T.; Tashiro, S.; Elledge, S. J.; Takata, M., FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nat Struct Mol Biol* **2008**, *15* (11), 1138-46.
22. Geng, L.; Huntoon, C. J.; Karnitz, L. M., RAD18-mediated ubiquitination of PCNA activates the Fanconi anemia DNA repair network. *J Cell Biol* **2010**, *191* (2), 249-57.
23. Cohn, M. A.; Kowal, P.; Yang, K.; Haas, W.; Huang, T. T.; Gygi, S. P.; D'Andrea, A. D., A UAF1-containing multisubunit protein complex regulates the Fanconi anemia pathway. *Molecular cell* **2007**, *28* (5), 786-97.
24. Samara, N. L.; Datta, A. B.; Berndsen, C. E.; Zhang, X.; Yao, T.; Cohen, R. E.; Wolberger, C., Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science* **2010**, *328* (5981), 1025-9.
25. Tonks, N. K., Redox redux: revisiting PTPs and the control of cell signaling. *Cell* **2005**, *121* (5), 667-70.
26. Pagano, G.; Degan, P.; d'Ischia, M.; Kelly, F. J.; Nobili, B.; Pallardo, F. V.; Youssoufian, H.; Zatterale, A., Oxidative stress as a multiple effector in Fanconi anaemia clinical phenotype. *Eur J Haematol* **2005**, *75* (2), 93-100.
27. Sohn, J.; Rudolph, J., Catalytic and chemical competence of regulation of cdc25 phosphatase by oxidation/reduction. *Biochemistry* **2003**, *42* (34), 10060-70.
28. Reddie, K. G.; Seo, Y. H.; Muse Iii, W. B.; Leonard, S. E.; Carroll, K. S., A chemical approach for detecting sulfenic acid-modified proteins in living cells. *Mol Biosyst* **2008**, *4* (6), 521-31.

29. Zhuang, Z.; Ai, Y., Processivity factor of DNA polymerase and its expanding role in normal and translesion DNA synthesis. *Biochim Biophys Acta* **2009**, *1804* (5), 1081-93.
30. Cohn, M. A.; Kee, Y.; Haas, W.; Gygi, S. P.; D'Andrea, A. D., UAF1 is a subunit of multiple deubiquitinating enzyme complexes. *The Journal of biological chemistry* **2009**, *284* (8), 5343-51.
31. Hampton, M. B.; Stamenkovic, I.; Winterbourn, C. C., Interaction with substrate sensitises caspase-3 to inactivation by hydrogen peroxide. *FEBS Lett* **2002**, *517* (1-3), 229-32.
32. Barrett, W. C.; DeGnore, J. P.; Konig, S.; Fales, H. M.; Keng, Y. F.; Zhang, Z. Y.; Yim, M. B.; Chock, P. B., Regulation of PTP1B via glutathionylation of the active site cysteine 215. *Biochemistry* **1999**, *38* (20), 6699-705.
33. Hong, L.; Fast, W., Inhibition of human dimethylarginine dimethylaminohydrolase-1 by S-nitroso-L-homocysteine and hydrogen peroxide. Analysis, quantification, and implications for hyperhomocysteinemia. *The Journal of biological chemistry* **2007**, *282* (48), 34684-92.
34. Case, A.; Stein, R. L., Mechanistic studies of ubiquitin C-terminal hydrolase L1. *Biochemistry* **2006**, *45* (7), 2443-52.
35. Polgar, L., Mercaptide-imidazolium ion-pair: the reactive nucleophile in papain catalysis. *FEBS Lett* **1974**, *47* (1), 15-8.
36. Brocklehurst, K.; Watts, A. B.; Patel, M.; Verma, C.; Thomas, E. W.; in Sinnott, M., *Comprehensive biological catalysis : a mechanistic reference*. Academic Press: San Diego, CA. pp. 381-423, 1998.
37. Polgar, L.; Bender, M. L., The nature of general base-general acid catalysis in serine proteases. *Proc Natl Acad Sci U S A* **1969**, *64* (4), 1335-42.
38. Bergmann, E. M.; James, M. N. G., in *Proteases of infectious agents* (Dunn, B. M., Ed.) Academic Press: London, pp 139-163, 1999.
39. Sarkany, Z.; Szeltner, Z.; Polgar, L., Thiolate-imidazolium ion pair is not an obligatory catalytic entity of cysteine peptidases: the active site of picornain 3C. *Biochemistry* **2001**, *40* (35), 10601-6.

40. Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L., 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry* **2004**, *43* (15), 4568-74.
41. Polgar, L., On the mode of activation of the catalytically essential sulfhydryl group of papain. *Eur J Biochem* **1973**, *33* (1), 104-9.
42. Barretto, N.; Jukneliene, D.; Ratia, K.; Chen, Z.; Mesecar, A. D.; Baker, S. C., The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. *J Virol* **2005**, *79* (24), 15189-98.
43. Lindner, H. A.; Fotouhi-Ardakani, N.; Lytvyn, V.; Lachance, P.; Sulea, T.; Menard, R., The papain-like protease from the severe acute respiratory syndrome coronavirus is a deubiquitinating enzyme. *J Virol* **2005**, *79* (24), 15199-208.
44. Rozema, D. B.; Poulter, C. D., Yeast protein farnesyltransferase. pKas of peptide substrates bound as zinc thiolates. *Biochemistry* **1999**, *38* (40), 13138-46.
45. Stone, E. M.; Costello, A. L.; Tierney, D. L.; Fast, W., Substrate-assisted cysteine deprotonation in the mechanism of dimethylargininase (DDAH) from *Pseudomonas aeruginosa*. *Biochemistry* **2006**, *45* (17), 5618-30.
46. Gauci, S.; Helbig, A. O.; Slijper, M.; Krijgsveld, J.; Heck, A. J.; Mohammed, S., Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. *Anal Chem* **2009**, *81* (11), 4493-501.
47. Matsuoka, S.; Ballif, B. A.; Smogorzewska, A.; McDonald, E. R., 3rd; Hurov, K. E.; Luo, J.; Bakalarski, C. E.; Zhao, Z.; Solimini, N.; Lerenthal, Y.; Shiloh, Y.; Gygi, S. P.; Elledge, S. J., ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **2007**, *316* (5828), 1160-6.
48. Dephoure, N.; Zhou, C.; Villen, J.; Beausoleil, S. A.; Bakalarski, C. E.; Elledge, S. J.; Gygi, S. P., A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci U S A* **2008**, *105* (31), 10762-7.
49. Cleland, W. W., Statistical analysis of enzyme kinetic data. *Methods Enzymol* **1979**, *63*, 103-38.
50. Glasoe, P. K.; Long, F. A., Use of Glass Electrodes to Measure Acidities in Deuterium Oxide. *Journal of Physical Chemistry* **1960**, *64* (1), 188-190.

## Chapter 4

### DISCOVERY OF SMALL-MOLECULE INHIBITORS AGAINST THE HUMAN USP1/UAF1 DEUBIQUITINATING ENZYME COMPLEX

#### 4.1 Introduction

Abnormal function of the ubiquitin-proteasome system has been linked to many human diseases, including neurological disorders, viral infection and cancer.<sup>1-3</sup> Ubiquitin is a versatile signaling molecule. Although initially being discovered in proteasome-mediated protein degradation, ubiquitin and ubiquitin-like proteins are now known to play important roles in a myriad of nonproteolytic processes, including membrane trafficking, transcription regulation and DNA damage response.<sup>4</sup>

In recent years, ubiquitin-proteasome system has attracted increasing attention as promising target for novel therapeutics. A series of proteasome inhibitors has been developed and tested in both cell and animal models.<sup>5,6</sup> In 2003, the protease inhibitor Bortezomib (originally called PS341 and then Velcade) was the first drug that targets a component of the ubiquitin system to be approved for clinical use in the United States.<sup>7</sup> It was approved by FDA for the treatment of multiple myeloma and mantle cell lymphoma. The success of Velcade has established the ubiquitin-proteasome system as a valid target for anti-cancer treatment. However, proteasome inhibitors in general suffer from a narrow therapeutic index because proteasome is at a convergence point downstream of many important cellular pathways.<sup>8</sup> A promising alternative to proteasome inhibitors is to target the enzymes upstream of proteasome-

mediated protein degradation, *i.e.* the ubiquitin conjugation and deconjugation system, to generate more specific, less toxic therapeutic agents.

Developing inhibitors for ubiquitin ligase E3 has attracted lots of attention due to its large number of existing E3s (more than 600) and its high specificity in ubiquitination. Several E3 ligase inhibitors have entered clinical trials. For example, Roche and Johnson & Johnson have found E3-Hdm2 inhibitors called Nutlin/R7112, JNJ26854165 respectively, for the treatment of blood cancers and solid tumors. Both compounds were used to increase p53 expression because they can interrupt the interaction between E3 ligase HDM2 and substrate p53, a tumor suppressor. Novartis developed LCL161 to inhibit E3-IAP (Apoptosis Proteins) to treat solid tumors. Seven antagonists of IAPs have entered clinical trials.<sup>9</sup>

Another ubiquitin-like protein, Nedd8, can also be covalently conjugated to target proteins in process called neddylation. Nedd8 is the closest relative to ubiquitin compared with other ubiquitin-like proteins. Nedd8 shares ~60% sequence identity, 76% sequence similarity and structural similarity with ubiquitin.<sup>10</sup> Nedd8 was attached to the target protein in a similar manner as in ubiquitination. Neddylation requires E1-activating enzyme (NAE-E1), E2-conjugating enzyme (Ube2M or Ube2F).<sup>9</sup> No Nedd8 specific E3 has yet been identified so far. The primary target for neddylation is Cullins that function as scaffolds for the assembly of multisubunit ubiquitin E3s. A compound called MLN4924 has been found by Millenium/Takeda company to specifically inhibit NAE-E1, the activating enzyme for neddylation.<sup>9</sup> This drug is in a phase II stage and was used to treat multiple myeloma and non-Hodgkin's lymphoma.<sup>9</sup>

Deubiquitinating enzymes (DUBs) are a class of enzymes that can cleave isopeptide bond formed between the C-terminal carboxylate of ubiquitin and a lysine side-chain amino group on the target protein. Besides cleaving the free polyubiquitin chain bearing different linkages, DUB can also remove a single ubiquitin moiety or ubiquitin chain from a target protein.<sup>11, 12</sup> There are close to 100 DUBs in the human proteome. DUBs can be classified into five families: the ubiquitin specific proteases (USP), the ubiquitin C-terminal hydrolases (UCH), the ovarian-tumor (OTU) domain DUBs, the Machado-Joseph (MJD) domain DUBs, and the Jab1/MPN metalloenzyme (JAMM) domain DUBs.<sup>6, 8</sup> Among them, ubiquitin specific proteases (USPs) constitute the largest DUB family. Human USPs are emerging as promising targets for pharmacological intervention because of their connection to a number of human diseases, including prostate, colon and breast cancer,<sup>13, 14</sup> pediatric acute lymphoblastic leukemia,<sup>15</sup> and familial cylindromatosis.<sup>16</sup> The advantage of inhibiting USPs lies in the potential specificity of therapeutic intervention that can lead to better efficacy and reduce nonspecific side effects.

So far only a few small molecule inhibitors were reported for the human DUBs. Through high throughput screening, Liu *et al.* identified a group of isatin *O*-acyl oximes compounds that inhibited UCH-L1 with an IC<sub>50</sub> ranging from 0.8 – 0.94 μM.<sup>17</sup> In cell-based assay, the selected isatin oximes induced the proliferation of the lung cancer cell, supporting an antiproliferative function of UCH-L1. Mermerian *et al.* reported another class of UCH-L1 inhibitors based on 3-amino-2-keto-7H-thieno[2,3-b]pyridin-6-one. This class of compounds inhibits UCH-L1 uncompetitively ( $K_i = 2.8 \mu\text{M}$ ).<sup>18</sup> Hirayama *et al.* carried out *in silico* virtual screening and identified several dihydro-pyrrole compounds that competitively inhibit UCH-L3

(IC<sub>50</sub> = 100 – 150 μM).<sup>19</sup> Over 60 USPs have been identified in the human proteome. However, only two small-molecule inhibitors have been reported targeting human USPs. Colland *et al.* identified a cyano-indenopyrazine derivative, HBX41,108, that inhibits USP7 with a submicromolar IC<sub>50</sub> value.<sup>20</sup> This compound inhibited deubiquitination of p53 both *in vitro* and *in vivo*. Progenra also has developed inhibitor of USP7, i.e. P5091(not reported). Besides USP7, Progenra shows interest in developing inhibitors targeting USP20, USP2a, USP33, and AMSH (associated molecule with the SH3 domain of STAM).<sup>9</sup> Another company, Novartis has patented inhibitors for USP2 and UCH-L3 (not reported).<sup>9</sup> More recently, Lee *et al.* reported a novel USP14 inhibitor 1-[1-(4-fluorophenyl)-2,5-dimethylpyrrol-3-yl]-2-pyrrolidin-1-ylethanone (IC<sub>50</sub> = 4 – 5 μM).<sup>21</sup> This compound was shown to accelerate protein degradation by enhancing the proteasomal activity in reconstituted assay and in mouse embryonic fibroblast cells.

Developing pharmaceutical drugs to inhibiting deubiquitinase is still at an early stage. To date, no deubiquitinase inhibitor has entered the clinical trial yet. Here we selected human USP1 for developing novel USP inhibitors. To our knowledge, the inhibitors for USP1 have never been reported or studied before. USP1 has been implicated in DNA damage response. Previous studies showed that disruption of USP1 in chicken DT40 cells resulted in increased sensitivity to DNA crosslinker.<sup>22</sup> Knockout of the murine USP1 gene in a mouse model resulted in hypersensitivity to mitomycin C.<sup>23</sup> These observations suggest that inhibiting the cellular activity of USP1 will likely sensitize cells to DNA crosslinker. USP1 is known to deubiquitinate two proteins, FANCD2 and PCNA, *in vivo*.<sup>24, 25</sup> Posttranslational modification of PCNA by ubiquitin, either monoubiquitination or K63-linked polyubiquitination, is central to the

normal DNA damage response process in eukaryotes. Reversible ubiquitination of PCNA is essential for translesion synthesis that promotes lesion-bypass synthesis across the damaged base. USP1 also deubiquitinates FANCD2, an important protein in the human Fanconi anemia pathway.<sup>24</sup> Fanconi anemia has been linked to human chromosomal instability and predisposition to cancer.<sup>22</sup> One hallmark of FA patients is the hypersensitivity to DNA-crosslinking agents, such as mitomycin C (MMC) and diepoxybutane (DEB).<sup>23</sup> Thirteen FA proteins are known to form a multi-subunit complex that ubiquitinates FANCD2.<sup>26</sup> Ubiquitination of FANCD2 was shown to be essential for the repair of interstrand crosslink caused by DNA crosslinkers. Because USP1 plays important roles in the two essential DNA damage response pathways, it represents a promising target for small molecule intervention to improve the efficacy of the commonly used DNA damaging drugs by modulating the cells' ability of repairing DNA lesions.

USP1 belongs to a class of DUBs that require interacting partner protein for normal enzymatic function.<sup>11</sup> It was reported that human USP1 forms a stable complex with UAF1.<sup>27</sup> UAF1 is a WD40 repeat containing protein and was reported to interact with USP1, USP12 and USP46.<sup>27, 28</sup> UAF1 may serve as an activator in regulating a variety of human DUBs. USP1 alone possesses low level of deubiquitinating activity. Upon formation of a USP1/UAF1 heterodimeric complex, the catalytic activity of USP1 is greatly stimulated. Given that many DUBs function as a complex *in vivo*, it has become imperative to identify inhibitors against the many DUB complexes. The different binding partners of DUBs could potentially provide the desired specificity toward an otherwise conserved catalytic site. In this chapter, we subjected the human USP1/UAF1 complex to high-throughput screening and

identified potent and specific inhibitors against USP1/UAF1. The best inhibitors for USP1/UAF1, pimozone and GW7647, inhibit USP1/UAF1 through a noncompetitive mechanism. We also demonstrated that the USP1/UAF1 inhibitors sensitize the cisplatin-resistant cancer cells to DNA crosslinker, cisplatin.

## **4.2 Results**

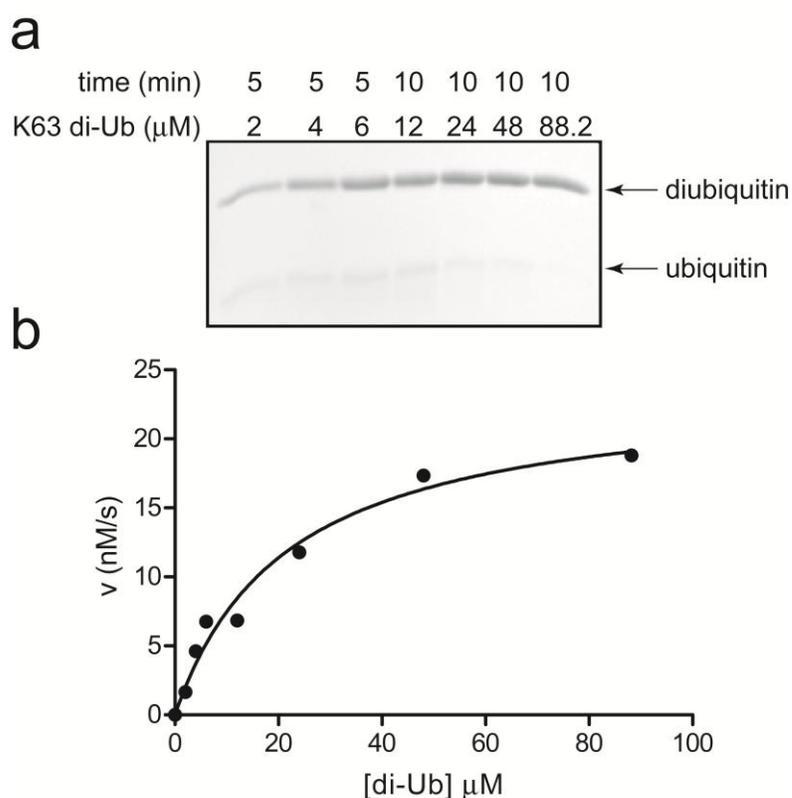
### **4.2.1 qHTS against USP1/UAF1 using Ub-Rho as a substrate**

The qHTS was done by Dr. Anton Simeonov group at NIH Chemical Genomics Center. To date, ubiquitin-7-amido-4-methylcoumarin (Ub-AMC) has been the substrate of choice for most DUB high-throughput screening efforts.<sup>20, 21</sup> However, coumarin-based fluorescence assays are limited by unfavorable spectroscopic properties, which can lead to fluorescence artifacts from test compounds.<sup>29</sup> To alleviate some of the assay interference arising from fluorescent compounds, we utilized the recently identified ubiquitin-rhodamine 110 (Ub-Rho), which exhibits red-shifted fluorescence compared to Ub-AMC, as a substrate.<sup>30</sup> For high-throughput screening, we used the recombinant USP1/UAF1 complex in a 4  $\mu$ L assay volume to screen several bioactive libraries in qHTS mode.<sup>31</sup> Consistently high  $Z'$  factors (average 0.8) were observed throughout the screen, indicating a robust performance. The screen yielded a range of active compounds associated with different potencies ( $IC_{50}$ ) and concentration-response curve quality.

### **4.2.2 Develop an orthogonal diubiquitin cleavage assay**

To validate the top hits using a more physiologically relevant substrate, we developed a gel-based assay using diubiquitin (di-Ub) as a substrate to evaluate the potency of the inhibitors. Diubiquitin as a substrate has been successfully used to

characterize the deubiquitinating activity of DUBs from several families.<sup>32-37</sup> The gel-based diubiquitin assay, being orthogonal to the fluorescence-based assay, served to rule out false positives that arose during the screening, likely due to the interference from compound fluorescence. We obtained quantitative kinetic data of USP1/UAF1 hydrolyzing K63-linked diubiquitin using the gel-based assay ( $k_{\text{cat}} = 0.24 \text{ s}^{-1}$  and  $K_m = 21.8 \text{ }\mu\text{M}$ ) (Figure 4.1). The kinetic values obtained are of the same order as those previously determined for several other DUBs.<sup>29, 33, 35, 37</sup>



**Figure 4.1:** (a) SDS-PAGE showing the progression of the diubiquitin cleavage at different substrate concentration by USP1/UAF1 at different time points. (b) Michaelis-Menten plot of the kinetic data for USP1/UAF1 with K63 diubiquitin as substrate. The USP1/UAF1 concentration was 100 nM.

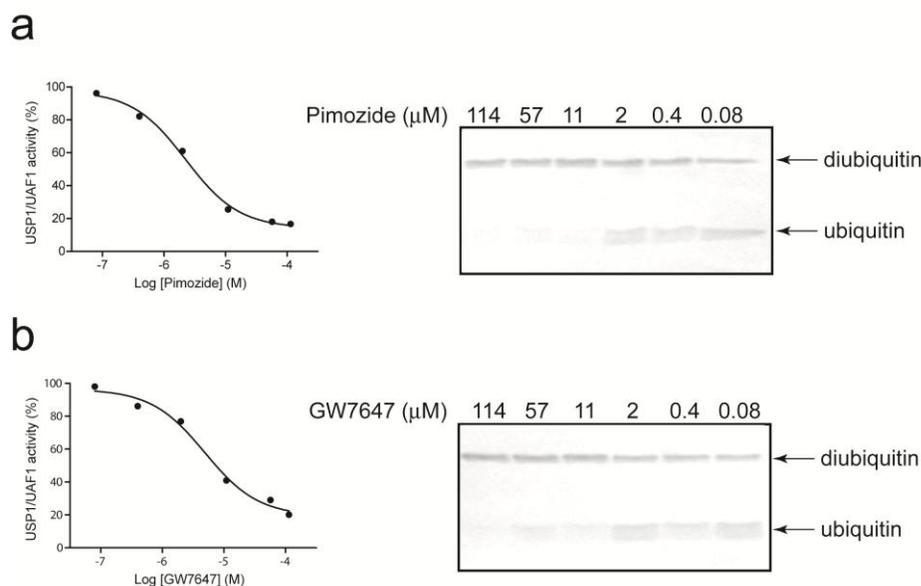
### 4.2.3 Validate the hits using the di-Ub cleavage assay

To evaluate the potency of the inhibitors identified from qHTS, we determined the IC<sub>50</sub> values of the selected top active compounds inhibiting USP1/UAF1-catalyzed cleavage of the K63-linked di-Ub (Table 4.1).

**Table 4.1 The IC<sub>50</sub> (μM) value of the top forty two compounds in inhibiting USP1/UAF1 by Ub-Rho or diubiquitin. NI, no significant inhibition was observed at the highest inhibitor concentration of 114 μM. FI, full inhibition was observed at the lowest inhibitor concentration of 0.08 μM.**

compound	Ub-Rho	diubiquitin	compound	Ub-Rho	diubiquitin
Primaquine	4.0	4.2	ZINC00943071	15.9	23
ZPCK	28.2	5.8	Trifluoperazine	50.1	7.8
Calcimycin	35.5	12.6	Eseroline	2.8	9.3
Flupenthixol	56.2	6.9	Lysuride	35.5	156
GW7647	44.7	4.9	4'-Epidaunomycin	31.6	8.4
CBCromo1_000149	50.1	91	Scoulerine	10.0	4.9
K185	50.1	9.7	BMY 45778	39.8	35.5
Lopac-L-133	50.1	108	SB 206553	56.2	35.8
MK-886	50.1	25.1	Tocris-1675	50.1	21.3
NCGC00162252-02	22.4	21.7	L783281	56.2	4.5
Methylergonovine	56.2	7.0	Tocris-0912	44.7	4.8
CHEBI:114131K	44.7	5.8	Tocris-1133	50.1	15.8
Pimozide	35.5	2.2	SB22489G	44.7	23.8
Quinacrine	79.4	NI	Tocris-1311	25.1	8.3
SU 6656	25.1	57.6	Ergometrine	56.2	9.0
rottlerin	79.4	7.9	bracteoline	39.8	6.5
Ritanserlin	39.8	6.1	oxoglaucine	12.6	3.7
Tosylphenylalanyl Chloromethyl Ketone	35.5	12.6	lysergol	56.2	NI
4-hydroxy-3,3-dimethyl-2H- benzo(g)indole-2,5(3H)-dione	0.9	FI	NCGC00091272-02	56.2	3.0
3-((4-(4-chlorophenyl)piperazin-1-yl) methyl)-1H-pyrrolo(2,3-b)pyridine	50.1	NI	Carmine	56.2	3.8
4-(2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl- 2-naphthalenyl)-1-propenyl)benzoic acid	31.6	5.1	Tocris-1856	56.2	8.9

Out of the 42 compounds, we selected the top five compounds with  $IC_{50}$  values ranging from 2.2  $\mu$ M to 7.9  $\mu$ M. Among them, pimozone and GW7647 are the most potent inhibitors with an  $IC_{50}$  of 2.2  $\mu$ M and 4.9  $\mu$ M, respectively. As shown in Figure 4.2, we observed concentration-dependent inhibition of di-Ub cleavage by



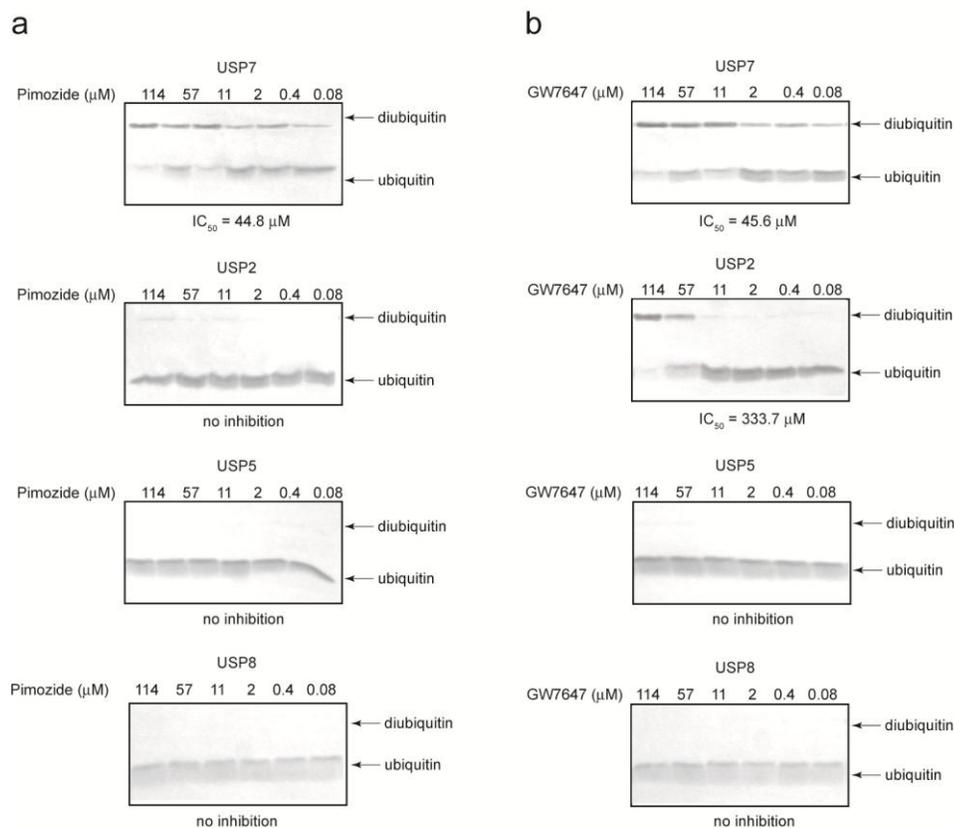
**Figure 4.2: Inhibition of USP1/UAF1 by pimozone (a) and GW7647 (b). Dose-dependent inhibition of USP1/UAF1 activity (left) and SDS-PAGE analysis (right) of the cleavage of K63-linked diubiquitin in the presence of different concentrations of inhibitors are shown.**

pimozone and GW7647. The  $IC_{50}$  was obtained by fitting the dose-response inhibition curve as described in section 4.4.2. Three other compounds, flupenthixol, rottlerin and trifluoperazine, also demonstrated potent inhibition against USP1/UAF1 with  $IC_{50}$  values less than 8  $\mu$ M. The  $IC_{50}$  value determined using di-Ub substrates are in general smaller compared to that determined using Ub-Rho as a substrate. Nonetheless, a good

correlation between the  $IC_{50}$  values determined using the two substrates were observed for the top active compounds.

#### 4.2.4 Specificity of the USP1/UAF1 inhibitors against human USPs

We then determined the specificity of the five compounds in inhibiting human USPs. We selected four most studied human USPs, *i.e.* USP7, USP2, USP5 and USP8, and subjected them to the diubiquitin cleavage assay. Figure 4.3 shows the SDS-PAGE analysis of the cleavage of K63-linked diubiquitin by various human USPs inhibition by Pimozide and GW7647.



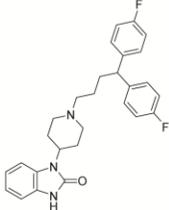
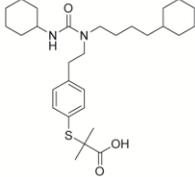
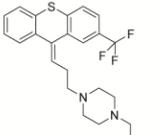
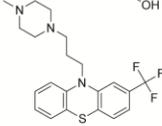
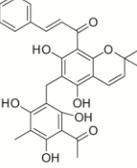
**Figure 4.3:** Pimozide (a) and GW7647 (b) tested as inhibitors against human USP7, USP2, USP5 and USP8 using K63 diubiquitin as substrate.

Based on the measured  $IC_{50}$  (Table 4.2), we found that pimozone and GW7647 are both highly specific for USP1/UAF1. Pimozone inhibits USP7 20-fold more weakly ( $IC_{50} = 44.8 \mu\text{M}$ ) than USP1/UAF1 ( $IC_{50} = 2.2 \mu\text{M}$ ). No inhibition was observed against USP2, USP5 and USP8 even with the highest inhibitor concentration tested (114  $\mu\text{M}$ ). GW7647 also showed good specificity against USP1/UAF1. The  $IC_{50}$  determined for USP7 and USP2 (45.6 and 333.7  $\mu\text{M}$  respectively) are 10 and 68 folds higher than that determined for USP1/UAF1 ( $IC_{50} = 4.9 \mu\text{M}$ ). Moreover, GW7647 does not inhibit USP5 and USP8. In contrast, flupenthixol and trifluoperazine are less specific inhibitors against USP1/UAF1. Both compounds inhibit USP7 with an  $IC_{50}$  comparable to that for USP1/UAF1. Nonetheless, no inhibition was observed against USP2 and USP5. The two compounds demonstrated different inhibition profile against USP8. While no inhibition was observed for trifluoperazine against USP8, a weak inhibition ( $IC_{50} = 19 \mu\text{M}$ ) was observed for flupenthixol. Among the top five inhibitors, rottlerin is the least specific inhibitor. It inhibits USP1/UAF1, USP7, USP2 and USP8 with  $IC_{50}$  ranging from 6.2 to 33.7  $\mu\text{M}$ . Inspection of the rottlerin structure revealed that it contains an  $\alpha, \beta$ -unsaturated carbonyl group, which is a potential Michael acceptor.

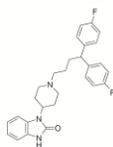
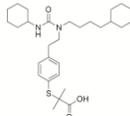
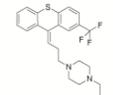
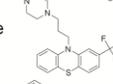
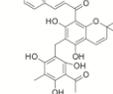
We also tested the top five compounds in inhibiting USP1 alone. USP1 is a poor enzyme and little activity was detected in cleaving the K63-linked di-Ub substrate. Thus, we used Ub-Rho as a substrate to assess the inhibition of USP1 by the top five inhibitors. Pimozone, GW7647, flupenthixol and trifluoperazine demonstrated low level of inhibition against USP1 alone. The estimated  $IC_{50}$  are approximately seven- to fourteen-fold greater compared to the  $IC_{50}$  values in inhibiting USP1/UAF1

determined using Ub-Rho as a substrate (Table 4.3). In contrast, rottlerin is a slightly more potent inhibitor against USP1 compared to USP1/UAF1.

**Table 4.2 The IC<sub>50</sub> (μM) value of the top five compounds in inhibiting human USPs determined using K63-linked diubiquitin substrate. NI, no significant inhibition was observed at the highest inhibitor concentration of 114 μM.**

compound	structure	USP1/UAF1 diubiquitin	USP7 diubiquitin	USP2 diubiquitin	USP5 diubiquitin	USP8 diubiquitin
Pimozide		2.2	44.8	NI	NI	NI
GW7647		4.9	45.6	333.7	NI	NI
Flupenthixol		6.9	12.9	NI	NI	19.0
Trifluoperazine		7.8	8.6	NI	NI	NI
Rottlerin		7.9	13.3	33.7	154.7	6.2

**Table 4.3 The IC<sub>50</sub> (μM) value of the top five compounds in inhibiting human DUBs and other cysteine proteases.**

compound	structure	USP1/UAF1 Ub-Rho	USP1 Ub-Rho	UCH-L3 Ub-Rho	UCH-L1 Ub-Rho	SENP1 SUMO1 -AMC	Caspase-1 Ac-WEHD -AFC	Caspase-3 Ac-DEVD -AFC	Caspase-6 Ac-VEID -AFC	Papain Z-FR -AMC
Pimozide		35.48	>500	>500	>500	>500	>250	>250	>250	>250
GW7647		44.67	>300	>200	>200	>300	41.6	>250	21.1	128.7
Flupenthixol		56.23	>500	>500	>500	>200	>250	>250	>250	>250
Trifluoperazine		50.12	>500	>500	>500	>500	>250	>250	>250	132.9
Rottlerin		79.43	22.6	21.7	12.1	39.1	8.7	4.6	24.6	44.9

#### 4.2.5 Specificity of the USP1/UAF1 inhibitors against the UCH-family DUBs and other cysteine proteases

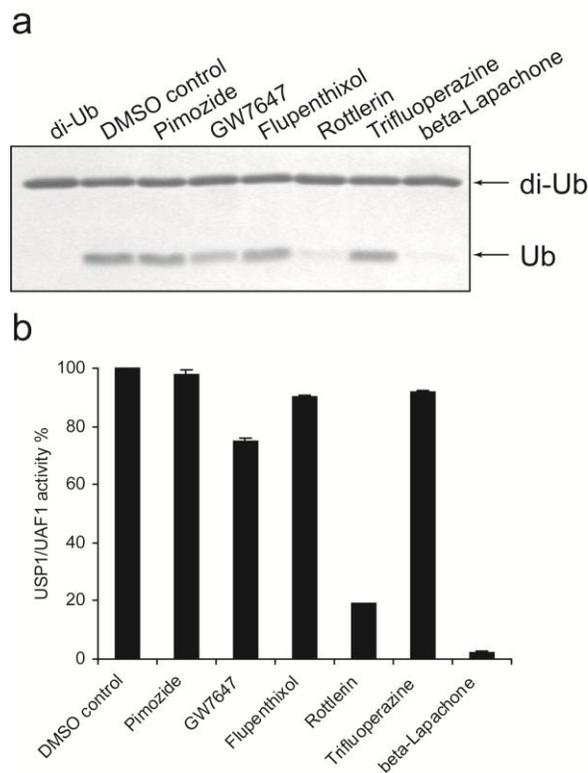
In addition to USPs, we also assessed the specificity of the top five compounds in inhibiting DUBs from the ubiquitin C-terminal hydrolase (UCH) family. Both human UCH-L1 and UCH-L3 were tested in the inhibition assay. Because UCH does not cleave diubiquitin substrate, we used Ub-Rho as a substrate for IC<sub>50</sub> determination. Pimozide, flupenthixol and trifluoperazine are poor inhibitors of UCH-L1 and UCH-L3 (IC<sub>50</sub> > 500 μM). Similarly, GW7647 inhibited both UCHs with IC<sub>50</sub> greater than 200 μM. Rottlerin demonstrated strong inhibition of both UCHs with IC<sub>50</sub> values of 12-21 μM. Thus, pimozide, GW7647, flupenthixol and trifluoperazine are specific inhibitors against USPs compared to the UCH family DUBs. We also assessed the inhibition of a deSUMOylating enzyme, SENP1, by the top five inhibitors

using SUMO1-AMC as a substrate. As summarized in Table 4.3, pimozone, GW7647, flupenthixol and trifluoperazine showed low level of inhibition against SENP1, while rottlerin inhibited SENP1 with an  $IC_{50}$  around 40  $\mu$ M.

We further tested the top five active compounds against other cysteine proteases, including caspases and papain (Table 4.3). Pimozone, flupenthixol and trifluoperazine showed poor inhibition of caspase-1, caspase-3 and caspase-6 ( $IC_{50} > 250 \mu$ M). GW7647 is also a poor inhibitor of caspase-3 ( $IC_{50} > 250 \mu$ M). However, it showed inhibition against caspase-1 and caspase-6 with  $IC_{50}$  values of 41.6  $\mu$ M and 21.1  $\mu$ M, respectively. Pimozone and flupenthixol did not inhibit papain ( $IC_{50} > 250 \mu$ M). GW7647 and trifluoperazine demonstrated little inhibition against papain with  $IC_{50}$  values around 130  $\mu$ M. In comparison, rottlerin inhibited all three caspases and papain with relatively strong inhibition ( $IC_{50}$  ranging from 8 – 45  $\mu$ M).

#### **4.2.6 Pimozone and GW7647 are reversible inhibitors**

We next assessed the reversibility of inhibition of USP1/UAF1 by the top five inhibitors. We followed a protocol that was developed by Copeland and coworkers and subsequently used in a number of studies to determine the reversibility of enzyme inhibition.<sup>38-41</sup> Briefly, USP1/UAF1 (10  $\mu$ M) was incubated with inhibitor at a concentration 10-fold greater than its  $IC_{50}$ . The incubated USP1/UAF1 was rapidly diluted (100-fold) into an assay mixture. If the inhibition is reversible, the remaining USP1/UAF1 activity is expected to be close to that determined in a control reaction, in which no inhibitor was included in the incubation. If the inhibition is irreversible, little recovery of the USP1/UAF1 activity will be observed upon dilution.



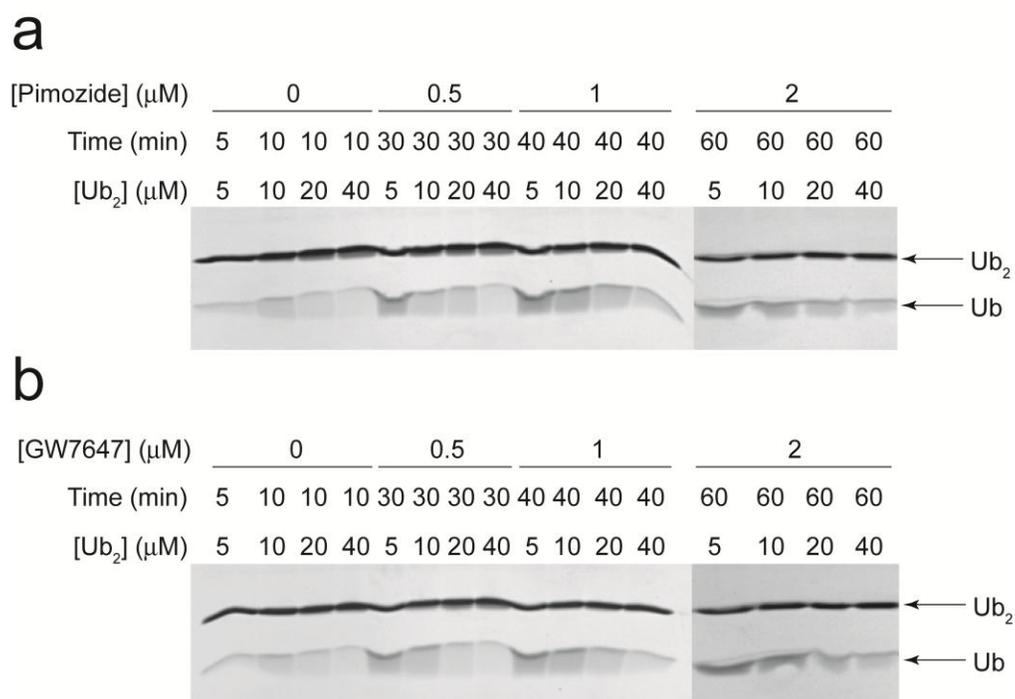
**Figure 4.4: Reversibility of inhibition by USP1/UAF1 inhibitors. (a) SDS-PAGE showing the diubiquitin cleavage by USP1/UAF1 that was pre-incubated with different inhibitors and then rapidly diluted into assay buffer. (b) The remaining activity of USP1/UAF1 treated with five inhibitors. The bar graph represents mean of three independent experiments  $\pm$  s.d. Pimozide, GW7647, flupenthixol and trifluoperazine are reversible inhibitors. Rottlerin is a largely irreversible inhibitor. Beta-Lapachone was tested as an irreversible inhibitor control here.**

Our results (Figure 4.4) indicated that all four inhibitors, *i.e.* pimozide, GW7647, flupenthixol and trifluoperazine are largely reversible inhibitors because USP1/UAF1 treated with these inhibitors exhibited rapid recovery of activity upon dilution. The measured activity of USP1/UAF1 following dilution is close to that determined in the control reaction. In contrast, rottlerin is an irreversible inhibitor of

USP1/UAF1 with only 18% enzymatic activity recovered following the same incubation-dilution protocol. This observation agrees with the notion that rottlerin acts as a Michael acceptor and inhibits USP1/UAF1 by modifying the active site cysteine in USP1/UAF1. Beta-Lapachone known as a Michael acceptor is used as a positive control to show the irreversibility of being the covalent modifier for USP1/UAF1. Negligible USP1/UAF1 remaining activity was observed after incubated with beta-Lapachone.

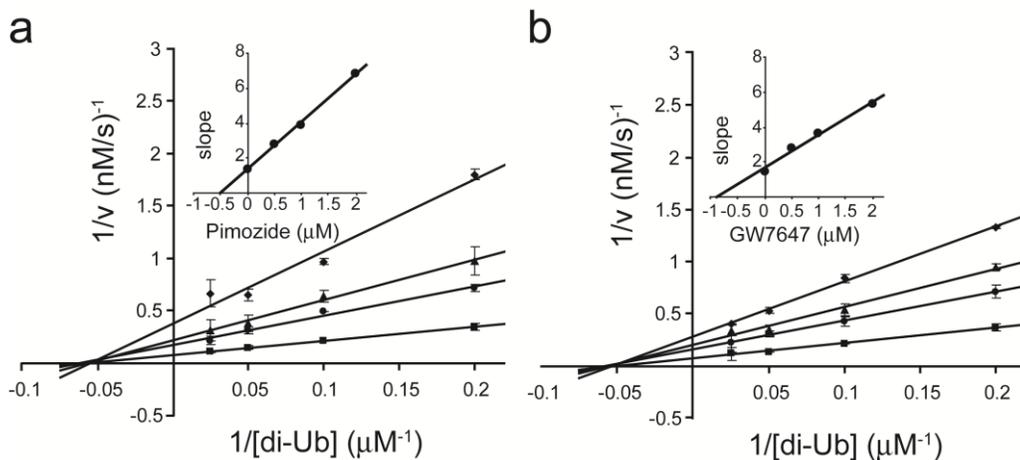
#### 4.2.7 Pimozide and GW7647 inhibit USP1/UAF1 by a noncompetitive mechanism

The inhibition mechanism of pimozide and GW7647 was assessed by measuring the steady-state rate constants at varied inhibitor concentrations using K63-linked di-Ub as a substrate (Figure 4.5).



**Figure 4.5: SDS-PAGE showing the progression of the diubiquitin cleavage of USP1/UAF1 in the presence of 0, 0.5  $\mu\text{M}$ , 1  $\mu\text{M}$  and 2  $\mu\text{M}$  (a) Pimozide and (b) GW7647.**

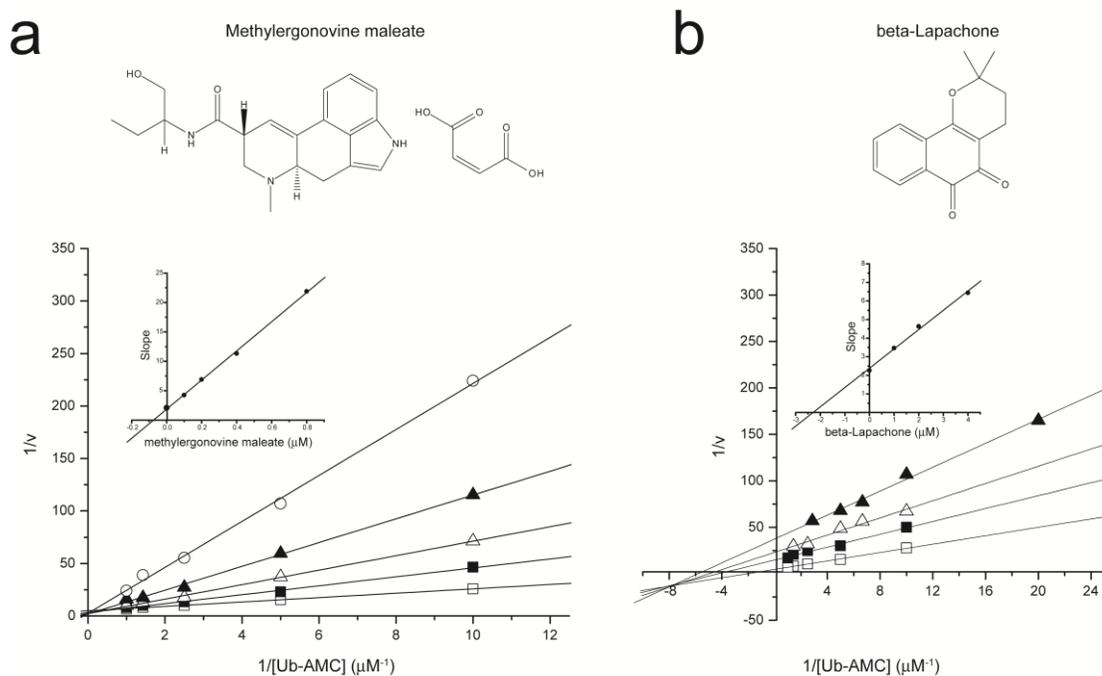
The Lineweaver-Burk plots for both pimozide and GW7647 (Figure 4.6) clearly indicate a noncompetitive inhibition mechanism. The measured inhibition constants ( $K_i$ ) for pimozide and GW7647 are 0.50  $\mu\text{M}$  and 0.75  $\mu\text{M}$ , respectively. These observations suggest that both pimozide and GW7647 bind at a site other than the active site of USP1/UAF1. We can also infer that both inhibitors bind the enzyme and enzyme substrate complex with equal affinity.



**Figure 4.6: Determine the inhibition mechanism of USP1/UAF1 by Pimozide and GW7647. (a) Lineweaver-Burk plots of USP1/UAF1 catalyzed hydrolysis of diubiquitin in the presence of 0 ( $\blacksquare$ ), 0.5 ( $\bullet$ ), 1 ( $\blacktriangle$ ), 2 ( $\blacklozenge$ )  $\mu\text{M}$  pimozide. (b) Lineweaver-Burk plots of USP1/UAF1 catalyzed hydrolysis of diubiquitin in the presence of 0 ( $\blacksquare$ ), 0.5 ( $\bullet$ ), 1 ( $\blacktriangle$ ), 2 ( $\blacklozenge$ )  $\mu\text{M}$  GW7647. Each data point represents mean of two independent experiments  $\pm$  s.d. The  $K_i$  values for pimozide ( $0.50 \pm 0.02 \mu\text{M}$ ) and GW7647 ( $0.75 \pm 0.04 \mu\text{M}$ ) were determined by plotting the slopes obtained from the Lineweaver-Burk plot against the inhibitor concentrations (inset).**

#### 4.2.8 Inhibition mechanism determination of USP1/UAF1 by methylergonovine and beta-Lapachone using Ub-AMC as the substrate

The inhibitory mechanism of methylergonovine and  $\beta$ -lapachone were characterized by adding various amounts of Ub-AMC to USP1/UAF1 in the presence of different concentrations of inhibitor. Lineweaver-Burk plot (Figure 4.7) indicated that methylergonovine was a competitive inhibitor with a  $K_i$  of 75 nM while  $\beta$ -Lapachone was a non-competitive inhibitor with a  $K_i$  of 2.29  $\mu$ M. These data showed that methylergonovine binds the active site of USP1/UAF1, which prevents the interaction of substrate to the enzyme. However,  $\beta$ -lapachone binds both the free enzyme and enzyme-substrate complex somewhere other than the active site. Methylergonovine is much more potent than  $\beta$ -lapachone based on the inhibition constant.



**Figure 4.7: Lineweaver-Burk plot of the inhibition of USP1/UAF1 with (a) methylergonovine maleate. Inhibition was by 0 (□), 100 (■), 200 (Δ), 400 (▲), 800 (○) nM methylergonovine maleate against varying concentrations of Ub-AMC. (b) Inhibition was by 0 (□), 1 (■), 2 (Δ), 4 (▲) μM beta-Lapachone against varying concentrations of Ub-AMC. The  $K_i$  for methylergonovine maleate and beta-Lapachone were determined from a plot of the slopes obtained from the Lineweaver-Burk plot at each concentration of inhibitors (inset).**

Also, the specificity (Table 4.4) of methylergonovine and beta-Lapachone to USP1/UAF1 were evaluated against several other DUBs from the major peptidase family by measuring the  $IC_{50}$ . Methylergonovine is only weakly active towards USP7 (2.55 μM), UBP15 (2.1 μM) and UCH-L3 (1.51 μM). It displayed a decent activity towards USP1 alone (627 nM). Beta-Lapachone showed very poor activity towards USP7 (> 20 μM). It is very weakly active towards USP1/UAF1, USP1 alone and UCH-L3. However, it did not show any inhibitory effect towards UBP15, a yeast ortholog of human USP7. Overall, these results suggest there is some selectivity for methylergonovine in the inhibition of USP1/UAF1 activity. More DUBs could be tested to further verify the specificity of this compound towards USP1/UAF1.

**Table 4.4 The  $IC_{50}$  (μM) value of methylergonovine and beta-Lapachone in inhibiting various DUBs using Ub-AMC as the substrate.**

DUBs	Methylergonovine maleate $IC_{50}$ (μM)	beta-Lapachone $IC_{50}$ (μM)
USP1/UAF1	0.29	1.39
USP1	0.67	5.37
USP7	2.55	> 20
UCH-L3	1.51	2.00
UBP15	2.10	N/A

#### **4.2.9 Validate the hits using the di-Ub cleavage assay for the second set of compounds**

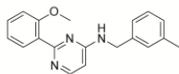
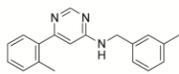
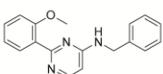
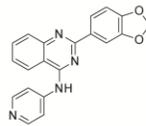
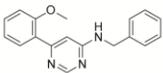
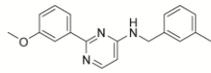
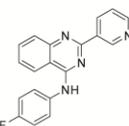
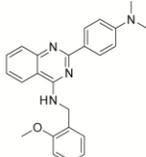
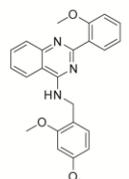
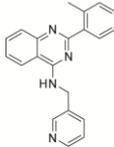
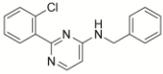
NCGC sent us a second set of compounds in which they all share a core structure containing a pyrimidine ring. We determined the  $IC_{50}$  values of the selected top twenty three active compounds inhibiting USP1/UAF1-catalyzed cleavage of the K63-linked di-Ub (Table 4.5). Out of the 23 compounds, we selected the top four compounds with  $IC_{50}$  values ranging from 2.4  $\mu$ M to 4.2  $\mu$ M (Table 4.6).

We then determined the specificity of the four compounds in inhibiting human USPs. We selected two human USPs, *i.e.* USP7, USP2, and subjected them to the diubiquitin cleavage assay. Based on the measured  $IC_{50}$  (Table 4.6), we found that NCGC00012848-01 is highly specific for USP1/UAF1. NCGC00012848-01 inhibits USP7 with an  $IC_{50}$  of 37.8  $\mu$ M, which is 12-fold higher than that measured for USP1/UAF1 ( $IC_{50}$  = 3.3  $\mu$ M). No inhibition was observed against USP2 even with the highest inhibitor concentration tested (114  $\mu$ M). NCGC00011621-01 also showed good specificity against USP1/UAF1. The  $IC_{50}$  determined for USP7 (19.4  $\mu$ M) are 5 fold higher than that determined for USP1/UAF1 ( $IC_{50}$  = 4.2  $\mu$ M). No inhibition was observed for USP2. In contrast, NCGC00013130-03 and NCGC00073294-02 are less specific inhibitors against USP1/UAF1. Both compounds inhibit USP7 with an  $IC_{50}$  comparable to that for USP1/UAF1. Nonetheless, no inhibition was observed against USP2 for NCGC00013130-03. While an  $IC_{50}$  value of 37.7  $\mu$ M was observed for NCGC00073294-02.

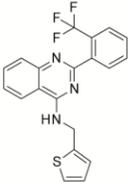
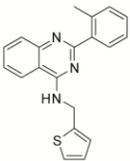
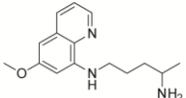
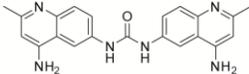
**Table 4.5 The IC<sub>50</sub> (μM) value of twenty three compounds in inhibiting USP1/UAF1 by diubiquitin. NI, no significant inhibition was observed at the highest inhibitor concentration of 114 μM.**

compound (NCGC number)	structure	diubiquitin	compound (NCGC number)	structure	diubiquitin
00012848-01		3.3	00010637-01		NI
00013130-03		2.4	00010067-01		21.3
00073294-02		2.6	00012591-01		26.6
00011621-01		4.2	00012429-01		4.2
00010459-01		NI	00010239-01		80.2
00011663-01		NI	00010434-01		157.9

Table continues next page.

compound (NCGC number)	structure	diubiquitin	compound (NCGC number)	structure	diubiquitin
00010568-01		18.8	00011779-01		26.3
00011249-01		16.9	00012009-01		28.6
00011364-01		10.9	00012196-01		13.0
00011438-01		NI	00012419-01		10.0
00011690-01		14.5	00012495-01		NI
00011749-01		20.0			

**Table 4.6** The IC<sub>50</sub> (μM) value of the top four compounds in inhibiting human USPs determined using K63-linked diubiquitin substrate. NI, no significant inhibition was observed at the highest inhibitor concentration of 114 μM.

compound (NCGC number)	structure	USP1/UAF1	USP7	USP2
		diubiquitin	diubiquitin	diubiquitin
00012848-01		3.3	37.8	NI
00011621-01		4.2	19.4	NI
00073294-02		2.6	7.4	37.7
00013130-03		2.4	2.6	NI

### 4.3 Discussion

Conventional protease inhibitor discovery relies on small molecules that bind to the protease active site. Although this approach has garnered some success, especially when the structural and specificity information of the protease is available, this approach faces challenges due to the common problem of overlapping activity or lack of specificity of the active site-targeting protease inhibitors. In recent years there

has been increasing realization that both specificity and potency of protease inhibitors can be achieved by targeting the protease exosite or allosteric site.<sup>42</sup>

One emerging feature of DUBs and the related deSUMOylating enzymes is that the efficient catalysis requires extensive interactions in the enzyme exosite. This is reflected by the observation that the DUB substrate based on the linear epitope recognized by the active site can only be cleaved in a rate 5 to 6 orders of magnitude lower compared to substrate that contains a complete ubiquitin or SUMO moiety.<sup>43-46</sup> Thus, targeting the DUB exosite is a promising alternative strategy for developing DUB inhibitors. Moreover, many DUBs are regulated by specific interaction with protein partner.<sup>11</sup> For some DUBs, such interaction is required for the DUB activity. Although the detailed mechanism of such regulation remains to be elucidated, allosteric regulation likely plays a central role in modulating the DUB enzymatic activity. Therefore, inhibitor discovery by targeting DUBs at sites other than active site will serve to validate the new paradigm of protease inhibitor development. High-throughput screening has the advantage over the rational inhibitor design that targets solely the active site, and will allow the identification of inhibitors targeting the exosite and allosteric site.

In this study we identified USP1/UAF1 inhibitors, pimozide and GW7647, through high throughput screening against ten thousand compounds. Our kinetic analyses revealed noncompetitive mechanism for both inhibitors, suggesting that both inhibitors bind at a site different from the active site. Interestingly, both compounds are poor inhibitors of USP1 alone, which suggests that they need interactions with UAF1 to achieve tight binding. This again supports the notion that both inhibitors bind outside the USP1 active site. This may be the basis of the

specificity of pimozone and GW7647 in inhibiting USP1/UAF1 among a group of human USPs tested. Moreover, no inhibition was observed for the deSUOMylating enzyme. We also confirmed that pimozone and GW7647 are both reversible inhibitors. Different from a previously reported USP7 inhibitor that act as an uncompetitive inhibitor,<sup>20</sup> pimozone and GW7647 bind to both the enzyme-substrate complex and the free enzyme. The inhibition mechanism of another recently reported inhibitor against USP14 has not been rigorously studied, although a competitive mechanism was suggested.<sup>21</sup>

We found that both pimozone and GW7647 were more potent inhibitors against USP1/UAF1 when the inhibition was assayed using K63-linked diubiquitin substrate when compared to Ub-Rho as a substrate. This raised the interesting possibility that the different binding modes of diubiquitin and Ub-Rho to USP1/UAF1 may determine the potency of the inhibitor. This is particularly true when the inhibitor acts noncompetitively and binds to the DUB-substrate complex. Recent structural determination of K63-linked diubiquitin binding to AMSH LP revealed specific interactions between the proximal ubiquitin and the AMSH core subunit as well as an inserted motif.<sup>32</sup> Therefore, it is possible that pimozone and GW7647 bind to a site on USP1/UAF1 that overlaps with the proximal ubiquitin binding site, thus contributing to the stronger inhibition observed for diubiquitin as a substrate. Another possibility is that pimozone and GW7647 may bind to an allosteric site that influences the binding of proximal ubiquitin due to conformation change in the proximal ubiquitin binding site. In both cases, the USP1/UAF1 inhibitors bind outside of the active site. This observation stresses the necessity of characterizing DUB inhibitors using a physiologically relevant substrate. To date Ub-AMC and Ub-Rho has been used

predominantly as a DUB substrate in high-throughput screening. However, the lack of interactions between ubiquitin acceptor protein and the targeted DUB may result in the false negative in the screening. Therefore, developing new fluorescence-based DUB substrate amenable for high-throughput screening will help to identify potent and specific inhibitors that target the exosite site.

Many anticancer drugs are DNA crosslinkers. The DNA crosslinkers exert effects by damaging the genomic DNA of cancer cells, which in turn triggers apoptosis of the cells. Cisplatin, commonly used as the first-line treatment for various types of cancers, acts by crosslinking the genomic DNA. Despite some success in treating certain cancers, the effectiveness of platinum drugs in clinics has been limited by the resistance that cancer cells usually develop during the course of therapy. Therefore, overcoming cancer cells' resistance to cisplatin and other platinum drugs are imperative for cancer therapy. Since USP1 is involved in both translesion synthesis and Fanconi anemia pathways, inhibiting the cellular USP1 activity will likely sensitize cancer cells to DNA damaging agents. Dr. Yongxing Ai found out that the USP1/UAF1 inhibitors, pimozone and GW7647, can sensitize the human NSCLC cells to cisplatin killing. Given that monoubiquitination of both PCNA and FANCD2 has been implicated in DNA damage tolerance, disruption of the deubiquitination process is likely the underlying mechanism for the enhancement of cisplatin-induced cell killing.

The synergy between cisplatin and USP1/UAF1 inhibitors was quantified by Dr. Yongxing Ai following the method developed by Chou and Talalay. This method allows for the evaluation of two or multiple agents at different concentrations and the different levels of cell killing. Using this methodology, synergy can be

distinguished from additivity and antagonism. Such analyses have been widely adapted for assessing the effectiveness of combination therapy. In this current study, he showed that pimozone alone had a moderate cell killing effect on H596 cells ( $IC_{50}$  of 21.5  $\mu$ M). A combination of cisplatin and pimozone at 1:4 ratio led to a  $IC_{50}$  as low as 2.1  $\mu$ M, which becomes comparable to that determined for the cisplatin-sensitive cell lines. This indicates that combination of USP1 inhibitors and cisplatin render the cisplatin-resistant cells normal sensitivity to cisplatin drug. This is desirable for the USP1/UAF1 inhibitors as a sensitizer of human cancer cell lines to cisplatin or other platinum-based anticancer drugs.

Remarkably, pimozone, a diphenylbutylpiperidine derivative, was also known for its neuroleptic property in treating chronic schizophrenic patients. However, evidence suggested that the pharmacological activity of pimozone may not be limited to neuroleptic function. It has been reported that pimozone inhibits the proliferation of human breast cancer, neuroblastoma, melanoma and non-small cell lung cancer cells.<sup>47-50</sup> Pimozone was also found to induce autophagy and promote long-lived protein degradation.<sup>51</sup> Therefore, it is possible that pimozone likely has multiple pharmacological activities in live cells. The cellular target of pimozone's antineoplastic activity has remained elusive. In this study, we reported that pimozone is a potent and specific inhibitor of a human DUB, USP1/UAF1. Another potent USP1/UAF1 inhibitor is GW7647. GW7647 was initially identified as an agonist of human peroxisome proliferator-activated receptor- $\alpha$  (PPAR $\alpha$ ), a ligand-activated transcription factor involved in the regulation of lipid homeostasis.<sup>52</sup> So far no cytotoxic activity against human cell lines was reported for GW7647. Our studies revealed a modest synergistic effect of GW7647 in killing human NCSLC when combined with cisplatin.

Together our work uncovered the potential new use of two well-characterized pharmacologically active agents.

USP1/UAF1 represents a large group of USPs that was recently found to specifically interact with WD40-repeat protein.<sup>53</sup> Our work in identifying specific, noncompetitive inhibitors against the USP1/UAF1 complex validated the notion of targeting USP complex, in particular the USP-WD40 repeat protein complex, for DUB inhibitor discovery. Moreover, we demonstrated that DUB inhibitors binding outside the active site can provide high selectivity against a large group DUBs in the human proteome. The availability of USP1/UAF1 inhibitors will be a valuable chemical tool for studying the human DNA damage response that involves translesion synthesis and Fanconi anemia pathway. The current USP1/UAF1 inhibitors also serve as a starting point for future medicinal chemistry optimization to improve the potency of the inhibitors.

#### **4.4 Materials and Methods**

##### **4.4.1 Kinetic parameters determination for USP1/UAF1 using K63 di-Ub as a substrate**

(2  $\mu$ M – 88.2  $\mu$ M) K63-linked diubiquitin was incubated in a buffer containing 50 mM HEPES, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA and 1 mM DTT. 100 nM USP1/UAF1 was added to each assay containing different concentrations of diubiquitin to initiate the reaction. The reaction was allowed for 5 min – 10 min at 37 °C, and quenched by the addition of Laemmli sample buffer. The reaction product was separated on a 20% denaturing SDS-PAGE gel and stained with Coomassie Blue. The intensity of the individual diubiquitin and monoubiquitin bands were quantified using

Quantity One 4.3.1 (Biorad, Hercules, CA). The percentage of the conversion was determined and used to calculate the reaction rate. Data were plotted and fit to the Michaelis-Menten equation using GraphPad Prism 5.0 (GraphPad Software, Inc. La Jolla, CA).

#### **4.4.2 Assay inhibition of USP1/UAF1 using K63 di-Ub as a substrate**

Inhibitors at varied concentrations (0.08  $\mu$ M – 114  $\mu$ M) were added to the assay solution containing USP at an appropriate concentration (100 nM for USP1/UAF1), 2  $\mu$ M K63-linked diubiquitin in a buffer containing 50 mM HEPES, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA and 1 mM DTT. The reaction was allowed for 1 hour at 37  $^{\circ}$ C, and quenched by the addition of Laemmli sample buffer. The reaction product was separated on a 20% denaturing SDS-PAGE gel and stained with Coomassie Blue. The intensity of the individual diubiquitin and monoubiquitin bands were quantified using Quantity One 4.3.1 (Biorad, Hercules, CA). The percentage of the conversion was determined and used to calculate the reaction rate. The rate was plotted against inhibitor concentrations and fitted to equation 1 to determine the IC<sub>50</sub> values using GraphPad Prism (GraphPad Software, Inc. La Jolla, CA).

$$Y = Y_0 + \frac{Y_{max} - Y_0}{1 + 10^{X - \text{Log}IC_{50}}} \quad \text{eq. 1}$$

Y is the enzymatic activity, X is equal to Log [I]

#### **4.4.3 Assess the specificity of USP1/UAF1 inhibitors**

The specificity of USP1/UAF1 inhibitors was tested against four human deubiquitinating enzymes, USP7, USP2, USP5 and USP8. Full length USP7 (HAUSP), full length USP5 (Isopeptidase T), full length USP8 and the catalytic domain of USP2 (residue 259 – 605) were purchased from Boston Biochem

(Cambridge, MA). The  $IC_{50}$  value was determined for each enzyme as described in section 4.4.1 using 5 nM USP7, 20 nM USP2, 5 nM USP5 and 80 nM USP8.

#### **4.4.4 Assay inhibition of USP1 and UCH using Ub-rhodamine**

Inhibitor at a range of concentrations (1  $\mu$ M – 500  $\mu$ M) was incubated with 100 nM USP1, 3 nM UCH-L3 or 50 nM UCH-L1 at room temperature for 5 min. UCH-L3 and UCH-L1 were purchased from Boston Biochem (Cambridge, MA). Reaction was initiated by adding 150 nM Ub-rhodamine (LifeSensors Inc., PA) in a buffer containing 50 mM HEPES, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA and 1 mM DTT. The fluorescence increase was monitored at 535 nm upon excitation at 485 nm. The initial velocity was measured and plotted against the inhibitor concentration. The plot was fit to equation 1 to obtain the  $IC_{50}$  value using GraphPad Prism 5 (GraphPad Software, Inc., La Jolla, CA).

#### **4.4.5 Assay inhibition of SENP1 using SUMO1-AMC**

Inhibitor at a range of concentrations (1  $\mu$ M – 500  $\mu$ M) was incubated with 0.8 nM SENP1 at room temperature for 5 min. SENP1 was purchased from Boston Biochem (Cambridge, MA) and it contains the catalytic domain consisting of residues 415 – 643. Reaction was initiated by adding 150 nM SUMO1-AMC (Boston Biochem, Cambridge, MA) in a buffer containing 50 mM HEPES, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA and 1 mM DTT. The fluorescence increase was monitored at 465 nm upon excitation at 340 nm. The initial velocity was measured and plotted against the inhibitor concentration. The plot was fit to equation 1 to obtain the  $IC_{50}$  value using GraphPad Prism 5 (GraphPad Software, Inc., La Jolla, CA).

#### 4.4.6 $K_i$ determination for top USP1/UAF1 inhibitors

The inhibition assay solution contained 100 nM USP1/UAF1, 5 – 40  $\mu$ M K63-linked diubiquitin, and inhibitors at varied concentrations in a buffer containing 50 mM HEPES, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA and 1 mM DTT. The reaction was allowed for 5 – 60 minutes at 37 °C, and quenched by the addition of Laemmli sample buffer at any given time point. The reaction product was separated on a 20% denaturing SDS-PAGE gel and stained with Coomassie Blue. The intensity of the individual diubiquitin and monoubiquitin bands were quantified using Quantity One 4.3.1 (Biorad, Hercules, CA). The percentage of the conversion was determined and used to calculate the reaction rate. The Lineweaver-Burk plot was obtained by plotting  $1/v$  against  $1/[di-Ub]$  at four different inhibitor concentrations. For noncompetitive inhibition,  $K_i$  was determined by fitting the data to equation 2 using Origin 8 (OriginLab Corp., Northampton, MA).

$$\frac{1}{v} = \left[ \frac{K_m}{V_{max}} \left( \frac{1}{[S]} \right) + \frac{1}{V_{max}} \right] \left( 1 + \frac{[I]}{K_i} \right) \quad \text{eq. 2}$$

$$\text{Slope} = \frac{K_m}{V_{max}} \left( 1 + \frac{[I]}{K_i} \right)$$

$$\text{Intercept} = \frac{1}{V_{max}} \left( 1 + \frac{[I]}{K_i} \right)$$

#### 4.4.7 Reversibility of inhibition

10  $\mu$ M USP1/UAF1 was incubated with inhibitor at a concentration 10x the  $IC_{50}$  value at room temperature for 15 minutes. This solution was then diluted by 100 fold into an assay solution (50 mM HEPES buffer, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA, 1 mM DTT) containing 40  $\mu$ M K63-linked diubiquitin and incubated at 37 °C. Aliquot was removed at 10 min and quenched by the addition of Laemmli

sample buffer. The reaction product was separated on a 20% denaturing SDS-PAGE gel and stained with Coomassie Blue. The intensity of the individual diubiquitin and monoubiquitin bands were quantified using Quantity One 4.3.1 (Biorad, Hercules, CA). In the control reaction, 10  $\mu\text{M}$  USP1/UAF1 was incubated with the vehicle (DMSO) and the activity measured for USP1/UAF1 following dilution was treated as 100%.

#### 4.4.8 $K_i$ determination for USP1/UAF1 inhibitor methylergonovine and beta-Lapachone using Ub-AMC as the substrate

The enzyme inhibition assay was carried out by incubating different concentrations of inhibitor (0 – 4  $\mu\text{M}$ ) with 2 nM USP1/UAF1 at room temperature for 5 min. The reaction was initiated by adding various amounts of Ub-AMC (50 nM – 1  $\mu\text{M}$ ) in a buffer containing 50 mM HEPES, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA and 1 mM DTT. Ub-AMC was purchased from Boston Biochem (Cambridge, MA). Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence increase at 435 nm with excitation at 350 nm. The initial velocity was measured for each reaction and the Lineweaver-Burk plot was obtained by plotting  $1/v$  against  $1/[\text{Ub-AMC}]$  at three or four different inhibitor concentrations. For competitive inhibition,  $K_i$  value was determined by fitting the inhibition data to the equation 3 using Origin 8 (OriginLab Corp., Northampton, MA). For noncompetitive inhibition,  $K_{is}$  and  $K_{ii}$  were determined by fitting the data to equation 4.

$$\frac{1}{v} = \frac{K_m}{V_{max}} \left( 1 + \frac{[I]}{K_i} \right) \left( \frac{1}{[S]} \right) + \frac{1}{V_{max}} \quad \text{eq. 3}$$

$$\text{slope} = \frac{K_m}{V_{max}} \left( 1 + \frac{[I]}{K_i} \right)$$

$$\frac{1}{v} = \left[ \frac{K_m}{V_{max}} \left( \frac{1}{[S]} \right) + \frac{1}{V_{max}} \right] \left( 1 + \frac{[I]}{K_i} \right) \quad \text{eq. 4}$$

$$\text{slope} = \frac{K_m}{V_{max}} \left( 1 + \frac{[I]}{K_{is}} \right)$$

$$\text{intercept} = \frac{1}{V_{max}} \left( 1 + \frac{[I]}{K_{ii}} \right)$$

#### 4.4.9 Assess the specificity of USP1/UAF1 inhibitors methylergonovine and beta-Lapachone

Inhibitor at a range of concentrations (0.1 nM – 1 mM) was incubated with DUBs (2 nM for USP1/UAF1, 30 nM for USP1, 50 pM for UCH-L3, 0.5 nM for USP7, 0.5 nM for UBP15) at room temperature for 5 min. UCH-L3 and USP7 (full length) were purchased from Boston Biochem (Cambridge, MA). Deubiquitylating assay was initiated by adding 100 nM Ub-AMC in a buffer containing 50 mM HEPES, pH 7.8, 0.1 mg/ml BSA, 0.5 mM EDTA and 1 mM DTT. Release of AMC (7-amido-4-methylcoumarin) was monitored by fluorescence increase at 435 nm with excitation at 350 nm. The initial velocity was measured and plotted against each inhibitor concentration. To determine the IC<sub>50</sub> value the plot was fitted to equation 1 using GraphPad Prism 5 (GraphPad Software, Inc. La Jolla, CA).

## REFERENCES

1. Hoeller, D.; Hecker, C. M.; Dikic, I., Ubiquitin and ubiquitin-like proteins in cancer pathogenesis. *Nat Rev Cancer* **2006**, *6* (10), 776-788.
2. Rubinsztein, D. C., The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature* **2006**, *443* (7113), 780-786.
3. Gao, G.; Luo, H., The ubiquitin-proteasome pathway in viral infections. *Can J Physiol Pharmacol* **2006**, *84* (1), 5-14.
4. Kirkin, V.; Dikic, I., Role of ubiquitin- and Ubl-binding proteins in cell signaling. *Curr Opin Cell Biol* **2007**, *19* (2), 199-205.
5. Adams, J., The proteasome: a suitable antineoplastic target. *Nat Rev Cancer* **2004**, *4* (5), 349-360.
6. Adams, J.; Palombella, V. J.; Sausville, E. A.; Johnson, J.; Destree, A.; Lazarus, D. D.; Maas, J.; Pien, C. S.; Prakash, S.; Elliott, P. J., Proteasome inhibitors: a novel class of potent and effective antitumor agents. *Cancer Res* **1999**, *59* (11), 2615-2622.
7. Adams, J.; Kauffman, M., Development of the proteasome inhibitor Velcade (Bortezomib). *Cancer Invest* **2004**, *22* (2), 304-11.
8. Guedat, P.; Colland, F., Patented small molecule inhibitors in the ubiquitin proteasome system. *BMC Biochem* **2007**, *8 Suppl 1*, S14.
9. Cohen, P.; Tcherpakov, M., Will the ubiquitin system furnish as many drug targets as protein kinases? *Cell* **2010**, *143* (5), 686-93.
10. Rabut, G.; Peter, M., Function and regulation of protein neddylation. 'Protein modifications: beyond the usual suspects' review series. *EMBO Rep* **2008**, *9* (10), 969-76.
11. Reyes-Turcu, F. E.; Ventii, K. H.; Wilkinson, K. D., Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu Rev Biochem* **2009**, *78*, 363-397.
12. Amerik, A. Y.; Hochstrasser, M., Mechanism and function of deubiquitinating enzymes. *Biochim Biophys Acta* **2004**, *1695* (1-3), 189-207.

13. Priolo, C.; Tang, D.; Brahamandan, M.; Benassi, B.; Sicinska, E.; Ogino, S.; Farsetti, A.; Porrello, A.; Finn, S.; Zimmermann, J.; Febbo, P.; Loda, M., The isopeptidase USP2a protects human prostate cancer from apoptosis. *Cancer Res* **2006**, *66* (17), 8625-32.
14. Popov, N.; Wanzel, M.; Madiredjo, M.; Zhang, D.; Beijersbergen, R.; Bernards, R.; Moll, R.; Elledge, S. J.; Eilers, M., The ubiquitin-specific protease USP28 is required for MYC stability. *Nat Cell Biol* **2007**, *9* (7), 765-774.
15. De Pitta, C.; Tombolan, L.; Campo Dell'Orto, M.; Accordi, B.; te Kronnie, G.; Romualdi, C.; Vitulo, N.; Basso, G.; Lanfranchi, G., A leukemia-enriched cDNA microarray platform identifies new transcripts with relevance to the biology of pediatric acute lymphoblastic leukemia. *Haematologica* **2005**, *90* (7), 890-898.
16. Kovalenko, A.; Chable-Bessia, C.; Cantarella, G.; Israel, A.; Wallach, D.; Courtois, G., The tumour suppressor CYLD negatively regulates NF-kappaB signalling by deubiquitination. *Nature* **2003**, *424* (6950), 801-805.
17. Liu, Y.; Lashuel, H. A.; Choi, S.; Xing, X.; Case, A.; Ni, J.; Yeh, L. A.; Cuny, G. D.; Stein, R. L.; Lansbury, P. T., Jr., Discovery of inhibitors that elucidate the role of UCH-L1 activity in the H1299 lung cancer cell line. *Chemistry & biology* **2003**, *10* (9), 837-846.
18. Mermerian, A. H.; Case, A.; Stein, R. L.; Cuny, G. D., Structure-activity relationship, kinetic mechanism, and selectivity for a new class of ubiquitin C-terminal hydrolase-L1 (UCH-L1) inhibitors. *Bioorg Med Chem Lett* **2007**, *17* (13), 3729-3732.
19. Hirayama, K.; Aoki, S.; Nishikawa, K.; Matsumoto, T.; Wada, K., Identification of novel chemical inhibitors for ubiquitin C-terminal hydrolase-L3 by virtual screening. *Bioorg Med Chem* **2007**, *15* (21), 6810-6818.
20. Colland, F.; Formstecher, E.; Jacq, X.; Reverdy, C.; Planquette, C.; Conrath, S.; Trouplin, V.; Bianchi, J.; Aushev, V. N.; Camonis, J.; Calabrese, A.; Borg-Capra, C.; Sippl, W.; Collura, V.; Boissy, G.; Rain, J. C.; Guedat, P.; Delansorne, R.; Daviet, L., Small-molecule inhibitor of USP7/HAUSP ubiquitin protease stabilizes and activates p53 in cells. *Molecular cancer therapeutics* **2009**, *8* (8), 2286-2295.

21. Lee, B. H.; Lee, M. J.; Park, S.; Oh, D. C.; Elsasser, S.; Chen, P. C.; Gartner, C.; Dimova, N.; Hanna, J.; Gygi, S. P.; Wilson, S. M.; King, R. W.; Finley, D., Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* **2010**, *467* (7312), 179-184.
22. Patel, K. J.; Joenje, H., Fanconi anemia and DNA replication repair. *DNA Repair (Amst)* **2007**, *6* (7), 885-890.
23. Auerbach, A. D.; Rogatko, A.; Schroeder-Kurth, T. M., International Fanconi Anemia Registry: relation of clinical symptoms to diepoxybutane sensitivity. *Blood* **1989**, *73* (2), 391-396.
24. Nijman, S. M.; Huang, T. T.; Dirac, A. M.; Brummelkamp, T. R.; Kerkhoven, R. M.; D'Andrea, A. D.; Bernards, R., The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. *Mol Cell* **2005**, *17* (3), 331-339.
25. Huang, T. T.; Nijman, S. M.; Mirchandani, K. D.; Galardy, P. J.; Cohn, M. A.; Haas, W.; Gygi, S. P.; Ploegh, H. L.; Bernards, R.; D'Andrea, A. D., Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat Cell Biol* **2006**, *8* (4), 339-347.
26. Alpi, A. F.; Patel, K. J., Monoubiquitylation in the Fanconi anemia DNA damage response pathway. *DNA Repair (Amst)* **2009**, *8* (4), 430-435.
27. Cohn, M. A.; Kowal, P.; Yang, K.; Haas, W.; Huang, T. T.; Gygi, S. P.; D'Andrea, A. D., A UAF1-containing multisubunit protein complex regulates the Fanconi anemia pathway. *Mol Cell* **2007**, *28* (5), 786-797.
28. Cohn, M. A.; Kee, Y.; Haas, W.; Gygi, S. P.; D'Andrea, A. D., UAF1 is a subunit of multiple deubiquitinating enzyme complexes. *J Biol Chem* **2009**, *284* (8), 5343-51.
29. Simeonov, A.; Jadhav, A.; Thomas, C. J.; Wang, Y.; Huang, R.; Southall, N. T.; Shinn, P.; Smith, J.; Austin, C. P.; Auld, D. S.; Inglese, J., Fluorescence spectroscopic profiling of compound libraries. *J Med Chem* **2008**, *51* (8), 2363-2371.
30. Hassiepen, U.; Eidhoff, U.; Meder, G.; Bulber, J. F.; Hein, A.; Bodendorf, U.; Lorthiois, E.; Martoglio, B., A sensitive fluorescence intensity assay for deubiquitinating proteases using ubiquitin-rhodamine110-glycine as substrate. *Anal Biochem* **2007**, *371* (2), 201-207.

31. Inglese, J.; Auld, D. S.; Jadhav, A.; Johnson, R. L.; Simeonov, A.; Yasgar, A.; Zheng, W.; Austin, C. P., Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proc Natl Acad Sci U S A* **2006**, *103* (31), 11473-11478.
32. Sato, Y.; Yoshikawa, A.; Yamagata, A.; Mimura, H.; Yamashita, M.; Ookata, K.; Nureki, O.; Iwai, K.; Komada, M.; Fukai, S., Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. *Nature* **2008**, *455* (7211), 358-362.
33. Yao, T.; Song, L.; Xu, W.; DeMartino, G. N.; Florens, L.; Swanson, S. K.; Washburn, M. P.; Conaway, R. C.; Conaway, J. W.; Cohen, R. E., Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1. *Nat Cell Biol* **2006**, *8* (9), 994-1002.
34. Amerik, A.; Swaminathan, S.; Krantz, B. A.; Wilkinson, K. D.; Hochstrasser, M., In vivo disassembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome. *EMBO J* **1997**, *16* (16), 4826-4838.
35. Virdee, S.; Ye, Y.; Nguyen, D. P.; Komander, D.; Chin, J. W., Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol* **2010**, *6* (10), 750-757.
36. Bremm, A.; Freund, S. M.; Komander, D., Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nat Struct Mol Biol* **2010**, *17* (8), 939-947.
37. Cooper, E. M.; Boeke, J. D.; Cohen, R. E., Specificity of the BRISC deubiquitinating enzyme is not due to selective binding to Lys63-linked polyubiquitin. *J Biol Chem* **2009**, *285* (14), 10344-10352.
38. Copeland, R. A.; ebrary Inc., Evaluation of enzyme inhibitors in drug discovery a guide for medicinal chemists and pharmacologists. In *Methods of biochemical analysis v. 46* [Online] Wiley-Interscience: Hoboken, N.J., 2005; pp. xvii, 271 p. <http://site.ebrary.com/lib/princeton/Doc?id=10114176>.
39. Liu, T.; Toriyabe, Y.; Kazak, M.; Berkman, C. E., Pseudoirreversible inhibition of prostate-specific membrane antigen by phosphoramidate peptidomimetics. *Biochemistry* **2008**, *47* (48), 12658-12660.

40. Anderson, K.; Lai, Z.; McDonald, O. B.; Stuart, J. D.; Nartey, E. N.; Hardwicke, M. A.; Newlander, K.; Dhanak, D.; Adams, J.; Patrick, D.; Copeland, R. A.; Tummino, P. J.; Yang, J., Biochemical characterization of GSK1070916, a potent and selective inhibitor of Aurora B and Aurora C kinases with an extremely long residence time1. *Biochem J* **2009**, *420* (2), 259-265.
41. Ma, X. Q.; Zhang, H. J.; Zhang, Y. H.; Chen, Y. H.; Wu, F.; Du, J. Q.; Yu, H. P.; Zhou, Z. L.; Li, J. Y.; Nan, F. J.; Li, J., Novel irreversible caspase-1 inhibitor attenuates the maturation of intracellular interleukin-1beta. *Biochem Cell Biol* **2007**, *85* (1), 56-65.
42. Drag, M.; Salvesen, G. S., Emerging principles in protease-based drug discovery. *Nat Rev Drug Discov* **2010**, *9* (9), 690-701.
43. Renatus, M.; Parrado, S. G.; D'Arcy, A.; Eidhoff, U.; Gerhartz, B.; Hassiepen, U.; Pierrat, B.; Riedl, R.; Vinzenz, D.; Worpenberg, S.; Kroemer, M., Structural basis of ubiquitin recognition by the deubiquitinating protease USP2. *Structure* **2006**, *14* (8), 1293-1302.
44. Mikolajczyk, J.; Drag, M.; Bekes, M.; Cao, J. T.; Ronai, Z.; Salvesen, G. S., Small ubiquitin-related modifier (SUMO)-specific proteases: profiling the specificities and activities of human SENPs. *J Biol Chem* **2007**, *282* (36), 26217-26224.
45. Drag, M.; Mikolajczyk, J.; Krishnakumar, I. M.; Huang, Z.; Salvesen, G. S., Activity profiling of human deSUMOylating enzymes (SENPs) with synthetic substrates suggests an unexpected specificity of two newly characterized members of the family. *Biochem J* **2008**, *409* (2), 461-9.
46. Reyes-Turcu, F. E.; Horton, J. R.; Mullally, J. E.; Heroux, A.; Cheng, X.; Wilkinson, K. D., The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* **2006**, *124* (6), 1197-208.
47. Strobl, J. S.; Kirkwood, K. L.; Lantz, T. K.; Lewine, M. A.; Peterson, V. A.; Worley, J. F., 3rd, Inhibition of human breast cancer cell proliferation in tissue culture by the neuroleptic agents pimozide and thioridazine. *Cancer Res* **1990**, *50* (17), 5399-5405.
48. Strobl, J. S.; Melkounian, Z.; Peterson, V. A.; Hylton, H., The cell death response to gamma-radiation in MCF-7 cells is enhanced by a neuroleptic drug, pimozide. *Breast Cancer Res Treat* **1998**, *51* (1), 83-95.

49. Wiklund, E. D.; Catts, V. S.; Catts, S. V.; Ng, T. F.; Whitaker, N. J.; Brown, A. J.; Lutze-Mann, L. H., Cytotoxic effects of antipsychotic drugs implicate cholesterol homeostasis as a novel chemotherapeutic target. *Int J Cancer* **2010**, *126* (1), 28-40.
50. Neifeld, J. P.; Tormey, D. C.; Baker, M. A.; Meyskens, F. L., Jr.; Taub, R. N., Phase II trial of the dopaminergic inhibitor pimozide in previously treated melanoma patients. *Cancer Treat Rep* **1983**, *67* (2), 155-157.
51. Zhang, L.; Yu, J.; Pan, H.; Hu, P.; Hao, Y.; Cai, W.; Zhu, H.; Yu, A. D.; Xie, X.; Ma, D.; Yuan, J., Small molecule regulators of autophagy identified by an image-based high-throughput screen. *Proc Natl Acad Sci U S A* **2007**, *104* (48), 19023-19028.
52. Brown, P. J.; Stuart, L. W.; Hurley, K. P.; Lewis, M. C.; Winegar, D. A.; Wilson, J. G.; Wilkison, W. O.; Ittoop, O. R.; Willson, T. M., Identification of a subtype selective human PPARalpha agonist through parallel-array synthesis. *Bioorg Med Chem Lett* **2001**, *11* (9), 1225-1227.
53. Sowa, M. E.; Bennett, E. J.; Gygi, S. P.; Harper, J. W., Defining the human deubiquitinating enzyme interaction landscape. *Cell* **2009**, *138* (2), 389-403.

## APPENDIX

Title: Chemically ubiquitylated  
histone H2B stimulates  
hDot1L-mediated  
intranucleosomal methylation

Author: Robert K. McGinty, Jaehoon  
Kim, Champak Chatterjee,  
Robert G. Roeder and Tom W.  
Muir

Publication Nature  
:

Publisher: Nature Publishing Group

Date: Apr 30, 2008

Copyright © 2008, Nature Publishing  
Group

Order Completed

Thank you very much for your order.

This is a License Agreement between Junjun Chen ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the [payment terms and conditions](#).

[Get the printable license](#).

License Number 2562281225512

License date Dec 05, 2010

Licensed content  
publisher Nature Publishing Group

Logged in as:  
Junjun Chen

LOGOUT

Licensed content publication	Nature
Licensed content title	Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation
Licensed content author	Robert K. McGinty, Jaehoon Kim, Champak Chatterjee, Robert G. Roeder and Tom W. Muir
Licensed content date	Apr 30, 2008
Type of Use	reuse in a thesis/dissertation
Volume number	453
Issue number	7196
Requestor type	academic/educational
Format	print and electronic
Portion	Full article
Circulation/distribution	<501
Author of this NPG article	no
NPG Code	
Title of your thesis / dissertation	INVESTIGATION OF CHEMICAL UBIQUITINATION OF PCNA AND MECHANISM, INHIBITION OF USP1/UAF1 & THE MOLECULAR RECOGNITION OF RNA BY THE PSEUDOURIDINE SYNTHASE RLUA
Expected completion date	Jan 2011
Estimated size (number of pages)	200
Total	0.00 USD

Title: Engineered diubiquitin synthesis reveals Lys29- isopeptide specificity of an OTU deubiquitinase

Author: Satpal Virdee, Yu Ye, Duy P Nguyen, David Komander, Jason W Chin

Publication Nature Chemical Biology  
:

Publisher: Nature Publishing Group

Date: Aug 29, 2010

Copyright © 2010, Nature Publishing Group

Logged in as:  
Junjun Chen

LOGOUT

Order Completed

Thank you very much for your order.

This is a License Agreement between Junjun Chen ("You") and Nature Publishing Group ("Nature Publishing Group"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the [payment terms and conditions](#).

[Get the printable license](#).

License Number 2562291253383

License date Dec 05, 2010

Licensed content publisher Nature Publishing Group

Licensed content publication Nature Chemical Biology

Licensed content title	Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase
Licensed content author	Satpal Virdee, Yu Ye, Duy P Nguyen, David Komander, Jason W Chin
Licensed content date	Aug 29, 2010
Type of Use	reuse in a thesis/dissertation
Volume number	6
Issue number	10
Requestor type	academic/educational
Format	print and electronic
Portion	Full article
Circulation/distribution	<501
Author of this NPG article	no
NPG Code	
Title of your thesis / dissertation	INVESTIGATION OF CHEMICAL UBIQUITINATION OF PCNA AND MECHANISM, INHIBITION OF USP1/UAF1 & THE MOLECULAR RECOGNITION OF RNA BY THE PSEUDOURIDINE SYNTHASE RLUA
Expected completion date	Jan 2011
Estimated size (number of pages)	200
Total	0.00 USD