

**UTILIZING PROTEIN CRYSTALLOGRAPHY IN THE STRUCTURAL  
CHARACTERIZATION OF PROTEIN COMPLEXES**

by

Elizabeth Diaz

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Chemistry and Biochemistry

Spring 2018

© 2018 Elizabeth Diaz  
All Rights Reserved

**UTILIZING PROTEIN CRYSTALLOGRAPHY IN THE STRUCTURAL  
CHARACTERIZATION OF PROTEIN COMPLEXES**

by

Elizabeth Diaz

Approved: \_\_\_\_\_  
Sharon Rozovsky, Ph.D.  
Professor in charge of thesis on behalf of the Advisory Committee

Approved: \_\_\_\_\_  
Brian J. Bahnson, Ph.D.  
Chair of the Department of Chemistry and Biochemistry

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

Approved: \_\_\_\_\_  
Ann L. Ardis, Ph.D.  
Senior Vice Provost for Graduate and Professional Education

## ACKNOWLEDGMENTS

There is a saying that goes “no one achieves anything alone,” and that phrase has never held more meaning to me than during this journey. I know for certain that earning this degree part-time while also juggling a career would not be possible without the wonderful and supportive people in my life. To my thesis advisor, Dr. Sharon Rozovsky, thank you for your patience, for your guidance and direction, and for your willingness to work around my schedule. I will always remember your drive and passion for the work. My deepest gratitude to the rest of the Rozovsky lab. You truly exemplify the meaning of teamwork through your willingness to help one another at all times, and you have made my life all the more enjoyable. I would also like to thank Dr. Brian Bahnson for all of his assistance with the crystallography and for always being willing to answer any questions.

Mom and Dad, words are not enough to express how blessed I am to have you. Your prayers, your never-ending encouragement, and your desire to see me reach my full potential have aided me throughout my lifetime. Thank you for always believing in me.

This also would not have been possible without my Chemours team. To my manager, Amber Wellman, your support and flexibility during these last few years has made this all possible. Last but certainly not least, a special thank you to my colleague and mentor, Alex Marchione, who went above and beyond to ease my workload whenever possible. Thank you for always being willing to lend an ear and offer sage advice.

## TABLE OF CONTENTS

LIST OF TABLES .....	vi
LIST OF FIGURES .....	vii
ABSTRACT .....	x

### Chapter

1	INTRODUCTION.....	1
1.1	Introduction to p97 .....	1
1.2	Introduction to Selenoprotein S.....	7
1.3	Introduction to p97 Complexes .....	10
1.4	Introduction to GB1 .....	13
1.5	Introduction to X-Ray Crystallography.....	16
1.5.1	Principles of X-Ray Crystallography.....	16
1.5.2	Vapor Diffusion Crystallization.....	18
1.5.3	Structure Solving and Data Processing.....	22
1.6	Conclusion.....	27
	REFERENCES .....	29
2	CRYSTALLIZATION OF FULL LENGTH P97 AND P97 ND1 .....	40
2.1	Introduction .....	40
2.1.1	Published Crystallization Conditions for p97 .....	40
2.1.1.1	Published Conditions for p97 ND1 .....	40
2.1.1.2	Published Conditions for Full Length p97 .....	42
2.2	Sample Purification and Preparation for Crystallization.....	43
2.3	Crystallization of p97 ND1 .....	44
2.3.1	Crystallization Conditions .....	44
2.3.2	Crystal Tray Monitoring .....	45
2.3.3	Results for ND1 without Polyhistidine-tag.....	46
2.3.4	Results for ND1 with Polyhistidine-tag.....	47
2.4	Crystallization of p97 Full Length.....	48

	2.4.1 Crystallization Conditions .....	48
	2.4.2 Crystal Tray Monitoring .....	49
	2.4.3 Results.....	50
	REFERENCES .....	71
3	CRYSTALLIZATION OF P97 ND1 IN COMPLEX WITH SELENOPROTEIN S C TERMINAL PEPTIDE .....	73
	3.1 Introduction.....	73
	3.1.1 Previous Crystallization Conditions of Complexed p97.....	73
	3.1.2 C Terminal Selenoprotein S Peptide.....	74
	3.1.3 Preparation of p97 ND1 and C Terminal Selenoprotein S Peptide Complex .....	74
	3.2 Crystallization of p97 ND1 in Complex with SELENOS Peptide .....	75
	3.2.1 Crystallization Conditions .....	75
	3.2.2 Crystal Tray Monitoring .....	76
	3.2.3 Results.....	77
	REFERENCES .....	92
4	CRYSTALLIZATION OF GB1 VARIANTS.....	94
	4.1 Crystallization of GB1 Variants.....	94
	4.1.1 Crystallization Conditions .....	94
	4.1.2 Crystal Tray Monitoring .....	95
	4.1.3 Results.....	95
	REFERENCES.....	107
5	CONCLUSION AND FUTURE WORK.....	108
	5.1 Crystallization of p97/SELENOS Complex .....	108
	5.2 Crystallization of Six GB1 Selenomethionine Variants .....	111

## LIST OF TABLES

Table 2.1	Published Crystallization Conditions for p97 ND1 .....	52
Table 2.2	Published Crystallization Conditions for Full Length p97 .....	53
Table 2.3	Conditions for In-House p97 ND1 Screening 1 (Without Polyhistidine-tag) .....	54
Table 2.4	Conditions for In-House p97 ND1 Screening 2 (Without Polyhistidine-tag) .....	57
Table 2.5	Conditions for In-House p97 ND1 Screening (With Polyhistidine-tag) .	59
Table 2.6	Conditions for In-House First Full Length p97 Screening .....	61
Table 2.7	Conditions for In-House Second Full Length p97 Screening .....	63
Table 3.1	Published Crystallization Conditions for complexed p97 .....	80
Table 3.2	Published p97 Complex Binding Conditions .....	81
Table 3.3	Crystallization Conditions for Coarse Screen 1 (1:5 protein peptide ratio and 1:1 drop ratio) .....	82
Table 3.4	Crystallization Conditions for Coarse Screen 2 (10mg/ml protein, 1:2 and 1:3 drop ratio) .....	84
Table 3.5	Crystallization Conditions for Coarse Screen 2 (10mg/ml protein, 1:3 drop ratio) .....	86
Table 3.6	Comparison of successful in-house Crystallization conditions .....	88
Table 4.1	Previously Published Conditions for GB1 .....	98
Table 4.2	Initial Coarse Screen Conditions for GB1 Variants .....	99
Table 4.3	Fine Screen Conditions used for All GB1 Variants .....	101

## LIST OF FIGURES

Figure 1.1	Structure of hexameric p97 wild type in complex with ADP .....	4
Figure 1.2	Sequence and secondary structure for p97 .....	5
Figure 1.3	Structure of hexameric p97 ND1 (residues 1-480) wild type in complex with ADP .....	6
Figure 1.4a	Amino Acid Sequence of Selenoprotein S .....	9
Figure 1.4b	Domain Structure of SELENOS.....	9
Figure 1.4c	SELENOS Domain Structure of Coiled-coil region .....	9
Figure 1.5	Minimal VIM Consensus Sequence .....	12
Figure 1.6	Structure and sequence of immunoglobulin binding domain B1 of streptococcal protein G (GB1) .....	15
Figure 1.7	Crystallization Phase Diagram .....	26
Figure 1.8	Graphical Representation of the Concentration Relationship of Protein and Reagents in Varying Drop Ratios .....	27
Figure 2.1	p97 ND1 (Without Polyhistidine-tag) Screen 1, Well A3, 11.25x Magnification .....	65
Figure 2.2	p97 ND1 (Without Polyhistidine-tag) Screen 2, Well B4, 11.25x magnification.....	65
Figure 2.3	p97 ND1 (Without Polyhistidine-tag) Screen 2, Well B8, 11.25x magnification.....	66
Figure 2.4	p97 ND1 (With Polyhistidine-tag), Well C2, 11.25x magnification.....	66
Figure 2.5	p97 ND1 (With Polyhistidine-tag), Well D7, 11.25x magnification .....	67
Figure 2.6	p97 ND1 (With Polyhistidine-tag), Well E4, 11.25x magnification.....	67
Figure 2.7	p97 Full Length, Well F1, 11.25x magnification .....	68
Figure 2.8	p97 Full Length, Well F2, 11.25x magnification .....	68
Figure 2.9	p97 Full Length, Well F5, 11.25x magnification .....	69

Figure 2.10	p97 Full Length, Well F6, 11.25x magnification .....	69
Figure 2.11	p97 Full Length, Well F7, 11.25x magnification .....	70
Figure 2.12	p97 Full Length, Well F8, 11.25x magnification .....	70
Figure 3.1	p97 ND1/CSELENOS Complex, Screen 2, Well C4, 11.25X Magnification .....	89
Figure 3.2	p97 ND1/CSELENOS Complex, Screen 2, Well B7, 11.25X Magnification .....	89
Figure 3.3	p97 ND1/CSELENOS Complex, Screen 3, Well B8, 11.25X Magnification .....	90
Figure 3.4	p97 ND1/CSELENOS Complex, Screen 3, Well D3, 11.25X Magnification .....	90
Figure 3.5	p97 ND1/CSELENOS Complex, Screen 3, Well D5, 11.25X Magnification .....	91
Figure 3.6	p97 ND1/CSELENOS Complex, Screen 3, Well E1, 11.25X Magnification .....	91
Table 4.1	V29Sem Plate Morphology (45% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.5, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, 2 mM TCEP).....	102
Figure 4.2	V29Sem Needle Morphology (45% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.5, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, 2 mM TCEP).....	102
Figure 4.3	V29Sem (46% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.5, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, 2 mM TCEP).....	103
Figure 4.4	I6Sem (46% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.7, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent).....	103
Figure 4.5	I6Sem (48% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.7, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, 1 mM TCEP).....	104

Figure 4.6	L5Sem (46% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.6, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent).....	104
Figure 4.7	L5Sem (46% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.9, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent).....	105
Figure 4.8	L5Sem (47% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.6, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent).....	105
Figure 4.9	Representative Image of A34Sem (47% and 49% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.9, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent .....	106

## **ABSTRACT**

The ATPase protein p97 plays an essential role in the healthy function of living cells. When defective proteins are detected by the cell, the endoplasmic reticulum-associated protein degradation pathway aids in the transport and ultimate degradation of these misfolded proteins before they can aggregate and cause problems for the host. P97 is a key player in this pathway by serving as the main source of energy by which these proteins are transported for destruction. Because p97 plays such an important role in the transport of defective proteins to maintain a healthy cell, it is important to not only study the protein itself, but also understand its relationship and binding with other proteins in the pathway. One such binding partner is Selenoprotein S, a protein that is responsible for recruiting p97 to the pathway. In this work, protein crystallography was utilized in an attempt to study the p97 and Selenoprotein S binding complex to ultimately solve the complex crystal structure. Outlined are preliminary successful crystallization conditions for p97 on its own and in complex with Selenoprotein S. These conditions can be used as starting points for further refinement to produce high quality protein crystals for X-Ray analysis.

# Chapter 1

## INTRODUCTION

### 1.1 Introduction to p97

In the endoplasmic reticulum (ER), ER-associated protein degradation (ERAD) is a cellular pathway that identifies and targets misfolded proteins and aids in the degradation of said proteins by the proteasome<sup>1-3</sup>. Defective proteins, and the aggregation thereof, have been linked to several disorders and diseases, making the ERAD pathway, and the various proteins associated with it, essential for the healthy function of the cell<sup>1</sup>. Once the misfolded proteins are recognized in the ER, they are transported to the cytosol where the ubiquitination of the proteins occurs through the ubiquitin–proteasome system (UPS)<sup>2,19</sup>.

The transportation of the misfolded proteins from the ER to the cytosol occurs through a transmembrane channel. A multitude of multi-spanning ER membrane proteins have been hypothesized to make up this channel, with evidence pointing to Derlin-1, -2, and -3 as the primary proteins<sup>19,21</sup>. While there are dozens of proteins that make up the ERAD pathway, p97 is of particular interest as a key protein in the transport of the defective proteins through this channel, along with the p97 recruiter protein Selenoprotein S<sup>19</sup>.

The AAA (ATPase associated with a variety of cellular activities) ATPase, p97, plays an essential role in the degradation pathway by acting as the main source of energy by which the misfolded proteins are transported from the ER to the cytoplasm<sup>3</sup>. Along

with its important function in the ERAD pathway, p97 plays a role in ubiquitin fusion degradation (a virtual route of the ubiquitin-proteasome pathway)<sup>4,81</sup> autophagy, transcriptional control, homotopic membrane fusion, and cell cycle regulation<sup>4</sup>. Because of p97's critical role in these biological processes, defects and mutations in p97 and its cofactors have been linked to several diseases, including inclusion body myopathy, Paget's disease of the bone and frontotemporal dementia (IBMPFD), as well as autosomal-dominantly inherited amyotrophic lateral sclerosis (ALS), a neurodegenerative disease characterized by motor neuron degeneration<sup>4-5</sup>.

Initial crystallography results for p97 show that it has a homo-hexameric structure, and each subunit is made up of an N domain, two hexameric ATPase domains (D1 and D2), and an unstructured C-terminus (Figure 1.1, Figure 1.2). The overall architecture of p97 is one of two stacked hexameric rings formed by the D1 (Figure 1.3) and D2 domains with the N domain on the outside of the D1 ring<sup>5-6</sup>. The N-terminal domain is needed for the interaction of the p97 with adaptor proteins, while the D1 and D2 AAA domains are necessary for the ATP hydrolysis and communication of the nucleotide state throughout the protein<sup>5-8</sup>.

ATP hydrolysis by p97 provides the energy for protein dislocation by the ERAD complex. Like other members of the AAA family, p97 utilizes the ATP hydrolysis cycle to produce conformational changes in substrate molecules<sup>3</sup>. Studies have found that p97 flips between two major conformations during this cycle: the preactivated (apo and ADP) and activated (ATP) states<sup>8</sup>. The six p97/VCP protomers form a pore, and there are two loops at the D2 end of p97 that line this pore. During ATP hydrolysis, both loops undergo conformation changes that coincide with the preactivated and activated states<sup>8-9</sup>.

A large part of what defines the biological processes in which p97 participates are the variety of cofactors that bind to the molecule<sup>10</sup>. These cofactors can typically be categorized as substrate-recruiting or substrate-processing cofactors<sup>11</sup>. The major factors in the substrate-recruiting category determine how p97 will participate in the various cellular pathways and then bind to it in a mutually exclusive fashion. Additional substrate-recruiting cofactors then can ‘fine tune’ interactions within the given pathway<sup>4</sup>. Substrate-processing cofactors usually are associated with enzymatic activity, and primarily are those playing a role in ERAD and UFD where they help regulate the fate of the p97 substrates<sup>4,10,13</sup>.

To date, there have been at least seven distinct p97-binding domains categorized: the UBX-domain, the PUB-domain, the UBX-like domain, the PUL-domain, the BS1/Shp box, the VBM domain, and the VIM domain. Several proteins that bind to p97 have already been characterized to fall into these categories<sup>13</sup>. One such protein is Selenoprotein S (SELENOS), which is responsible for recruiting p97 to the ERAD complex. Understanding its binding to p97 is the primary focus of this portion of work. These discoveries will not only help in better understanding how cofactors of p97 bind, but also shed light on the unique role of SELENOS as an essential p97 binding partner and contributor to the proper functionality of the ERAD pathway.

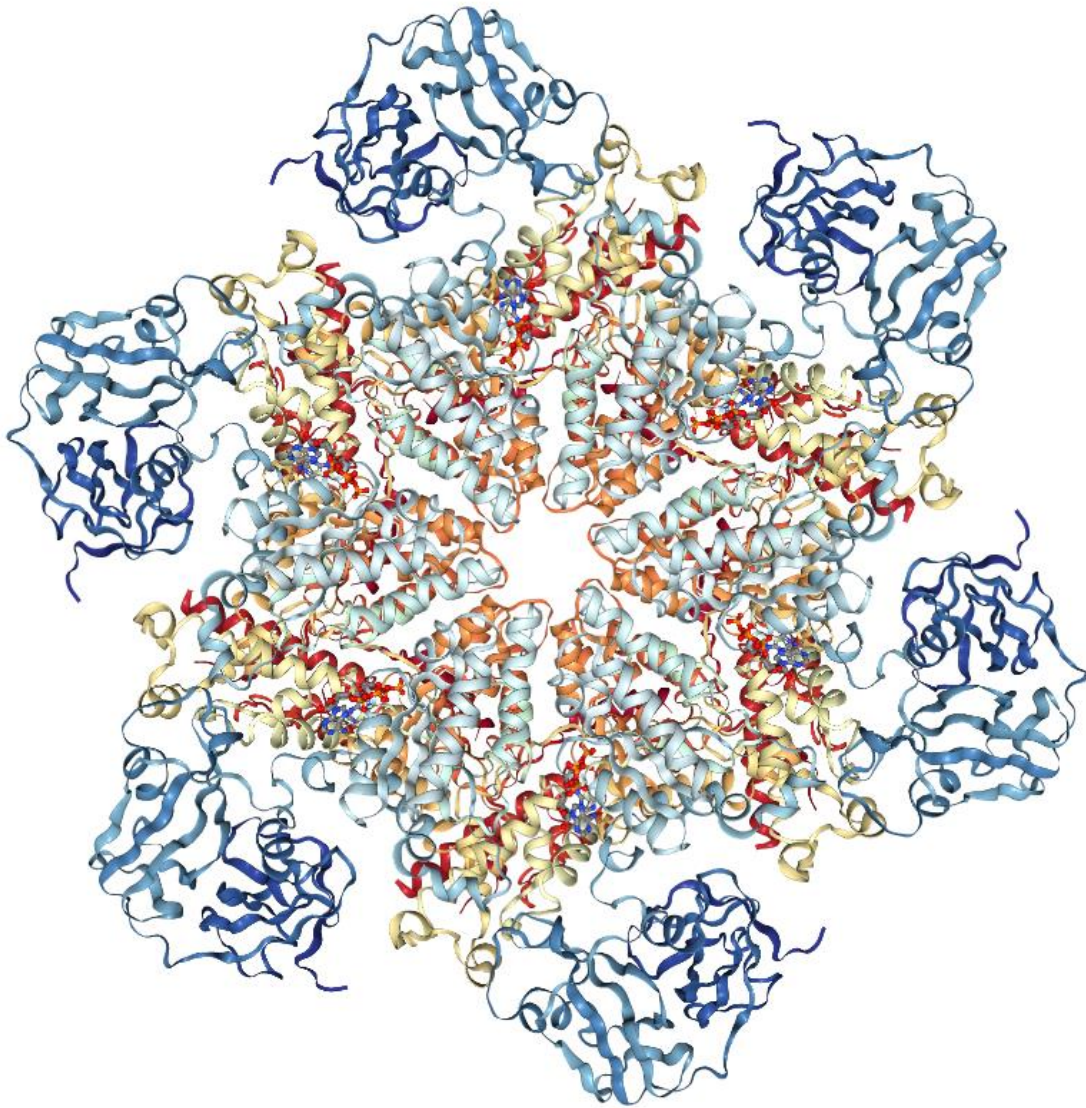


Figure 1.1: Structure of hexameric p97 wild type in complex with ADP. PDB ID: 3CF3.

Used with permission by Davies *et al*<sup>83</sup>. Image created using NGL viewer<sup>84</sup>.

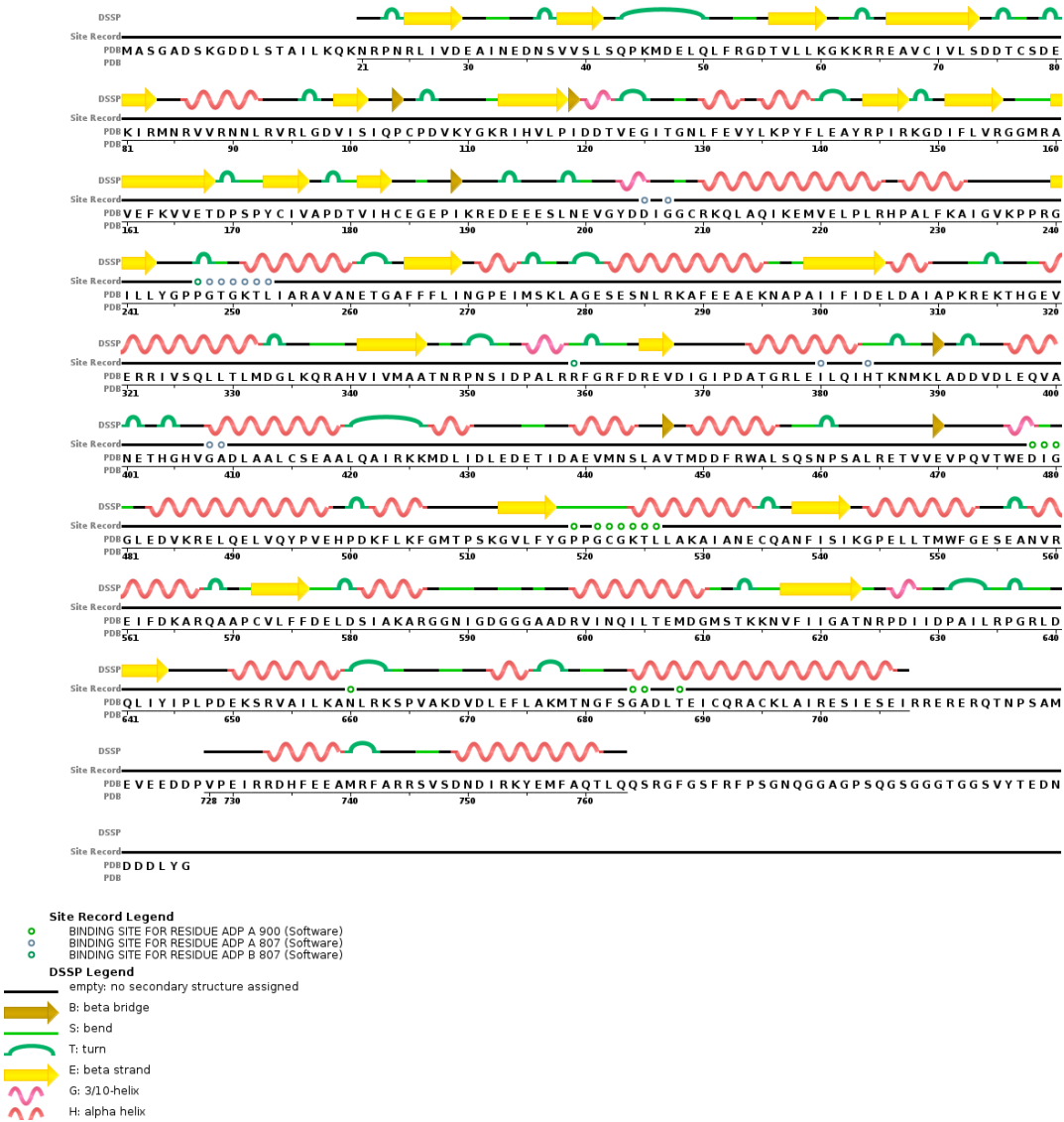


Figure 1.2: Sequence and secondary structure for p97. PDB ID: 3CF3. Used with permission by Davies et al<sup>83,87</sup>.

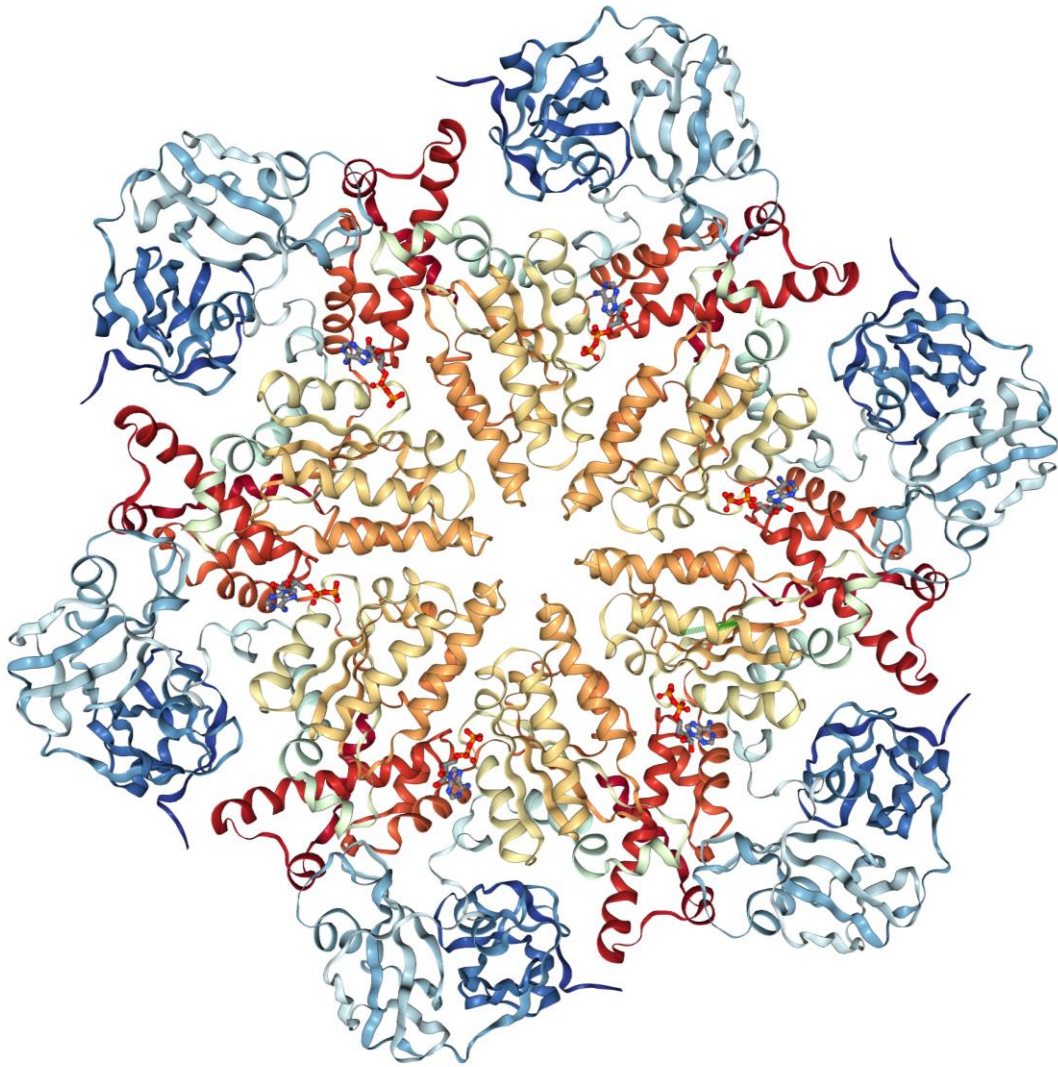


Figure 1.3: Structure of hexameric p97 ND1 (residues 1-480) wild type in complex with ADP. PDB ID: 5DYI. Used with permission by Tang *et al*<sup>85</sup>. Image created using NGL viewer<sup>86</sup>.

## 1.2 Introduction to Selenoprotein S

The subset of selenium containing proteins are known as selenoproteins, and this group is made up of twenty-five known selenoprotein genes in the human genome<sup>14-15</sup> with several thousand selenoproteins so far found in nature<sup>25</sup>. Selenium (Sec) is the 21<sup>st</sup> naturally occurring amino acid, and is analogous to the more commonly found cysteine residue, differing only by the replacement of the sulfur with a selenium atom<sup>25</sup>. Because Sec is encoded by the same codon as the stop codon, UGA, decoding for Sec requires a specific stem-loop structure in the 3'-UTR of eukaryotic selenoprotein mRNAs, designated the selenocysteine insertion sequence (SECIS) element, and a machinery to synthesize Sec on its own tRNA and insert it into proteins in response to the UGA codon<sup>19</sup>. Despite this high energy cost to synthesize Sec, it is believed that its continued incorporation is due to its lower pKa and stronger nucleophilicity than cysteine, which in turn has shown higher reactivity when Sec is utilized by enzymes<sup>19,24-25</sup>.

Selenoproteins themselves have a wide variety of structural motifs and are involved in various biological processes, particularly redox functions including glutathione peroxidase and thioredoxin reductase<sup>21</sup>. While the function of some selenoproteins is known, many are still unknown or poorly characterized, despite the fact that the physiological impacts of dysregulation have been studied<sup>14</sup>.

As previously mentioned, one of the known p97 binding partners is SELENOS. SELENOS, also known as SEPSI, VIMP, Tanis, and Sels, has been reported to be involved in the translocation of proteins from the ER to the cytosol, as well as playing a role in inflammatory responses and maintenance of ER homeostasis<sup>14-18</sup>. For the purposes of these experiments, the specific role of SELENOS as a p97 recruiter to the ERAD complex is being explored.

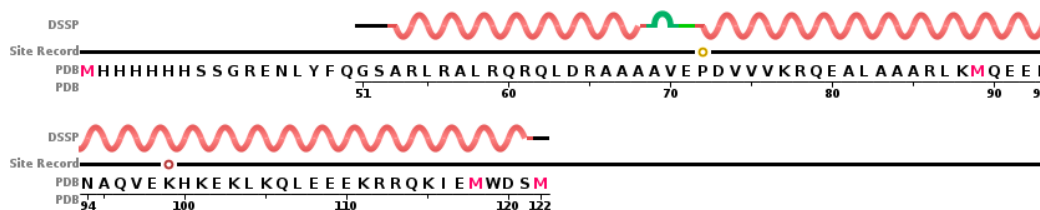
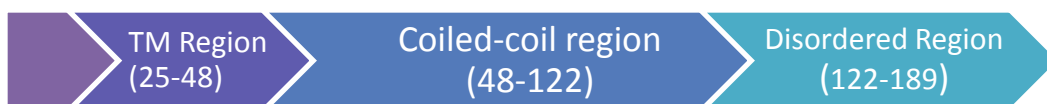
SELENOS itself is a 189 amino acid long protein (Figure 1.4) that contains one transmembrane domain in the N-terminus segment. It can be divided into three regions: the N-terminal ER luminal segment, the transmembrane domain encompassing residues 26-48, and the longer cytosolic tail with the Sec in the C-terminal penultimate position<sup>20,23</sup>. This tail contains a coiled coil domain that has been hypothesized to be responsible for binding other proteins, including p97, or SELENOS dimerization<sup>14,18-20</sup>. There is a well conserved, yet intrinsically disordered region in the C terminus that contains the active site residues of SELENOS that have redox properties<sup>23</sup>. The Sec in the active site of selenoproteins has been linked to redox catalysis<sup>16</sup>, and in SELENOS, the Cys<sup>174</sup> and Sec<sup>188</sup> form a selenenylsulfide bond and have been linked to the reductase and peroxidase activity *in vitro*<sup>24-25</sup>. Recent work exhibited that the Sec<sup>188</sup> residue is the peroxidase residue due to the inactivity of U188S and U188Δ mutants, where the U188Δ translation terminates after residue 187. This also led to the conclusion that Cys<sup>174</sup> is the resolving residue<sup>25</sup>. It was found that selenium was in fact required for the protein's reductase and peroxidase activity, with the SELENOS cysteine mutants showing no activity<sup>25</sup>. This poses additional questions as to the unique enzymatic nature of SELENOS, since many other selenoprotein cysteine mutants only show decreases in enzymatic activity as opposed to no activity whatsoever<sup>26</sup>.

Because of the unique nature of the SELENOS recruitment of p97, studying its interaction with p97 is essential in understanding its role in the ERAD complex. These sets of experiments aim to study the interactions between SELENOS and p97, most importantly by obtaining structural information through co-crystallization of p97 and SELENOS, particularly the c-terminal end of SELENOS. Further explanations regarding

known p97 complexes as well as successful crystallization conditions for p97 will be discussed in subsequent sections.

```

      10      20      30      40      50
MERQEESLSA RPALETEGLR FLHTTVGSLL ATYGWYIVFS CILLYVWFQK
      60      70      80      90     100
LSARLRALRQ RQLDRAAAAV EPDVVVKRQE ALAAARLKMQ EELNAQVEKH
      110     120     130     140     150
KEKLGQLEEE KRRQKIEMWD SMQEGKSYKG NAKKPQEEDS PGPSTSSVLK
      160     170     180
RKSDRKPLRG GGYNPLSGEG GGACSWRPGR RGPSSGGUG
  
```



**Site Record Legend**

- BINDING SITE FOR RESIDUE GOL A 3 (Software)
- BINDING SITE FOR RESIDUE CL A 1 (Software)

**DSSP Legend**

- empty: no secondary structure assigned
- S: bend
- T: turn
- H: alpha helix

Figure 1.4a: The Amino Acid sequence of SELENOS

Figure 1.4b: Domain structure of SELENOS. Amino acids 25-48 are the transmembrane region, 48-122 are the structured coiled-coil region, and amino acids 122-189 make up the disordered binding region

Figure 1.4c: SELENOS Domain Structure of coiled coiled region previously crystallized. PDB ID: 2QF2. Used with permission by Walker *et al*<sup>90</sup>.

### 1.3 Introduction to p97 Complexes

Evidence suggests that it is through binding to various adaptor proteins that p97 is directed to different cellular processes. Apart from SELENOS, several other adaptors have been identified and crystal structures of their complexes have been reported<sup>10,30-41</sup>. These include the UBX (ubiquitin regulatory X) domain of p47<sup>41</sup>, the UBX domain of Fas-associated factor 1 (FAF1)<sup>34,38</sup>, the ubiquitin regulatory X-like (UBXL) domain of the Ovarian Tumor Domain-containing Protein 1 (OTU1)<sup>37</sup>, the SHP motif of ubiquitin fusion degradation protein 1 (UFD1)<sup>39</sup>, the SHP motif of derlin-1<sup>36</sup>, the VCP-interacting motif (VIM) of the ubiquitin E3 ligase gp78<sup>30</sup>, and the VCP-binding motif (VBM) of rhomboid protease RHBDL4<sup>32</sup>. These UBX and UBXL domains, as well as the VBM, VIM, and SHP interacting motifs are shared across p97 adaptors, which means that the proteins that share similar binding motifs could also have similar binding mechanisms<sup>10</sup>.

SELENOS is also known as VIMP, or VIM protein, which means that it binds with p97 via the VCP interacting motif. The VIM was first discovered in the mammalian ERAD ubiquitin ligase gp78, recruiting p97 to the ER membrane<sup>42</sup>. Another protein, the small VCP-inhibiting protein (SVIP), was found to also interact with the p97 through a VIM as well, it's purpose to serve as a negative regulator of ERAD<sup>43</sup>. Thus, the original definition of the VIM sequence became based off of the 30 residue-long binding sites of these two p97 interacting proteins<sup>10,45</sup>. However, more recently, Buchberger *et al.* proposed a minimal VIM consensus sequence, RX<sub>5</sub>AAX<sub>2</sub><sup>42</sup>. Figure 1.5 is taken from that work, and shows the minimal VIM consensus sequence through comparison of well-known VIM containing proteins, including SELENOS (VIMP).

SELENOS has a predictable VIM sequence between residues 78 and 88, however, VIM deleted mutants were still shown to interact with p97 and recruit it to the ER. This same study found that deletions of the proline residues in positions 178 and 183 inhibit

binding, which indicates that this particular region of SELENOS plays a critical role<sup>45</sup>. These proline residues are within the disordered region of SELENOS, where the Cys<sup>174</sup> and Sec<sup>188</sup> residues form a selenylsulfide bond<sup>46</sup>. It is believed the proline residues here play a key role in the structural integrity of this region, and since mutations to these proline residues lead to a loss of binding activity *in vivo*, it points to the selenosulfide bond as playing a crucial role in the p97/SELENOS complex<sup>47-48</sup>.

This work intends to better understand the binding of p97 to SELENOS, and, ultimately, solve the crystal structure of the complex. Because it is known that the p97 binding region is located in the ND1 domain and that the disordered C terminus of SELENOS is critical for binding, initial experiments focused on those regions. Crystallization conditions were initially based off published conditions for various p97 complexes to avoid the tedious task of blind screens that could take months. We were able to find supposed ideal conditions for binding of the complex, as well as successful crystallization reservoir solutions and drop ratios which grew protein crystals.

gp78	AMFR2	HUMAN	VTI <b>RRR</b> ML <b>AAA</b> A <b>ERR</b> L <b>OK</b>
	Q28GC0	XENTR	TTV <b>RRR</b> ML <b>AAA</b> A <b>ERR</b> I <b>QM</b>
	O6PFU8	DANRE	VTI <b>RRR</b> RL <b>AAA</b> A <b>ERR</b> IL <b>R</b>
	C3YFB3	BRAFL	VSV <b>RRR</b> IL <b>AAA</b> A <b>ERR</b> IL <b>O</b>
	Q55J61	CRYNE	DNDV <b>RRQ</b> V <b>AD</b> <b>AA</b> L <b>RR</b> L <b>GG</b>
SVIP	SVIP	HUMAN	I <b>EE</b> K <b>R</b> AK <b>L</b> A <b>E</b> <b>AA</b> E <b>RR</b> <b>O</b> K <b>E</b>
	SVIP	MOUSE	E <b>EE</b> K <b>R</b> E <b>K</b> L <b>A</b> E <b>AA</b> E <b>RR</b> <b>O</b> K <b>E</b>
	B5X187	SALSA	E <b>TR</b> <b>RR</b> R <b>OL</b> A <b>E</b> <b>AA</b> E <b>K</b> <b>R</b> <b>O</b> K <b>E</b>
	Q7PUC6	ANOGA	K <b>EV</b> <b>RR</b> Q <b>OO</b> R <b>E</b> <b>AA</b> E <b>RR</b> <b>I</b> E
	Q8MZ33	DROME	E <b>ERR</b> Q <b>Q</b> L <b>D</b> <b>AA</b> K <b>K</b> <b>R</b> <b>O</b> E
VIMP	SELS	HUMAN	V <b>V</b> V <b>K</b> R <b>O</b> E <b>A</b> L <b>AA</b> R <b>I</b> K <b>M</b> <b>O</b> E
	SELS	XENTR	E <b>I</b> V <b>R</b> R <b>O</b> E <b>A</b> V <b>I</b> A <b>A</b> R <b>I</b> <b>R</b> <b>M</b> <b>O</b> E
	B7PTY9	IXOSC	V <b>V</b> V <b>O</b> R <b>O</b> V <b>E</b> M <b>D</b> A <b>ARR</b> <b>R</b> <b>M</b> <b>O</b> A
	B3S7J3	TRIAD	E <b>A</b> L <b>A</b> R <b>NE</b> K <b>M</b> L <b>AA</b> R <b>A</b> K <b>L</b> <b>O</b> E
	A9V266	MONBE	P <b>V</b> Q <b>A</b> S <b>Q</b> E <b>S</b> V <b>L</b> A <b>A</b> R <b>A</b> R <b>O</b> <b>Q</b> E
ANKZF1 (ZNF744)	ANKZ1	HUMAN	A <b>L</b> S <b>D</b> R <b>E</b> K <b>R</b> A <b>L</b> <b>AA</b> E <b>RR</b> L <b>AA</b>
	Q56X22	ARATH	M <b>A</b> A <b>O</b> R <b>E</b> K <b>R</b> A <b>AA</b> E <b>RR</b> <b>M</b> <b>A</b> S
	YOL4	SCHPO	M <b>R</b> I <b>E</b> R <b>E</b> K <b>R</b> A <b>AA</b> M <b>K</b> <b>R</b> <b>M</b> <b>O</b> T
	YD049	YEAST	R <b>R</b> L <b>M</b> R <b>E</b> <b>O</b> R <b>A</b> R <b>AA</b> E <b>RR</b> <b>M</b> <b>K</b> <b>K</b>
	C4YKY5	CANAL	M <b>R</b> V <b>L</b> R <b>E</b> <b>O</b> R <b>A</b> R <b>AA</b> E <b>ARR</b> <b>M</b> <b>K</b> <b>R</b>
Wss1	WSS1	YEAST	G <b>N</b> S <b>P</b> R <b>E</b> L <b>AA</b> F <b>AA</b> E <b>RR</b> <b>R</b> <b>Y</b> <b>R</b> <b>D</b>
	Q8X095	NEUCR	G <b>L</b> T <b>R</b> R <b>E</b> I <b>L</b> A <b>K</b> A <b>E</b> E <b>RR</b> <b>M</b> <b>K</b> <b>K</b>
	C8VTV9	EMENI	G <b>G</b> G <b>M</b> R <b>E</b> V <b>L</b> A <b>R</b> A <b>E</b> E <b>ER</b> <b>R</b> <b>A</b> <b>R</b> <b>R</b>
	Q4P739	USTMA	T <b>A</b> S <b>A</b> R <b>E</b> L <b>R</b> A <b>Q</b> A <b>L</b> L <b>R</b> L <b>AA</b>
	WSS1	SCHPO	S <b>M</b> R <b>G</b> R <b>E</b> A <b>R</b> I <b>AA</b> L <b>L</b> R <b>V</b> <b>D</b> <b>N</b>
Ygl108c	YGK8	YEAST	K <b>V</b> S <b>P</b> K <b>E</b> A <b>A</b> R <b>L</b> <b>AA</b> E <b>K</b> <b>R</b> <b>F</b> <b>O</b> E
	O6FU03	CANGA	K <b>I</b> S <b>A</b> K <b>E</b> A <b>A</b> R <b>L</b> <b>AA</b> E <b>K</b> <b>R</b> <b>Y</b> <b>O</b> <b>D</b>
	O6CKI2	KLULA	T <b>L</b> S <b>P</b> R <b>E</b> A <b>A</b> R <b>V</b> A <b>E</b> <b>O</b> R <b>N</b> <b>A</b> <b>K</b>
	c.albicans		T <b>L</b> N <b>A</b> K <b>E</b> A <b>S</b> A <b>K</b> A <b>E</b> E <b>ER</b> L <b>Q</b> <b>K</b>
	c.tropicalis		P <b>M</b> N <b>A</b> R <b>E</b> A <b>S</b> A <b>R</b> A <b>E</b> E <b>RY</b> <b>K</b> <b>K</b>
UBXD1	UBXN6	HUMAN	P <b>T</b> N <b>E</b> A <b>O</b> M <b>AA</b> A <b>A</b> L <b>A</b> R <b>I</b> E <b>O</b>
	Q1LYI0	DANRE	P <b>T</b> A <b>G</b> A <b>O</b> M <b>A</b> G <b>AA</b> L <b>A</b> R <b>I</b> E <b>O</b>
	B4LJ58	DROVI	I <b>S</b> K <b>E</b> A <b>R</b> A <b>AA</b> N <b>A</b> L <b>A</b> R <b>I</b> E <b>R</b>
	Q7PP13	ANOGA	I <b>S</b> A <b>E</b> A <b>K</b> A <b>AA</b> N <b>A</b> L <b>A</b> R <b>E</b> <b>G</b>
	B3RNJ8	TRIAD	P <b>T</b> S <b>D</b> A <b>S</b> R <b>A</b> G <b>Q</b> A <b>A</b> L <b>R</b> <b>L</b> <b>D</b> <b>N</b>
ZFAND2b	ZFN2B	HUMAN	E <b>G</b> H <b>P</b> T <b>S</b> R <b>A</b> G <b>L</b> <b>AA</b> T <b>S</b> R <b>A</b> <b>O</b> A
	Q0P3R6	XENLA	K <b>S</b> A <b>P</b> I <b>S</b> R <b>A</b> G <b>H</b> V <b>A</b> L <b>L</b> R <b>S</b> <b>O</b> A
	C3XWL7	BRAFL	T <b>G</b> R <b>A</b> V <b>S</b> N <b>A</b> G <b>AA</b> T <b>L</b> R <b>N</b> <b>K</b> <b>G</b>
	O7SXI4	DANRE	D <b>N</b> K <b>P</b> V <b>S</b> K <b>S</b> G <b>H</b> A <b>A</b> L <b>M</b> R <b>A</b> <b>O</b> <b>G</b>
	Q5DEG5	SCHJA	N <b>S</b> R <b>K</b> L <b>S</b> A <b>A</b> G <b>T</b> <b>AA</b> L <b>Y</b> R <b>A</b> <b>I</b> <b>F</b>
consensus			<b>R</b> xxxxxx <b>AA</b> xx <b>R</b>

**Figure 1.5:** Minimal VIM consensus sequence taken from Buchberger *et al* and used with permission. The most highly conserved residues are boxed in red, minimal consensus sequence RX5AAX2R shown at the bottom. X refers to positions with no significant conservation<sup>42</sup>.

## 1.4 Introduction to GB1

Protein crystallography was not only utilized in the study of p97 and its cofactor SELENOS, but it was also used to look at another protein of interest, the immunoglobulin binding domain B1 of streptococcal protein G (GB1). While unrelated to the p97 work, the growth of GB1 crystals had its own application in using GB1 variants to build a  $^{77}\text{Se}$  NMR tensor library.

A wide variety of bacteria have surface proteins that have been shown to bind with high affinity to host proteins like immunoglobulin. One such protein, Protein G in group G *Streptococcus*, is a multi-domain surface protein that selectively binds to all subclasses of mammalian immunoglobulin G (IgG) in the Fc region, making it more selective than other bacterial agents, like staphylococcal protein A. Depending on the strain, there are two or three repeating domain units, B1, B2, and B3 that bind IgG, while another two or three bind to serum albumin. It is hypothesized that these binding interactions are what aid the bacteria in eluding the immune response of the host<sup>49-50</sup>.

GB1, has long served as a model system for protein design and folding. While small (56 residues) and containing only one domain, GB1 has a very stable fold, and contains one  $\alpha$ -helix, a four strand  $\beta$ -sheet, and two hairpin loops(Figure 1.6)<sup>50-51</sup>. Because of this, it is often used for studies of protein stability and folding. The structure of GB1 has been studied extensively by X-ray diffraction and solid-state NMR<sup>52-55</sup>, making information about its structure and stability readily available.

Current studies in the Rozovsky group used GB1 as a model protein for introducing selenomethionine (Sem) and using it as an NMR probe for local environments due to the  $\frac{1}{2}$  spin of the  $^{77}\text{Se}$  nucleus. Six Sem GB1 variants were expressed and purified, V29Sem, I6Sem, L5Sem, V39Sem, V34Sem, and V54Sem, with the intent of using solid-state NMR to establish the beginnings of a library of  $^{77}\text{Se}$  chemical shift

tensors for Sem by probing the various environments. In order to do this, it is important to also know the exact structures of each mutant to account for things like alternative side chain conformations. Solving the structures of these variants through X-ray crystallography will be cardinal for establishing the relationship between  $^{77}\text{Se}$  chemical shift and the nuclear environment.

Outlined in Chapter 4 are the crystallization conditions based off of published data and successful conditions achieved in-house for each of the GB1 variant.

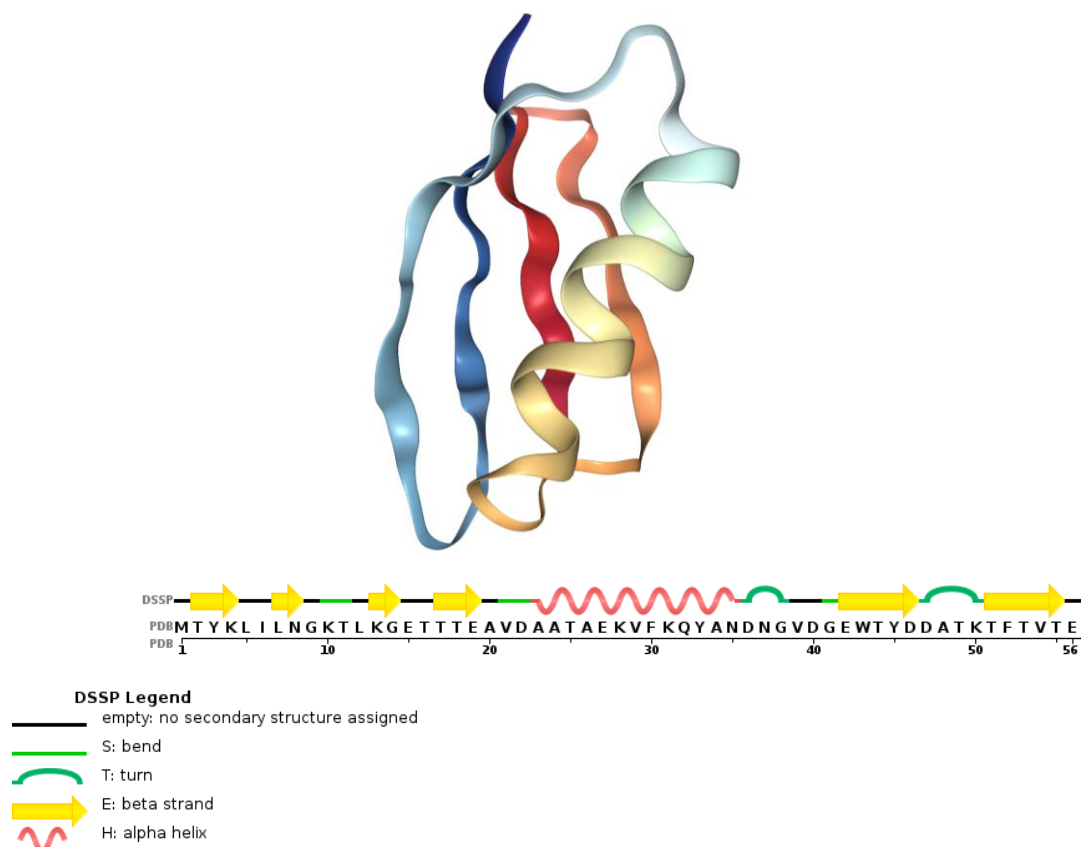


Figure 1.6: Structure and sequence of immunoglobulin binding domain B1 of streptococcal protein G (GB1) . PDB ID: 1GB1. Used with permission by Gronenborn *et al*<sup>88</sup>. Image created using NGL viewer<sup>89</sup>

## 1.5 Introduction to X-Ray Crystallography

### 1.5.1 Principles of X-Ray Crystallography

Over one hundred years ago, the ground work was laid for what would become modern macromolecular crystallography with the discovery of X-rays in 1895<sup>55-58</sup>. Nearly twenty years later, Max von Laue first showed that the wavelength of X-rays could be comparable to the interatomic distances of crystals<sup>59</sup>, and W.H. Bragg and W.L. Bragg truly pioneered the application of X-rays as tools for studying crystal structure. W.L. Bragg mathematically explained the Laue diffraction images through the equation now known as Bragg's Law,  $n\lambda = 2d\sin\theta$ , which describes the relationship between the angles of diffraction ( $\theta$ ), the wavelength of the X-rays ( $\lambda$ ) and the interplanar spacings ( $d$ ) in the crystal lattice<sup>55,60-65</sup>. These earned both Laue and the Braggs Nobel Prizes for their discoveries, and since then, over forty Nobel Prizes in the field of X-ray crystallography have been awarded<sup>55,66</sup>.

Similarly, breakthroughs in the crystallization of macromolecules also paved the way for modern protein crystallography. As far back as 1840, books can be found that describe crystals of earthworm hemoglobin<sup>66</sup>. By 1909, there were publications with over 600 hemoglobin crystals from 200 different organisms<sup>67</sup>, and the biggest breakthrough came in 1946 when James Sumner was the first to crystallize an enzyme, urease<sup>68</sup>.

However, it wasn't until 1960 that these two fields were able to be combined to yield the first 3-dimensional protein crystal structures of myoglobin and hemoglobin<sup>55,69</sup>. Since then, macromolecules of interest have been crystallized and their structures solved by X-ray diffraction. The Protein Data Bank, a data repository for structural models, now contains over 100,000 models, over 90% of which were solved using X-ray crystallography<sup>70-71</sup>.

Solving a molecular structure using crystallography is an often time consuming and difficult, multi-step approach. The first step involves the expression and purification of the protein(s) of interest. Protein purity is of the utmost importance in crystallography, as impurities can inhibit good crystal growth. Once protein of high enough purity has been obtained, the next step is the actual growing of the protein crystal. This is perhaps the most tedious step, with ideal conditions taking months or years to identify. Protein crystals are grown by first dissolving the protein in a buffer solution and then slowly precipitating the protein out of solution<sup>71-72</sup>.

Precipitation of a protein requires a careful balance of protein concentration, precipitants, and growing environments to find conditions that do not denature the protein and cause it to crash out of solution in a non-crystalline form, but also bring the concentration of the protein above the saturation limit. The most common method of growing protein crystals is vapor diffusion, and is discussed at length in the next section.

Once ideal conditions have been met and crystals believed to be of diffraction quality are grown, the crystals are typically harvested by mounting on special loops, flash freezing in liquid nitrogen for preservation, and then submitted for X-ray analysis. X-ray diffraction data is collected by mounting the crystal between the source and detector. As the electrons in the molecules are being hit with X-rays, the rays are diffracted and form a unique pattern of reflections. There are obstacles to overcome in turning these diffraction patterns into usable data to build an electron density map, and these are discussed in more detail in section 1.5.3. However, once the data has been phased to give an electron density map, there are several programs and software suites that aid in the model building, refinement, and ultimate interpretation of the final protein model<sup>70-72</sup>.

### 1.5.2 Vapor Diffusion Crystallization

To grow protein crystals, the most common approach is to prepare the protein in an aqueous buffer and then mix with some sort of precipitant(s). The precipitant concentration should be just below what is needed to physically precipitate the protein of interest, and over time, water is evaporated, raising both the concentration of precipitant and protein in solution. This should result in the precipitation of the protein, ideally in crystalline form<sup>70,72</sup>.

The overall principle of crystallization remains the same despite the preferred method of water evaporation. The ultimate goal is to bring the solution to a supersaturated state where crystal growth can occur<sup>73</sup>. Figure 1.7, taken from Chayen *et al* paper on protein crystallography<sup>82</sup>, depicts a phase diagram for crystallization with the concentration of the protein as a function of precipitant. The solubility curve in the diagram is meant to represent the “line” between undersaturation of the protein in solution as a function of precipitant and the supersaturation region where growth can occur. Nucleation (the initial formation of molecular clusters occurring in higher precipitant conditions) and growth (no additional nucleation and precipitant conditions at ideal concentrations) can occur in the precipitation zone, and in ideal circumstances, continue to grow in the metastable zone. However, being pushed too far into the precipitation zone can lead to aggregates and amorphous precipitation as opposed to crystal formation<sup>70-73,82</sup>.

Also highlighted in the figure are the different crystallization methods and how these various approaches take their own unique pathway to crystal formation through the precipitation and metastable zones. The most popular approach, and the technique highlighted in this study, is vapor diffusion, where the solution of the protein and precipitant equilibrates in a closed system against a larger reservoir of precipitant at a given concentration. Volatiles will be exchanged and the drop will evaporate until

equilibrium is reached. This puts the protein concentration just barely over into the supersaturation zone for initial nucleation, and then as the system reaches equilibrium, growth can occur in the metastable zone<sup>70-73</sup>. Sitting drop and hanging drop vapor diffusion are the most common methods, hanging drop being the approach used in this study.

Hanging drop vapor diffusion is typically carried out in multi-well plates, each well filled with a different reservoir solution to allow for screening of multiple conditions at a time. The reservoir solution is made up of a buffer at a given pH and salt concentration, any additives to stabilize the protein, the precipitant(s), and in some cases, the cryoprotectant to later allow for flash freezing. On a coverslip, volumes of the protein solution and reservoir solution are mixed together prior to the coverslip being inverted over the well, the drop then suspended over the reservoir to equilibrate<sup>70-73</sup>.

Varying the volume of protein and reservoir solution in the drop is just one of the ways to optimize growth conditions. In doing so, the starting and ending concentrations of protein in the hanging drop vary and take differing routes to supersaturation. By definition, the drop ratio is the volume of protein compared to volume of precipitant in the hanging drop. Figure 1.8 gives three examples of how varying the drop ratios impacts the starting and final concentrations of protein and starting concentration of reagents. Because the system reaches an equilibrium, the ending reagent concentrations are always the same. These variations allow for testing a multitude of initial concentrations without having to change the concentration of the protein stock solution. It also covers a wider range of supersaturation conditions and equilibrium paths<sup>74-75</sup>. This method has proven particularly helpful in cases where amorphous precipitant was seen almost immediately after forming the drop. By diluting the protein in the drop with reservoir buffer, a lower

initial and final concentration of protein was achieved, and the path to supersaturation was slower to better allow for crystal growth.

The time consuming but critical component in successful crystallizations are determining the specific conditions for successful crystal growth. Ionic strength, pH, concentration of protein, concentration of precipitants, incubation temperature, stabilizers, additives, ligands, inhibitors, reducing or oxidizing conditions, and rate of equilibrium are just some of the variables that can be the difference between an empty drop, a fully formed protein crystal, or amorphous precipitant. Differences in these components do not need to be drastic, either. A few pH points, degrees in incubation temperature, or precipitant concentrations off by just a few percentage points can make enough of an impact<sup>70-73</sup>. It is still not possible to calculate these variables and they each need to be experimentally tested.

The process by which ideal conditions are found is commonly referred to as screening. Since there are any number of possible buffer solutions, pH ranges, precipitants, and incubation temperatures, this process can be lengthy. There are commercial screening kits available with pre-determined reservoir solutions to perform coarse screens, and there are robotic setups that can automatically lay trays in a fraction of the time it would take a human.

During the screening process, the trays are laid with the different conditions, incubated, and the drops are monitored regularly to track progress. Crystals can form in days or weeks, but it is imperative to track regularly. Even if crystals do not form, there are clues in the drops that can point to better conditions. A clear droplet typically means that the protein in solution is still under saturated, and an increase in the protein concentration and/or precipitants is needed. On the other hand, amorphous protein

precipitating out can mean that your initial concentration of protein and/or precipitant is too high, causing your protein to denature. Other clues like oiling or phase separation are indicators of proteins that are just below supersaturation. In some cases, protein crystals can still grow out of these droplets, or at the very least, be counted as conditions from which finer screens can be developed.

There are circumstances where crystals might be observed that are actually salt crystals, not protein. Salt crystals typically have a different morphology than protein, but if unsure, there are methods to test besides structural confirmation by X-ray diffraction. Salt crystals are more resilient than protein crystals, so if they crush easily, it is a good indicator that they are protein. Also, protein crystals have large solvent channels that salt crystals do not possess. If one introduces dye to the drop and it is absorbed by the crystal (after a few hours or days), then it is most likely protein. If the crystals are salt, then this could be a sign the salt concentration in the reservoir buffer is too high<sup>74</sup>.

When observables like phase separation, microcrystals, or crystals close to but not quite diffraction quality are seen, it is a sign that optimal crystallization conditions might be close. Screening around these conditions by varying pH, incubation temperature, and precipitant concentration often lead to full-size protein crystals<sup>70-73</sup>.

Because screening for ideal conditions is a time consuming process and there are any number of starting points, the best, first step is to research published conditions for similar proteins. In the cases of p97 and GB1, crystal structures of the wild type and other mutants or complexes were already reported, along with their crystallization conditions. These served as helpful starting points to develop our own screening conditions in-house.

Once diffraction quality crystals are grown, harvested, and frozen, they can be analyzed by X-ray diffraction, and the process of structural determination can begin.

### 1.5.3 Structure Solving and Data Processing

In going from solution state to crystalline, the molecules assume one or very few defined orientations. This ordered 3-dimensional array of molecules can further be thought of as divided into identical unit cells combining to form a lattice. By imagining the crystal as an array of stacked, repeating unit cells, each cell representing the minimal volume element representative of the crystal as a whole, what is obtained through crystallography is a pattern of the diffracted light as it interacts with the atoms of the crystal. This then can be translated as an image of the electron cloud surrounding the molecules in that average unit cell of the crystal<sup>73</sup>. Based on the size and shape of the unit cell, the scattering of the X-rays is enhanced in certain directions but not others. Therefore, the overall intensity of the diffracted X-ray is dependent upon the arrangement of the atoms in the unit cell, which corresponds to the crystal structure. The number of reflections is directly related to the size of the unit cell, and the absolute position of all atoms within the structure effects the intensities of the reflections. This means that the entire crystal must be modelled in order to solve the structure, not just portions of it<sup>70,72</sup>.

In the X-ray analysis of crystals, the waves diffracted by the atoms obey Bragg's law, which states that angles of diffracted beams from the crystal can be calculated by treating said diffraction as reflections from equivalent, parallel planes of atoms in the crystal<sup>73</sup>. The crystal contains multiple copies of the protein molecule, and the diffraction pattern represents the average reflections across the copies in the lattice<sup>70-73</sup>.

To collect the diffraction data, the sample is mounted in the path of the X-ray, and the diffracted light intensity is then measured as a function of the angle of diffraction,  $\theta$ . A typical dataset is made up of hundreds of diffraction images, each image collected after the crystal is rotated in small, angular increments. The intensity of the reflections is not

only a function of the crystal content, but also overall crystal quality, intensity of the beam, exposure time, and the total time the crystal is in the reflection condition<sup>70</sup>.

The final dataset of coordinates and intensities represent the dimensions and symmetries of the diffraction pattern, and therefore, the unit cell. Once this data has been obtained, the next step is to see how these values translate to the final molecular structure. Each reflection is assigned three coordinates, indices, in the theoretical three dimensional reciprocal space of the pattern. So not only are the intensities of each reflection measured, but also their positioning relative to the origin. In combining these different images, the distribution of the intensities of all equivalent reflections is gauged by residual factor,  $R_{merge}$ .  $R$  is often expressed as a fraction or percentage, and is the measure of agreement amongst the sets of values. Ideally, the lower the  $R$  value, the better the data set. The  $R_{merge}$  factor is based off of the averaging process, known as scaling and merging, and the end result is one set of reflection intensities per point, along with its associated margin of error across the multiple values per reflection. Multiple scans equate to multiple observations of the same reflection points, and allows for the possibility to discard outliers caused by things like instrument error<sup>70-72</sup>.

Once the reflections have been indexed to obtain their geometric arrangement, the intensities of each reflection integrated, and these values reduced, the next step is to construct the electron density map. Each reflection is defined by its amplitude and phase, and through Fourier transform, these values can be used to calculate the electron density. However, in crystallography, there is something referred to as the “phase problem” because the only data available is the amplitude, defined as the root square of the measured intensity. This phase problem is one of the biggest challenges to overcome in solving crystal structures, these phases now having to be solved indirectly. There are

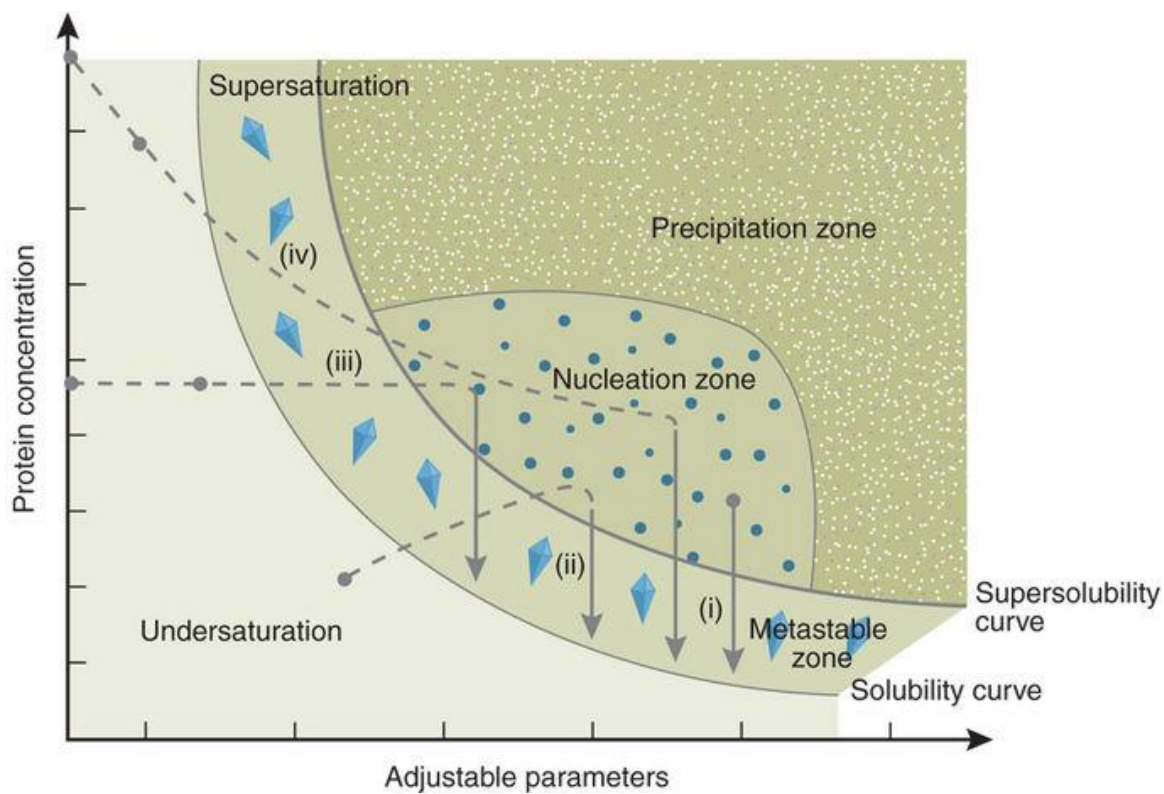
multiple indirect and direct methods to obtain the missing phases, perhaps the most common being molecular replacement. Here, a model of a similar protein with a known electron density (and therefore phase) is rotated and translated to fit the data from the new crystal. If this is successful, the amplitudes and phases from the model can be combined with your data. With the approximate phase data acquired, the electron density map can be calculated<sup>70-73,76-78</sup>.

With the electron density map at hand the structure of the protein can be refined. In the higher resolution data sets, things like secondary structure can be identified, and the backbone and some side chains can be assigned. Map fitting is where the atomic model of the protein is inserted into the density map based on the known sequence. The next steps typically involve manual refinement of the model compared against the density map, which includes placing the atoms in the correct locations within the electron density map. The ultimate goal of the refinement stage is to account for discrepancies in the density map. For example, because the data acquired is an “average” of unit cells, there are some cases where a reflection comprises multiple side chain conformations, and they will be reflected in the density map. Things like added water or ligands can also explain densities not accounted for by the protein alone. The more the model is refined through proper atomic placement so that it best fits the density map, the better the statistical parameters<sup>70-73,76-78</sup>.

Typically, this refinement step is done until acceptable refinement parameters are met. These values can vary based on a number of factors related to the quality of the data or how dependent the refinement was on prior knowledge of geometry as opposed to current experimental data. Finally, once the model has been refined well enough so that

the statistical parameters fall within the desired ranges, the structure is considered solved<sup>76-78</sup>.

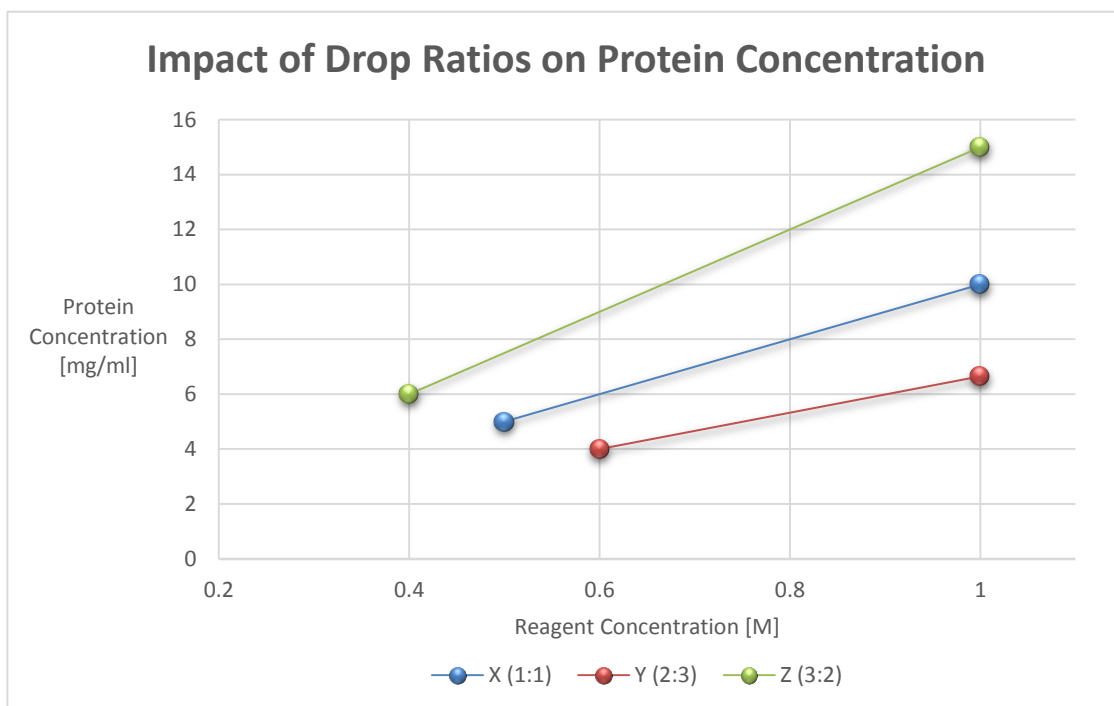
While the steps in processing the diffraction data appear to be daunting, there are a wide variety of computer programs available to solve structures. Software suites like CCP4 and PHENIX contain programs that can be used to automatically index and integrate the reflections, reduce the data and generate statistical parameters, solve the phase and electron density using molecular replacement, build the protein model from a preexisting amino acid sequence, and then refine and structurally analyze the model. Many of these steps are now fully automated, however, there is the opportunity for manual analysis of the data and structural refinement<sup>79-80</sup>.



**Figure 1.7:** Crystallization Phase Diagram. Taken with permission from Chayen, N.E., & Saridakis, E.

Protein Crystallization: From Purified Protein to Diffraction Quality Crystal. *Nature Methods*, 5, 157-153

(2008).



**Figure 1.8:** Graphical Representation of the concentration relationship of protein and reagent in varying drop ratios (1:1, 2:3: and 3:2 protein to reagent drop ratios).

## 1.6 Conclusion

Start to finish, the process to solve a macromolecular structure based on X-ray diffraction can be long and frustrating. From crystal growing to model refining, there is no shortage of opportunities to learn from trial and error. However, the information that can be learned from a fully realized structural model is invaluable. The work summarized in this thesis lays the foundation for the crystallization of AAA ATPase p97 in complex with one of its recruiter proteins, Selenoprotein S. Initial crystallization conditions for p97 ND1, full length p97, and p97 ND1 in complex with the Selenoprotein S C terminal peptide (CSELENOS) are reported, with the ultimate goal to eventually obtain diffraction quality crystals of the p97 ND1/CSELENOS complex and send for X-ray diffraction. In

solving the structure, we hope to better understand the mechanism by which p97 is recruited to the ERAD pathway by SELENOS, an essential function for a healthy cell.

Protein crystallography was also applied to the study of various Sem containing GB1 variants. These variants can serve as spectroscopic probes of local environment, and a structural model will aid in the creation of a library for  $^{77}\text{Se}$  chemical shift tensors. A fine screen of crystallization conditions applied to all six reported GB1 variants is reported, along with crystal candidates sent for analysis by X-ray diffraction.

## REFERENCES

1. Meusser, B., Hirsch, C., Jarosch, E., & Sommer, T. ERAD: The Long Road to Destruction. *Nature Cell Biology*. **7**, 766-772 (2005).
2. Clechanover, A. The Ubiquitin-Proteasome Proteolytic Pathway. *Cell*. **79**. 13-21 (1994).
3. Ye, Y., Meyer, H. H., Rapoport, T.A. The AAA ATPase Cdc48/p97 and its Partners Transport Proteins from the ER into the Cytosol. *Nature*. **414**. 652- 656 (2001).
4. Hanzelmann, P., & Schindelin, H. The Structural and Functional Basis of the p97/VCP-Interacting Motif (VIM) Mutually Exclusive Binding of Cofactors to the n-terminal Domain of p97. *J. Biol. Chem*. **286**, 28679-38690 (2011).
5. DeLaBarre, B., & Bringer, A.T. Complete Structure of p97/Vasolin-Containing Protein Reveals Communication Between Nucleotide Domains. *Nat Struct Biol*. **10**, 856-863 (2003).
6. Zhang, X., Shaw, A., Bates, P. A., Newman, R. H., Gowen, B., Orlova, E., Gorman, M. A., Kondo, H., Dokurno, P., Lally, J., Leonard, G., Meyer, H., van Heel, M., & Freemont, P.S. Structure of the AAA ATPase p97. *Molecular Cell*, **6**, 1473-1484 (2000).
7. DeLaBarre, B., Christianson, J.C., Kopito, R.R., & Bringer, A.T. Central Pore Residues Mediate the p97/VCP Activity Required for ERAD. *Molecular Cell*. **22**, 451-462 (2006).
8. DeLaBarre, B., & Bringer, A.T. Nucleotide Dependent Motion and Mechanism of Action of p97/VCP. *J. Mol. Biol*. **347**, 437–452 (2005).

9. Lupas, A.N. & Martin, J. AAA proteins. *Curr. Opin. Struct. Biol.* **12**, 746–753 (2002).
10. Yeung, H. O., Klopsteck, P., Niwa, H., Isaacson, R. L., Matthews, S., Zhang, X., and Freemont, P. S. Insights into adaptor binding to the AAA protein p97. *Biochem Soc Trans.* **36**, 62-67 (2008).
11. Jentsch, S. & Rumpf, S. Cdc48 (p97): A ‘Molecular Gearbox’ In The Ubiquitin Pathway? *TRENDS in Biochemical Sciences.* **32**, 6-11 (2007).
12. Schuberth, C. & Buchberger, A. UBX Domain Proteins: Major Regulators of the AAA ATPase cdc48/P97. *Cell.Mol. Life. Sci.* **65**, 2360-2371 (2008).
13. Madsen, L., Seeger, M., Semple, C.A., & Hartmann-Petersen, R. New ATPase Regulators-p97 Goes to the PUB. *Int. J. of Biochem Cell Biol* **41**, 2380-2388 (2009).
14. Turanov, A.A., Shchedrina, V.A., Everley, R.A., Lobanov, A.V., Yim, S.H., Marino, S.M., Gygi, S.P., Hatfield, D.L. & Gladyshev, V.N. Selenoprotein S is involved in maintenance and transport of multiprotein complexes. *Biochem. J.* **462**, 555-565 (2014).
15. McCann, J. C. A. and Bruce, N. Adaptive Dysfunction of Selenoproteins from the Perspective of the Triage Theory: Why Modest Selenium Deficiency May Increase Risk of Diseases of Aging. *FASEB J.* **25**, 1793–1814 (2011)
16. Papp, L. V., Holmgren, A. and Khanna, K. K. (2010) Selenium and Selenoproteins in Health and Disease. *Antioxid. Redox Signal.* **12**, 793–795
17. Ye, Y.H., Shibata, Y., Yun, C., Ron, D. & Rapoport, T.A. A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. *Nature* **429**, 841-847 (2004).

18. Ye, Y.H., Shibata, Y., Kikkert, M., van Voorden, M., Wiertz, E., & Rapoport, T.A. Recruitment of the p97 ATPase and ubiquitin ligases to the site of retrotranslocation at the endoplasmic reticulum membrane. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 14132-14138 (2005).
19. Shchedrina, V.A., Everley, R.A., Zhang, Y., Gygi, S.P., Hatfield, D.L. & Gladyshev, V.N. Selenoprotein K binds multiprotein complexes and is involved in the regulation of endoplasmic reticulum homeostasis. *J. Biol. Chem.* **286**, 42937-42948 (2011).
20. Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R. and Gladyshev, V. N. Characterization of mammalian selenoproteomes. *Science* **300**, 1439–1443 (2003).
21. Hatfield, D.L., & Gladyshev, V.N. How Selenium Has Altered Our Understanding of the Genetic Code. *Mol. Cell. Biol.* **22**, 3565-3576 (2002).
22. Hebert, D.N., Bernasconi, R., & Molinary, M. ERAD Substrates: Which Way Out? *Sem in Cell & Dev Biol.* **21**, 526-532 (2010).
23. Christensen, L.C., Jensen, N.W., Vala, A., Kamarauskaite, J., Johansson, L., Winther, J.R., Hofmann, K., Teilum, K. & Ellgaard, L. The human selenoprotein VCP-interacting membrane protein (VIMP) is non-globular and 39 harbors a reductase function in an intrinsically disordered region. *J. Biol. Chem.* **287**, 26388-26399 (2012).
24. Liu, J., Li, F. and Rozovsky, S. The intrinsically disordered membrane protein selenoprotein S is a reductase in vitro. *Biochemistry* **52**, 3051–3061 (2013)
25. Liu, J. and Rozovsky, S. The contribution of selenocysteine to the peroxidase activity of selenoprotein S. *Biochemistry* **52**, 5514–5516 (2013)

26. Arner, E.S.J. Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? *Exp. Cell Res.* **316**, 1296–1303 (2010).
27. Meusser, B., Hirsch, C., Jarosch, E., & Sommer, T. ERAD: The Long Road to Destruction. *Nature Cell Biology.* **7**, 766-772 (2005).
28. Clechanover, A. The Ubiquitin-Proteasome Proteolytic Pathway. *Cell.* **79**. 13-21 (1994).
29. Ye, Y., Meyer, H. H., Rapoport, T.A. The AAA ATPase Cdc48/p97 and its Partners Transport Proteins from the ER into the Cytosol. *Nature.* **414**. 652- 656 (2001).
30. Hanzelmann, P., & Schindelin, H. The Structural and Functional Basis of the p97/VCP-Interacting Motif (VIM) Mutually Exclusive Binding of Cofactors to the n-terminal Domain of p97. *J. Biol. Chem.* **286**, 28679-38690 (2011).
31. DeLaBarre, B., & Bringer, A.T. Complete Structure of p97/Vasolin-Containing Protein Reveals Communication Between Nucleotide Domains. *Nat Struct Biol.* **10**, 856-863 (2003).
32. Lim, J.J., Lee, Y., Ly, T.T., Kang, J.Y., Lee, J.-G., An, J.Y., Youn, H.-S., Park, K.R., Kim, T.G., Yang, J.K., Jun, Y., Eom, S.H. Structural Insights into the Interaction of p97 N-terminus Domain and VBM Motif in Rhomboid Protease, RHBDL4. *Biochem. J.* **473**, 2863-2880 (2016).
33. Haenzelmann, P., Schindelin, H. Crystal structure of p97N in complex with the C-terminus of gp78. *J. Biol. Chem.* **286**, 38679-38690 (2011).
34. Hanzelmann, P., Buchberger, A., Schindelin, H. Hierarchical Binding of Cofactors to the AAA ATPase p97. *Structure.* **19**, 833-843 (2011).

35. Le, L.T.M., Kang, W., Kim, J.Y., Le, O.T.T., Lee, S.Y., Yang, J.K. Structural Details of Ufd1 Binding to p97 and Their Functional Implications in ER-Associated Degradation. *PLoS ONE*. **11**, e0163394-e0163394 (2016).
36. Lim, J.J., Lee, Y., Yoon, S.Y., Ly, T.T., Kang, J.Y., Youn, H.S., An, J.Y., Lee, J.G., Park, K.R., Kim, T.G., Yang, J.K., Jun, Y., Eom, S.H. Structural insights into the interaction of human p97 N-terminal domain and SHP motif in Derlin-1 rhomboid pseudoprotease. *FEBS Lett.* **590**, 4402-4413 (2016).
37. Kim, S.J., Cho, J., Song, E.J., Kim, S.J., Kim, H.M., Lee, K.E., Suh, S.W., Kim, E.E. Structural Basis for Ovarian Tumor Domain-containing Protein 1 (OTU1) Binding to p97/Valosin-containing Protein (VCP). *J. Biol. Chem.* **289**, 12264-12274 (2014).
38. Kim, K.H., Kang, W., Suh, S.W., Yang, J.K. Crystal Structure of FAF1 UBX Domain in Complex with p97/VCP N Domain Reveals a Conformational Change in the Conserved FcisP Touch-Turn Motif of UBX Domain. *Proteins* **79**, 2583-2587 (2011).
39. Hanzelmann, P., Schindelin, H. Characterization of an Additional Binding Surface on the p97 N-Terminal Domain Involved in Bipartite Cofactor Interactions. *Structure* **24**, 140-147 (2016).
40. Arumughan, A., Roske, Y., Barth, C., Forero, L.L., Bravo-Rodriguez, K., Redel, A., Kostova, S., McShane, E., Opitz, R., Faelber, K., Rau, K., Mielke, T., Daumke, O., Selbach, M., Sanchez-Garcia, E., Rocks, O., Panakova, D., Heinemann, U., Wanker, E.E. Quantitative Interaction Mapping Reveals an Extended UBX Domain in ASPL That Disrupts Functional p97 Hexamers. *Nat. Commun.* **7**, 13047-13047 (2016).

41. Dreveny, I., Kondo, H., Uchiyama, K., Shaw, A., Zhang, X., Freemont, P.S. Structural Basis of the Interaction Between the AAA ATPase p97/VCP and its Adaptor Protein p47. *Embo Journal*. **23**, 989-1216 (2004).
42. Stapf, C., Cartwright, E., Bycroft, M., Hofmann, K., Buchberger, A. The General Definition of the p97/Valosin-containing Protein (VCP)-interacting Motif (VIM) Delineates a New Family of p97 Cofactors. *J. Biol. Chem.* **286**, 38670-38678 (2011).
43. Ballar, P., Zhong, Y., Nagahama, M., Tagaya, M., Shen, Y., and Fang, S. SVIP Induces Localization of p97/VCP to the Plasma and Lysosomal Membranes and Regulates Autophagy. *J. Biol. Chem.* **282**, 33908–33914 (2007).
44. Ballar, P., Shen, Y., Yang, H., Fang, S. The Role of a Novel p97/vasolin-containing protein-interacting motif of gp48 in Endoplasmic Reticulum Associated degradation. *J. Biol. Chem.* **281**, 35359–35368 (2006).
45. Lee, J.H., Kwon, J.H., Jeon, Y.A., Ko, K.T., Lee, S.R., Kim, I.Y. Pro178 and Pro183 of Selenoprotein S Are Essential Residues for Interaction with p97(VCP) during Endoplasmic Reticulum-associated Degradation. *J. Biol. Chem.* **289**, 13758-13768 (2014).
46. Christensen, L. C., Jensen, N. W., Vala, A., Kamarauskaite, J., Johansson, L., Winther, J. R., Hofmann, K., Teilum, K., and Ellgaard, L. The Human Selenoprotein VCP-Interacting Membrane Protein (VIMP) is Nonglobular and Harbors a Reductase Function in an Intrinsically Disordered Region. *J. Biol. Chem.* **287**, 26388–26399 (2012).
47. Williamson, M. P. The Structure and Function of Proline-rich Regions in Proteins. *Biochem. J.* **297**, 249–260 (1994).

48. Morgan, A. A., and Rubenstein, E. Proline: the Distribution, Frequency, Positioning, and Common Functional Roles of Proline and Polyproline Sequences in the Human Proteome. *PLoS One*. **8**, e53785 (2013).
49. Gronenborn, A.M. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* **253**, 657–661 (1991).
50. Gallagher, T., Alexander, P., Bryan, P. & Gilliland, G.L. Two crystal structures of the  $\beta$ 1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. *Biochemistry* **33**, 4721–4729 (1994).
51. Gronenborn, A.M., Clore, G.M. Structural Studies of Immunoglobulin-Binding Domains of Streptococcal Protein G. *Immunomethods*. **2**, 3-8 (1993).
52. Gallagher, T., Alexander, P., Bryan, P., Gilliland, G.L. Two Crystal Structures of the B 1 Immunoglobulin-Binding Domain of Streptococcal Protein G and Comparison with NMR. *Biochem.* **33**, 4721-4729 (1994).
53. Frank, M.K., Dyda, F., Dobrodumoc, A., Gronenborn, A.M. Core Mutations Switch Monomeric Protein GB1 into an Intertwined Tetramer. *Nature Str. Biol.* **11**, 877-885 (2002).
54. Frericks-Schmidt, H.L., Sperling, L.J., Gao, Y,G., Wylie, B.J., Boettcher, J.M., Wilson, S.R., Rienstra, C.M. Crystal Polymorphism of Protein GB1 Examined by Solid-State NMR Spectroscopy and X-ray Diffraction. *J. Phys. Chem.* **111**, 14362-14369 (2007).
55. Richardson, J.S., Richardson, D.D. Biophysical Highlights from 54 Years of Macromolecular Crystallography. *Biophysical Journal*, **106**, 510-525 (2014).

56. Jaskolski, M., Dauter, Z., Wlodawer, A. A Brief History of Macromolecular Crystallography, Illustrated by a Family Tree and its Nobel Fruits. *The FEBS Journal*. **281**, 3985-4009 (2014).
57. Röntgen, W.C. On a new kind of rays. *Science* **3**, 227–231 (1896).
58. Röntgen, W.C. A new form of radiation. *Science* **3**, 726–729 (1896).
59. Friedrich, W., Knipping, P., Laue, M. Interferenz-Erscheinungen bei Röntgenstrahlen. *Sitz Bayer Akad Wiss*, 303–322 (1912) .
60. Bragg, W.L. The Diffraction of Short Electromagnetic Waves by a Crystal. *Proc Camb Philos Soc* **17**, 43–57 (1913) .
61. Bragg, W.H. X-rays and crystals. *Nature*. **90**, 219 (1912)
62. Bragg ,WH & Bragg, W.L. The reflection of X-rays by crystals. *Proc Roy Soc Lond A*. **88**, 428–438 (1913) .
63. Bragg, W.H. The Reflection of X-rays by Crystals (II). *Proc Roy Soc Lond A* **89**, 246–248 (1913) .
64. Bragg, W.L. The Structure of Some Crystals as Indicated by their Diffraction of X-rays. *Proc R Soc Lond A* **89**, 248–277 (1913) .
65. Bragg WH & Bragg WL (1913) The structure of the diamond. *Proc R Soc Lond A* **89**, 277–291.
66. Hünefeld FL (1840) Die Chemismus in der tierescher Organisation, p. 160, FA Brockhouse, Leipzig.
67. Reichert, E.T., Brown, A.P. The Differentiation and Specificity of Corresponding Proteins and their Vital Substances in Relation to Biological Classification and Organic Evolution. Carnegie Institution of Washington, Washington DC (1909).
68. Sumner JB (1926) The isolation and crystallization of the enzyme urease. *J Biol Chem* **69**, 435–441.

69. Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G & North ACT Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. *Nature* **185**, 416–421 (1960).
70. Klostermeier, D., Rudolph, M.G. X-Ray Crystallography. *Biophysical Chemistry*. CRC Press. 531-570 (2017).
71. Wlodawer, A., Minor, W., Dauter, Z., Jaskolski, M. Protein Crystallography for Aspiring Crystallographers or How to Avoid Pitfalls and Traps in Macromolecular Structure Determination. *The FEBS Journal*. **280**, 5705-5736 (2013)
72. Rhodes, G. Crystallography Made Clear, 3<sup>rd</sup> Edition. Associated Press, Elsevier. (2006).
73. Benvenuti, M., Mangani, S. Crystallization of Soluble Proteins in Vapor Diffusion for X-ray Crystallography. *Nat, Protoc.* **2**, 1633-1651 (2007).
74. Crystal Growth 101. Hampton Research Corp, 1991-2016.  
[https://hamptonresearch.com/documents/growth\\_101/39.pdf](https://hamptonresearch.com/documents/growth_101/39.pdf)
75. Luft, J.R., Wolfley, J.R., Said, M.I., Nagel, R.M., Lauricella, A.M., Smith, J.L., Thayer, M.H., Veatch, C.K., Snell, E.H., Malkowski, M.G., DeTitta, G.T. Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature. *Protein Sci.* **16**, 715-722 (2007).
76. Kleywegt ,G. J. Validation of Protein Crystal Structures. *Acta Crystallographica*. **56**, 249-265 (2000).
77. Kleywegt ,G. J. On Vital Aid: the Why, What and How of Validation, *Acta Crystallographica*. **65**, 134-139 (2009).

78. Read, R. J., Adams, P. D., Arendall, W. B., Brunger, A. T., Emsley, P., Joosten, R. P., Zwart, P. H. A New Generation of Crystallographic Validation Tools for the Protein Data Bank. *Structure*, **19**, 1395–1412 (2011).
79. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallographica*, **67**, 235-242 (2011).
80. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Echols, N., Headd, J. J., Zwart, P. H.. The Phenix Software for Automated Determination of Macromolecular Structures. *Methods*. **55**, 94–106 (2011).
81. Zhang, M., Ishii, K., Hisaeda, H., Murata, S., Chiba, T., Tanaka, K., Himeno, K. Ubiquitin-fusion degradation pathway plays an indispensable role in naked DNA vaccination with a chimeric gene encoding a syngeneic cytotoxic T lymphocyte epitope of melanocyte and green fluorescent protein. *Immunology*, **112(4)** 567–574 (2004).
82. Chayen, N.E., & Saridakis, E. Protein Crystallization: From Purified Protein to Diffraction Quality Crystal. *Nature Methods*, **5**, 157-153 (2008).
83. Davies, J.M., Brunger, A.T., Weis, W.I. Improved Structures of Full-Length p97, an AAA ATPase: Implications for Mechanisms of Nucleotide-Dependent Conformational Change. *Structure*. **16**, 715-726 (2008).
84. PDB ID 3CF3, Rose, A.S, Hildebrand, P.W. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res.* 43:W576-579 (2015). RCSB PDB ([www.rcsb.org](http://www.rcsb.org))

85. Tang, W.K., Li,D., Li, C.C., Esser, L., Dai, R., Guo, L., Xia, D. A novel ATP-dependent conformation in p97 N-D1 fragment revealed by crystal structures of disease-related mutants. *Embo J.* 29, 2217-2229 (2010).
86. PDB ID 5DYI, Rose, A.S, Hildebrand, P.W. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res.* 43:W576-579 (2015). RCSB PDB ([www.rcsb.org](http://www.rcsb.org))
87. Image from the RCSB PDB ([www.rcsb.org](http://www.rcsb.org)) of PDB ID 3CF3. Davies, J.M., Brunger, A.T., Weis, W.I. Improved Structures of Full-Length p97, an AAA ATPase: Implications for Mechanisms of Nucleotide-Dependent Conformational Change. *Structure.* **16**, 715-726 (2008).
88. Gronenborn, A.M., Filpula, D.R., Essig, N.Z., Achari, A., Whitlow, M., Wingfield, P.T., Clore, G.M. *Science.* **253**, 657-661(1991).
89. PDB ID 1GB1, Rose, A.S, Hildebrand, P.W. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res.* 43:W576-579 (2015). RCSB PDB ([www.rcsb.org](http://www.rcsb.org))
90. PDB ID 2Q2F, Walker, J.R., Paramanathan, R., Butler-Cole, C., Weigelt, J., Sundstrom, M., Arrowsmith, C.H., Edwards, A.M., Bochkarev, A., Dhe-Paganon, S. The Structure of Human Selenoprotein S (VCP-interacting membrane protein).

## Chapter 2

### CRYSTALLIZATION OF P97 ND1 AND FULL LENGTH P97

#### 2.1 Introduction

##### 2.1.1 Published Crystallization Conditions for p97

Success in growing protein crystals relies on a number of environmental factors, such as protein and reservoir buffer solutions, salt concentrations, pH, precipitant concentrations, and even incubation time and temperature. Slight variances in these conditions could be the difference between a clear well, an amorphous precipitant, or a diffraction quality protein crystal. Screening different conditions could take weeks or months, even with screening kits, before conditions are found that could be fine-tuned around the protein(s) of interest.

While the p97/SELENOS protein complex crystal structure has not previously been published, several other crystal conditions for full length p97 and p97 ND, both on their own and in various complexes, have been reported.

As a starting point for these experiments, all published conditions made available in the Protein Data Bank for the crystallization of p97 ND1 and full length p97 were compiled and compared to one another (Table 2.1-2.2). Based on these parameters and how closely they correlated with our initial protein conditions, coarse, and ultimately fine-tuned screens were performed.

##### 2.1.1.1 Published Conditions for p97 ND1

As the believed minimal p97 unit that binds to SELENOS, crystallization experiments were performed on the ND1 domain of p97 first. The belief is

that attempts at co-crystallization would be more successful using the shorter binding unit as opposed to the much larger full length protein.

Table 2.1 outlines the crystallization conditions for nine PDB entries for uncomplexed p97 ND1 published by two different groups. Seven of these conditions were for human p97 ND1, two from mice, and there were a mixture of mutants and ATP or ADP bound conditions. The sequence of mammalian p97 is conserved across species, and there is no sequential difference between human and mouse. However, the structure for mouse p97 ND1 was solved using residues 1-458 as opposed to the human p97 ND1, which was solved using residues 1-480.

Looking at similarities amongst the various conditions, incubation temperature of the crystal trays was the most consistent for uncomplexed p97 ND1, all of the reported temperatures for the crystals trays being incubated between 15-16°C<sup>1-3</sup>. However, focusing on the human p97 ND1 only, all eight published conditions were in that incubation range<sup>1-3</sup>.

For protein solution conditions of human p97, these appeared to remain consistent across all p97 ND1 samples despite various mutations or bound nucleotides (25 mM Tris (pH 7.6) 0.3 mM NaCl, and 1 mM Dithiothreitol (DTT))<sup>1-3</sup>.

The biggest differences, as expected, were found in the reservoir solution conditions, but even these varied only slightly between labs. All but one condition for human p97 ND1 used a 100 mM sodium citrate buffer, the pH ranges differing only slightly (pH 5.6-6). NaCl in concentrations ranging between 100-300 mM were used for all citrate buffer conditions, as well as the precipitant PEG 3350 between 7-13.6%, and 20% glycerol cryoprotectant. All but two conditions added benzamidine as a trypsin

inhibitor (0.5-6%). The outlier was the condition for PDB entry 5DYG, which used 3.7 M sodium formate (pH 6) buffer with 8% glycerol<sup>1</sup>.

Because there was little variation among the majority of the uncomplexed p97 ND1 conditions, this served as the starting point for the coarse and subsequently, fine screen conditions outlined in section 2.3.

#### **2.1.1.2 Published Conditions for Full Length p97**

In Table 2.2, eight different PDB entries for uncomplexed p97 are listed. These entries include both human and mouse p97, various amino acid mutations, loop deletion, and ATP or ADP bound conditions. In these entries, there were some conditions that remained constant throughout the crystallization conditions published by differing labs. All crystals were grown using the hanging drop vapor diffusion method, and seven out of the eight crystals were grown at an incubation temperature of 25°C, while the other was grown at a 20°C. Differences in protein buffer solution appeared to be organism specific, with the Hanzelmann lab using 10 mM HEPES (pH 8), 100 mM NaCl, 5 mM MgCl<sub>2</sub>, and 1 mM tris(2-carboxyethyl)phosphine (TCEP) for all human protein variants<sup>5</sup>, and the Weiss lab using 25 mM Tris (pH 7.5), 150 mM KCl, 10 mM β-mercaptoethanol (BME), 10 mM MgCl<sub>2</sub>, and 10 mM Adenosine 5'-[γ-thio]triphosphate (ATP-γ-S) for all mouse protein variants<sup>6</sup>. ATP-γ-S is a commonly used non-hydrolyzed analogue of ATP where one of the gamma-phosphate oxygens is replaced with sulfur. Because of this, the anion hydrolyzes at a much slower rate than ATP<sup>8</sup>. The Freemont lab only had one PDB entry, and used 25 mM Hepes, 250 mM KCl, 2 mM BME, 1 mM MgCl<sub>2</sub> for their mouse p97 entry<sup>7</sup>.

Our work focused specifically on human p97, therefore those conditions were of particular interest. The Hanzelmann group used 100 mM MES buffer for all p97 full

length crystallizations, whether mutated and ATP or ADP bound. The differences in crystallizations conditions appeared to be a function of what nucleotide the protein was bound to as opposed to variations on the protein itself. All ATP bound conditions used 100 mM MES buffer at a pH of 5.75 with 0.4-0.6 M potassium acetate and 6.5-7% PEG 4000 as the precipitant. The ADP bound condition utilized 100 mM MES buffer at pH 6.5 with 150 mM ammonium sulfate and 12% PEG 4000 as the precipitant. All conditions where the protein was not peptide bound used a protein concentration of 8.5mg/ml. These ADP bound conditions served as the basis for our in-house screening outlined in section 2.4.

## **2.2 Sample Purification and Preparation for Crystallization**

As previously discussed, protein purity is of the utmost importance to ensure optimal crystal growth. In addition to standard purification procedures for p97<sup>1-7</sup>, an additional step was added whereby the protein was purified on an FPLC using a GE Superdex 200 column (120mL capacity). This allowed for the further separation of p97 (applicable for both full length, ND1 without the Polyhistidine-tag, and ND1 with Polyhistidine-tag) from any other impurities, particularly oligomeric impurities and aggregates.

For the full length p97 trials, the protein buffer solution used in the FPLC purification was 10 mM HEPES (pH 8), 100mM NaCl, and 5mM MgCl<sub>2</sub>. The protein buffer solution for the p97 ND1 without the Polyhistidine-tag was a 25mM Tris buffer (pH 8) with 100mM NaCl, and the p97 ND1 protein with the Polyhistidine-tag intact used a 25mM Tris buffer (pH 7.5) with 300mM NaCl.

Five milliliters of the protein sample were injected onto the Superdex 200 column, and the eluent was collected in a fraction collector in 3mL increments, with the buffer

flow rate at 1.0ml/min. A UV-Vis detector at 280nm was utilized to determine the fractions with the protein of interest, the peak for the aggregates/impurities typically appearing just before the p97 peak. These protein fractions were collected, combined, and concentrated.

Prior to crystallization, the purified proteins were treated with 1mM reducing agent (TCEP for p97 full length and DTT for ND1) and incubated with 0.5mM ADP at least 30 minutes prior to setting the crystal trays.

## **2.3 Crystallization of p97 ND1**

### **2.3.1 Crystallization Conditions**

While the ultimate goal of this work was to crystallize the p97/SELENOS protein complex, the first step was to first attempt to crystallize the portion of p97 that was believed to be directly involved in the binding of SELENOS, the ND1 domain. Discussed in 2.1.1 were the previously published conditions of ND1 from which we based our initial crystallization trials. Crystals for p97 ND1 with the C-terminal Polyhistidine-tag already cleaved were attempted first, however, since there was the possibility of needing to keep the Polyhistidine-tag intact for future purification of the p97/SELENOS complex, attempts were also made to see if p97 ND1 could also crystallize with the Polyhistidine-tag still present. Polyhistidine tags are common affinity tags used in the purification of recombinant proteins, where at least six histidine residues are added, or tagged, to the N- or C- terminus of a protein. The tagged protein can be purified by being run through an immobilized metal affinity column (IMAC), where the histidine residues interact with the metal ions, the other protein impurities without the histidine tag passing through the

column<sup>9</sup>. Because of the possible need to keep the tag intact for further purification, it was beneficial to attempt to crystallize both with and without the Polyhistidine tag.

Two of the main components in a hanging drop crystallization procedure are the reservoir buffer (with various additives or stabilizers) and the precipitants. The most commonly used reservoir buffer in published conditions was 100 mM sodium citrate in a pH range of 5.5-6. The pH of 5.8 was the most prevalent, and therefore was the pH used in subsequent successful trials. The salt content of the reservoir buffer in our trials varied between 100 and 300mM NaCl for both the Polyhistidine-tag containing and Polyhistidine-tag cleaved proteins. PEG 3350 in concentrations ranging from 5-20% was used as a precipitant, benzamidine in concentrations ranging from 2-6% was used as a trypsin inhibitor, and glycerol in concentrations ranging from 15-20% was used as a cryoprotectant (Tables 2.3 and 2.4).

All crystallization attempts utilized the hanging drop vapor diffusion method, carried out in 48 well crystallization plates. The buffer solution, glycerol, benzamidine, and PEG were added in various concentrations to a final volume of 400 $\mu$ L in the well. On each coverslip, 1  $\mu$ L of the reservoir solution was mixed with 1 $\mu$ L of the protein solution at concentrations of 7 mg/ml or 14 mg/ml. The mixture was suspended as a droplet on the coverslip, inverted over each well, and sealed using the pre-greased trays. The trays were labelled and incubated at 16°C.

### **2.3.2 Crystal Tray Monitoring**

The crystal trays were monitored using an optical stereoscope, starting at Day 0, to observe the conditions and appearance of each drop at the time of incubation. The drops were monitored every two-three days within the first week, and then at minimum once a week thereafter for up to six weeks. Crystals showed no noticeable change between the

first week and the third, however, typically after 3 weeks, the drops started to show slightly more amorphous precipitant and oiling. The trays were then checked sporadically for approximately one year. After several months, the crystals were smaller than previously observed or disappeared completely, and the drops showed an increase in salt crystals and amorphous precipitant. Crystal tracking sheets were utilized to note the appearance of each well and compare changes day to day. Evidence of phase separation, amorphous precipitants, salt crystals, or possible proteins crystals were all noted in the tracking sheets.

### **2.3.3 Results for ND1 without Polyhistidine-tag**

Two protein concentrations were initially utilized for this crystallization: 7 mg/ml and 14 mg/ml, as well as two salt concentrations in the buffer solution, 100 and 300 mM NaCl (Table 2.3). The same variations in precipitant were consistent across both salt and protein concentrations and compared to one another. It was observed that most of the higher protein wells saw amorphous protein within two days, indicating conditions well beyond the supersaturation zone and into the precipitation zone. It was also noted that the higher salt concentrations saw the formation of what appeared to be salt crystals, not protein crystals (confirmed by application of the dye test). The lower protein concentrations, however, saw more phase separation, a sign that the conditions are approaching the supersaturation zone. Consistent with these observations, the condition to yield what appeared to be protein crystals was the 100 mM NaCl citrate buffer solution using 7 mg/ml protein, with 13% PEG 3350, 20% glycerol, and 2% benzamidine (Table 2.3, well A3). These were whisker shaped microcrystals approximately 31 $\mu$ m in diameter that appeared after day 5 and were still observable after day 18.

Fine screening around these conditions were then performed (Table 2.4), using 100 mM NaCl citrate buffer solution (pH 5.8) with 7 mg/ml protein, 10-20% PEG 3350, 15-30% glycerol, and 1-3% benzamidine. This fine screen yielded crystals more than twice the size of the previous trial in wells B4 (20% PEG 3350, 15% glycerol, and 2% benzamidine) and B8 (20% PEG 3350, 20% glycerol, and 2% benzamidine). The crystals in well B4 and B6 appeared after 4 days and grew to full size by day 6, growing to ~65 microns and ~130 microns in each well respectively.

Because full, diffraction quality crystal formation of p97 ND1 was not the ultimate goal, only proof of capability, further optimizations around these conditions were not done. Instead, these parameters were then used as starting points for the p97/SELENOS complex crystallizations.

#### **2.3.4 Results for ND1 with Polyhistidine-tag**

To determine if p97 ND1 could be crystallized with the Polyhistidine-tag still intact, an attribute that would aid in the purification of the p97/SELENOS complex, an approach similar to what was used for p97 ND1 without the Polyhistidine-tag was attempted. Using the previously optimized conditions as a starting point, as well as Polyhistidine-tag containing published conditions in Table 2.1, only 7 mg/ml protein concentrations were used, and the same 100 mM sodium citrate buffer (pH 5.8) was also utilized with the 100 and 300 mM NaCl variations. The reservoir solution was made up of PEG 3350 in concentrations ranging between 5-20%, benzamidine in ranges between 2-6%, and glycerol in ranges between 15-20%(Table 2.5).

Optimizing around the previous ND1 conditions, it was observed that there were fewer wells with exclusively amorphous precipitate. Instead, there were more wells with microcrystals or phase separation, and three separate conditions yielded possible protein

crystals within the first 6 days, and each of those three were somewhat different in crystal shape.

Well C2 (100 mM citrate buffer (pH 5.8), 100 mM NaCl, 10% PEG 3350, 15% glycerol, 6% benzamidine) produced crystals similar in size and shape to those grown in previous ND1 trials. Well D7 (100 mM citrate buffer (pH 5.8), 300 mM NaCl, 15% PEG 3350, 20% glycerol, 2% benzamidine) produced cluster shaped crystals, while well E4 (100 mM citrate buffer (pH 5.8), 300 mM NaCl, 20% PEG 3350, 15% glycerol, 4% benzamidine) produced needle-like crystals. The crystals in well C2 and D7 were similar in size, measuring approximately 30  $\mu\text{m}$  in length, while the needle-like crystals in well E4 were closer to 130  $\mu\text{m}$ .

Again, since these trials were to prove capability and not optimization of the ND1 crystal on its own, further optimizations were not performed. Instead, these proved that the p97 ND1 could be crystallized both with or without the Polyhistidine-tag under these conditions, and these parameters could serve as starting points in the crystallization of the p97 ND1/SELENOS complex.

## **2.4 Crystallization of p97 Full Length**

### **2.4.1 Crystallization Conditions**

If it was found that the p97/SELENOS complex was more stable and/or tightly bound using the full length protein as opposed to p97 ND1, it was also important to investigate the crystallization of that protein on its own. Once again, the starting conditions were based on published results from successful p97 crystallization attempts, using buffer solutions and precipitant concentrations used by the most labs successfully. In this case, the buffer for initial screens was a 100 mM MES buffer at both pH 6.5 and 7,

both containing ammonium sulfate in concentration ranges between 100 and 200 mM, and using PEG 4000 as a precipitant in concentration ranges between 11-18% (Table 2.6-2.7).

All crystallization attempts utilized the hanging drop vapor diffusion method, carried out in 48 well crystallization plates. The buffer solutions and PEG 4000 were added in various concentrations and pHs to a final volume of 400 $\mu$ L in each well. On the coverslips, 1  $\mu$ L of the reservoir solution was mixed with 1 $\mu$ L of the protein solution at concentration of 5 mg/ml, 7.5 mg/ml, or 14 mg/ml protein. The mixture was suspended as a droplet on the coverslip, inverted over each well, and sealed using the pre-greased trays. The trays were labelled and incubated at 20°C.

#### **2.4.2 Crystal Tray Monitoring**

The crystal trays were monitored using an optical stereoscope, starting at Day 0, to observe the conditions and appearance of each drop at the time of incubation. The drops were monitored every two-three days within the first week, and then at minimum once a week thereafter for up to six weeks. Crystals showed no noticeable change between the first week and the third. However, typically after 3 weeks, the drops started to show slightly more amorphous precipitant and oiling. The trays were then checked sporadically for approximately one year. After several months, the crystals were smaller than previously observed and the drops showed an increase in salt crystals and amorphous precipitant. Crystal tracking sheets were utilized to note the appearance of each well and compare changes day to day. Evidence of phase separation, amorphous precipitants, salt crystals, or possible proteins crystals were all noted in the tracking sheets.

### 2.4.3 Results

Trials were performed using the full length p97 without expression tags. The largest variations being the protein concentration and possible protein purity. The first of the two main trials used protein in concentrations of 5 and 10 mg/ml (Table 2.6), whereas the second trial used a protein concentration of 7.5 mg/ml (Table 2.7). Trial 1 results are not listed in detail as there were no suitable crystals. For the second trial, however, there were several promising wells, particularly the wells in row F.

Overall, many of the wells in trial 2 showed evidence of phase separation and some precipitation by day 2, and by day 4, there was evidence of possible protein microcrystals, with little change between days 4-7. None of the crystals grew noticeably in size after day 7 for the six weeks they were regularly monitored, and some diminished in size after a few months. The most successful row, F, had consistent buffer conditions (100 mM MES with 200 mM ammonium sulfate at pH 7) and the only variation was the PEG 4000 concentration ranging from 11% to 18%. Like the p97 ND1, these were whisker shaped microcrystals, with a measurement of typically less than 25 microns across. Of note was that while only row F (F1, F2, F5, F6, F7, F8) yielded crystals, several other wells showed evidence of phase separation, which is typically indicative of conditions close to the supersaturation zone.

The differences in the two trials could be prime examples as to the importance of protein concentration and protein purity in crystallography. While the reservoir conditions remained consistent across both trials, the protein concentration varied, with the second trial choosing a concentration right in the middle of the range. The second trial also used fresher protein preparation, and as a result of these two factors, there was less amorphous precipitate in the wells and observable, albeit small, crystals were able to form.

Since these were merely proof of capability experiments, additional screens to optimize for diffraction quality crystals were not performed. As with the ND1 experiments, these conditions would be used as starting points for any possible future full length p97/SELENOS complex crystallization attempts.

### **Acknowledgements**

The author would like to thank Zhengqi Zhang, Vicki Wallace, and Rujin Cheng of the Rozovsky group for their work in expressing and purifying the various p97 proteins used for crystallization, as well as Mackenzie Lauro and Brian Bahnson for their assistance in the protein crystallography.

Table 2.1: Published Crystallization Conditions for p97 ND1

FDB entry	Organism	Mutation	Tag	Nucleotide Bound	Method	Temp (C)	Protein Solution	pH	[Protein]	Reservoir Solution				
5DY1 <sup>1</sup>	human	none	C-terminal His6	ADP	vapor diffusion sitting drop	16	25 mM Tris, pH 7.6, 0.3 mM NaCl, 1 mM DTT	6	7 mg/mL	100 mM Sodium Citrate	15.2% PEG 3350	2% benzamidine	20 % glycerol	300 mM NaCl
5DYG <sup>1</sup>	human	L198W	C-terminal His6	ADP	vapor diffusion sitting drop	16	25 mM Tris, pH 7.6, 0.3 mM NaCl, 1 mM DTT	6	7 mg/mL	3.7 M sodium formate			8% glycerol	
4KLN <sup>2</sup>	human	A232E	C-terminal His6	ATP-g-S	vapor diffusion sitting drop	15	25 mM Tris, pH 7.6, 0.3 mM NaCl, 1 mM DTT	5.8	N/A	100 mM Citrate	13.6% PEG 3350	0.525% benzamidine	20% glycerol	300 mM NaCl
4KO8 <sup>2</sup>	human	R155H	C-terminal His6	ATP-g-S	vapor diffusion hanging drop	16	25 mM Tris, pH 7.6, 0.3 mM NaCl, 1 mM DTT	5.6	N/A	100 mM Citrate	7% PEG 3350	6% benzamidine	20% glycerol	
4KOD <sup>2</sup>	human	R155H	C-terminal His6	ADP	vapor diffusion sitting drop	16	25 mM Tris, pH 7.6, 0.3 mM NaCl, 1 mM DTT	5.8	N/A	100 mM Citrate	13.6% PEG 3350	0.625% benzamidine	20% glycerol	300 mM NaCl
3HU1 <sup>3</sup>	human	R95G	C-terminal His6	ATP-g-S	vapor diffusion hanging drop	15	25 mM Tris, pH 7.6, 0.3 mM NaCl, 1 mM DTT	5.7	5 mg/mL	100 mM Citrate	13.5% PEG 3350		20% glycerol	100 mM NaCl
3HU2 <sup>3</sup>	human	R86A	C-terminal His6	ATP-g-S	vapor diffusion hanging drop	15	25 mM Tris, pH 7.6, 0.3 mM NaCl, 1 mM DTT	5.8	5 mg/mL	100 mM Citrate	13.6% PEG 3350	4% benzamidine	20% glycerol	100 mM NaCl
3HU3 <sup>3</sup>	human	R155H	C-terminal His6	ATP-g-S	vapor diffusion hanging drop	15	25 mM Tris, pH 7.6, 0.3 mM NaCl, 1 mM DTT	5.6	7 mg/mL	100 mM Citrate	7% PEG 3350	6% benzamidine	20% glycerol	100 mM NaCl
1E32 <sup>4</sup>	mouse	none	None	ADP	N/A	N/A	20 mM HEPES (pH 7.5), 10 mM MgCl <sub>2</sub> , 600 mM KCl, and 2 mM β-mercaptoethanol	5.5-6	16 mg/mL	3.8-4.2 M Na Formate	5% PEG 600		10% glycerol	

Table 2.2 Published Crystallization Conditions for Full Length p97

PDB entry	Organism	Protein	Mutation	Tag	Nucleotide Bound	Method	Temp. C	Protein Solution	pH	[Protein]	Reservoir Solution			
5C18 <sup>5</sup>	human	p97 Δ709-728	Δ709-728	None	ATP-g-S	vapor diffusion hanging drop	25	10 mM HEPES (pH 8), 100 mM NaCl, 5 mM MgCl <sub>2</sub> , 1 mM TCEP	5.75	8.5 mg/mL	0.4-0.6 M potassium acetate	6.5-7% PEG 4000	100 mM MES	
5C1A <sup>5</sup>	human	p97 mutant	N750D/R753D/M757D/Q760D	None	ATP-g-S	vapor diffusion hanging drop	25	10 mM HEPES (pH 8), 100 mM NaCl, 5 mM MgCl <sub>2</sub> , 1 mM TCEP	5.75	8.5 mg/mL	0.4-0.6 M potassium acetate	6.5-7% PEG 4000	100 mM MES	
5C19 <sup>5</sup>	human	p97	N750D/R753D/M757D/Q760D	None	ADP	vapor diffusion hanging drop	25	10 mM HEPES (pH 8), 100 mM NaCl, 5 mM MgCl <sub>2</sub> , 1 mM TCEP	6.5	8.5 mg/mL	150 mM ammonium sulfate	12% PEG 4000	100 mM MES	
3CF0 <sup>6</sup>	mouse	p97 D2	none	His6 (location not listed)	ADP	vapor diffusion hanging drop	25	25 mM Tris (pH 7.5), 150 mM KCl, 10 mM beta MEtOH, 10 mM MgCl <sub>2</sub> , 10 mM ATPgamma S	7.1	20 mg/mL	100 mM Tris		2.45 M ammonium sulfate	200 mM lithium sulfate
3CF1 <sup>6</sup>	mouse	p97 (refined from 1E32 and 3CF0 supercedes 1OZ4 and 1YQ0)	none	His6	ADP	vapor diffusion hanging drop	25	25 mM Tris (pH 7.5), 150 mM KCl, 10 mM beta MEtOH, 10 mM MgCl <sub>2</sub> , 10 mM ATPgamma S	7	14 mg/mL	50 mM NH <sub>4</sub> F	5% PEG 3350	100 mM HEPES	5 mM DTT
3CF2 <sup>6</sup>	mouse	p97 (supercedes 1YPW)	none	His6 (location not listed)	ADP/AMP-PNP	vapor diffusion hanging drop	25	25 mM Tris (pH 7.5), 150 mM KCl, 10 mM beta MEtOH, 10 mM MgCl <sub>2</sub> , 10 mM ATPgamma S	5	14 mg/mL	50 mM NH <sub>4</sub> F	6% PEG 400	0.75 M sodium dihydrogen phosphate	100 mM citrate buffer
3CF3 <sup>6</sup>	mouse	p97 (supercedes 1YQ1)	none	His6 (location not listed)	ADP	vapor diffusion hanging drop	25	25 mM Tris (pH 7.5), 150 mM KCl, 10 mM beta MEtOH, 10 mM MgCl <sub>2</sub> , 10 mM ATPgamma S	7	14 mg/mL	50 mM NH <sub>4</sub> F	5% PEG 3350	5 mM DTT	100 mM HEPES
1R7R <sup>7</sup>	mouse	p97	none	C-terminal His6	ADP	vapor diffusion hanging drop	20	25 mM Hepes, 250 mM KCl, 2 mM Bme, 1 mM MgCl <sub>2</sub> then added either 10 mM AMP-PNP or 2 mM 2'-BrAMP-PNP	5.6	7 mg/mL	1.0-1.4 M ammonium phosphate monobasic	1-5% PEG 400	2 mM beta MEtOH	100 mM sodium citrate

Table 2.3 Conditions for In-House p97 ND1 Screening 1 (Without Polyhistidine-tag)

Well	Buffer	pH	PEG 3350%	Benzamidine %	Glycerol %
A1	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	7	2	20
A2	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	2	20
A3	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	13	2	20
A4	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	2	20
A5	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	7	2	20
A6	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	2	20
A7	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	13	2	20
A8	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	2	20
B1	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	7	4	20
B2	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	4	20
B3	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	13	4	20
B4	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	4	20
B5	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	7	4	20
B6	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	4	20
B7	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	13	4	20
B8	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	4	20
C1	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	7	6	20
C2	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	6	20
C3	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	13	6	20

Well	Buffer	pH	PEG 3350%	Benzamidine %	Glycerol %
C4	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	6	20
C5	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	7	6	20
C6	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	6	20
C7	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	13	6	20
C8	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	6	20
D1	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	7	2	20
D2	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	10	2	20
D3	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	13	2	20
D4	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	15	2	20
D5	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	7	2	20
D6	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	10	2	20
D7	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	13	2	20
D8	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	15	2	20
E1	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	7	4	20
E2	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	10	4	20
E3	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	13	4	20
E4	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	15	4	20
E5	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	7	4	20
E6	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	10	4	20
E7	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	13	4	20

Well	Buffer	pH	PEG 3350 %	Benzamidine %	Glycerol %
E8	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	15	4	20
F1	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	7	6	20
F2	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	10	6	20
F3	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	13	6	20
F4	100mM Citrate Buffer; 100mM NaCl;14mg/ml protein	5.8	15	6	20
F5	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	7	6	20
F6	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	10	6	20
F7	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	13	6	20
F8	100mM Citrate Buffer; 300mM NaCl;14mg/ml protein	5.8	15	6	20

Table 2.4 Conditions for In-House p97 ND1 Screening 2 (Without Polyhistidine-tag)

Well	Buffer (7mg/ml protein)	pH	PEG 3350 %	Benzamidine %	Glycerol %
A1	100mM Citrate Buffer; 100mM NaCl	5.8	10	1	15
A2	100mM Citrate Buffer; 100mM NaCl	5.8	13	1	15
A3	100mM Citrate Buffer; 100mM NaCl	5.8	16	1	15
A4	100mM Citrate Buffer; 100mM NaCl	5.8	20	1	15
A5	100mM Citrate Buffer; 100mM NaCl	5.8	10	1	20
A6	100mM Citrate Buffer; 100mM NaCl	5.8	13	1	20
A7	100mM Citrate Buffer; 100mM NaCl	5.8	16	1	20
A8	100mM Citrate Buffer; 100mM NaCl	5.8	20	1	20
B1	100mM Citrate Buffer; 100mM NaCl	5.8	10	2	15
B2	100mM Citrate Buffer; 100mM NaCl	5.8	13	2	15
B3	100mM Citrate Buffer; 100mM NaCl	5.8	16	2	15
B4	100mM Citrate Buffer; 100mM NaCl	5.8	20	2	15
B5	100mM Citrate Buffer; 100mM NaCl	5.8	10	2	20
B6	100mM Citrate Buffer; 100mM NaCl	5.8	13	2	20
B7	100mM Citrate Buffer; 100mM NaCl	5.8	16	2	20
B8	100mM Citrate Buffer; 100mM NaCl	5.8	20	2	20
C1	100mM Citrate Buffer; 100mM NaCl	5.8	10	3	15
C2	100mM Citrate Buffer; 100mM NaCl	5.8	13	3	15
C3	100mM Citrate Buffer; 100mM NaCl	5.8	16	3	15
C4	100mM Citrate Buffer; 100mM NaCl	5.8	20	3	15
C5	100mM Citrate Buffer; 100mM NaCl	5.8	10	3	20
C6	100mM Citrate Buffer; 100mM NaCl	5.8	13	3	20
C7	100mM Citrate Buffer; 100mM NaCl	5.8	16	3	20
C8	100mM Citrate Buffer; 100mM NaCl	5.8	20	3	20
D1	100mM Citrate Buffer; 100mM NaCl	5.8	10	1	25
D2	100mM Citrate Buffer; 100mM NaCl	5.8	13	1	25
D3	100mM Citrate Buffer; 100mM NaCl	5.8	16	1	25
D4	100mM Citrate Buffer; 100mM NaCl	5.8	20	1	25
D5	100mM Citrate Buffer; 100mM NaCl	5.8	10	1	30
D6	100mM Citrate Buffer; 100mM NaCl	5.8	13	1	30
D7	100mM Citrate Buffer; 100mM NaCl	5.8	16	1	30
D8	100mM Citrate Buffer; 100mM NaCl	5.8	20	1	30
E1	100mM Citrate Buffer; 100mM NaCl	5.8	10	2	25
E2	100mM Citrate Buffer; 100mM NaCl	5.8	13	2	25
E3	100mM Citrate Buffer; 100mM NaCl	5.8	16	2	25

Well	Buffer (7mg/ml protein)	pH	PEG 3350%	Benzamidine %	Glycerol %
E4	100mM Citrate Buffer; 100mM NaCl	5.8	20	2	25
E5	100mM Citrate Buffer; 100mM NaCl	5.8	10	2	30
E6	100mM Citrate Buffer; 100mM NaCl	5.8	13	2	30
E7	100mM Citrate Buffer; 100mM NaCl	5.8	16	2	30
E8	100mM Citrate Buffer; 100mM NaCl	5.8	20	2	30
F1	100mM Citrate Buffer; 100mM NaCl	5.8	10	3	25
F2	100mM Citrate Buffer; 100mM NaCl	5.8	13	3	25
F3	100mM Citrate Buffer; 100mM NaCl	5.8	16	3	25
F4	100mM Citrate Buffer; 100mM NaCl	5.8	20	3	25
F5	100mM Citrate Buffer; 100mM NaCl	5.8	10	3	30
F6	100mM Citrate Buffer; 100mM NaCl	5.8	13	3	30
F7	100mM Citrate Buffer; 100mM NaCl	5.8	16	3	30
F8	100mM Citrate Buffer; 100mM NaCl	5.8	20	3	30

Table 2.5 Conditions for In-House p97 ND1 Screening (With Polyhistidine-tag)

Well	Buffer (7mg/ml protein)	pH	PEG 3350 %	Benzamidine %	Glycerol %
A1	100mM Citrate Buffer; 100mM NaCl	5.8	10	2	15
A2	100mM Citrate Buffer; 100mM NaCl	5.8	13	2	15
A3	100mM Citrate Buffer; 100mM NaCl	5.8	16	2	15
A4	100mM Citrate Buffer; 100mM NaCl	5.8	20	2	15
A5	100mM Citrate Buffer; 100mM NaCl	5.8	10	2	20
A6	100mM Citrate Buffer; 100mM NaCl	5.8	13	2	20
A7	100mM Citrate Buffer; 100mM NaCl	5.8	16	2	20
A8	100mM Citrate Buffer; 100mM NaCl	5.8	20	2	20
B1	100mM Citrate Buffer; 100mM NaCl	5.8	10	4	15
B2	100mM Citrate Buffer; 100mM NaCl	5.8	13	4	15
B3	100mM Citrate Buffer; 100mM NaCl	5.8	16	4	15
B4	100mM Citrate Buffer; 100mM NaCl	5.8	20	4	15
B5	100mM Citrate Buffer; 100mM NaCl	5.8	10	4	20
B6	100mM Citrate Buffer; 100mM NaCl	5.8	13	4	20
B7	100mM Citrate Buffer; 100mM NaCl	5.8	16	4	20
B8	100mM Citrate Buffer; 100mM NaCl	5.8	20	4	20
C1	100mM Citrate Buffer; 100mM NaCl	5.8	10	6	15
C2	100mM Citrate Buffer; 100mM NaCl	5.8	13	6	15
C3	100mM Citrate Buffer; 100mM NaCl	5.8	16	6	15
C4	100mM Citrate Buffer; 100mM NaCl	5.8	20	6	15
C5	100mM Citrate Buffer; 100mM NaCl	5.8	10	6	20
C6	100mM Citrate Buffer; 100mM NaCl	5.8	13	6	20
C7	100mM Citrate Buffer; 100mM NaCl	5.8	16	6	20
C8	100mM Citrate Buffer; 100mM NaCl	5.8	20	6	20
D1	100mM Citrate Buffer; 300mM NaCl	5.8	10	2	15
D2	100mM Citrate Buffer; 300mM NaCl	5.8	13	2	15
D3	100mM Citrate Buffer; 300mM NaCl	5.8	16	2	15
D4	100mM Citrate Buffer; 300mM NaCl	5.8	20	2	15
D5	100mM Citrate Buffer; 300mM NaCl	5.8	10	2	20
D6	100mM Citrate Buffer; 300mM NaCl	5.8	13	2	20
D7	100mM Citrate Buffer; 300mM NaCl	5.8	16	2	20
D8	100mM Citrate Buffer; 300mM NaCl	5.8	20	2	20
E1	100mM Citrate Buffer; 300mM NaCl	5.8	10	4	15
E2	100mM Citrate Buffer; 300mM NaCl	5.8	13	4	15
E3	100mM Citrate Buffer; 300mM NaCl	5.8	16	4	15

Well	Buffer (7mg/ml protein)	pH	PEG 3350 %	Benzamidine %	Glycerol %
E4	100mM Citrate Buffer; 300mM NaCl	5.8	20	4	15
E5	100mM Citrate Buffer; 300mM NaCl	5.8	10	4	20
E6	100mM Citrate Buffer; 300mM NaCl	5.8	13	4	20
E7	100mM Citrate Buffer; 300mM NaCl	5.8	16	4	20
E8	100mM Citrate Buffer; 300mM NaCl	5.8	20	4	20
F1	100mM Citrate Buffer; 300mM NaCl	5.8	10	6	15
F2	100mM Citrate Buffer; 300mM NaCl	5.8	13	6	15
F3	100mM Citrate Buffer; 300mM NaCl	5.8	16	6	15
F4	100mM Citrate Buffer; 300mM NaCl	5.8	20	6	15
F5	100mM Citrate Buffer; 300mM NaCl	5.8	10	6	20
F6	100mM Citrate Buffer; 300mM NaCl	5.8	13	6	20
F7	100mM Citrate Buffer; 300mM NaCl	5.8	16	6	20
F8	100mM Citrate Buffer; 300mM NaCl	5.8	20	6	20

Table 2.6 Conditions for In-House First Full Length p97 Screening

Well	Buffer	pH	PEG 4000 %
A1	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10mg/ml protein	7.0	11
A2	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10mg/ml protein	7.0	12
A3	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10mg/ml protein	7.0	13
A4	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10mg/ml protein	7.0	14
A5	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10mg/ml protein	7.0	15
A6	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10mg/ml protein	7.0	13
A7	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10mg/ml protein	7.0	17
A8	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10mg/ml protein	7.0	18
B1	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	11
B2	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	12
B3	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	13
B4	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	14
B5	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	15
B6	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	13
B7	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	17
B8	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	18
C1	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	11
C2	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	12
C3	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	13
C4	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	14
C5	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	15
C6	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	13
C7	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	17
C8	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 10 mg/ml protein	7.0	18
D1	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	10
D2	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	13
D3	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	16
D4	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	20
D5	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	10
D6	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	13
D7	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	16
D8	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	20
E1	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	10
E2	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	13
E3	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	16

Well	Buffer	pH	PEG 4000 %
E4	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	20
E5	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	10
E6	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	13
E7	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	16
E8	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	20
F1	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	10
F2	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	13
F3	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	16
F4	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	20
F5	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	10
F6	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	13
F7	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	16
F8	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 5 mg/ml protein	7.0	20

Table 2.7 Conditions for In-House Second Full Length p97 Screening

Well	Buffer	pH	PEG 4000 %
A1	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5mg/ml protein	6.5	11
A2	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5mg/ml protein	6.5	12
A3	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5mg/ml protein	6.5	13
A4	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5mg/ml protein	6.5	14
A5	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5mg/ml protein	6.5	15
A6	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5mg/ml protein	6.5	13
A7	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5mg/ml protein	6.5	17
A8	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5mg/ml protein	6.5	18
B1	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	11
B2	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	12
B3	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	13
B4	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	14
B5	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	15
B6	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	13
B7	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	17
B8	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	18
C1	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	11
C2	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	12
C3	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	13
C4	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	14
C5	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	15
C6	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	13
C7	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	17
C8	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	6.5	18
D1	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	11
D2	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	12
D3	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	13
D4	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	14
D5	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	15
D6	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	13
D7	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	17
D8	100mM MES Buffer; 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	18
E1	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	11
E2	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	12
E3	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	13

Well	Buffer	pH	PEG 4000 %
E4	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	14
E5	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	15
E6	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	16
E7	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	17
E8	100mM MES Buffer; 150mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	18
F1	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	11
F2	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	12
F3	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	13
F4	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	14
F5	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	15
F6	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	13
F7	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	17
F8	100mM MES Buffer; 200mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ; 7.5 mg/ml protein	7.0	18

## Crystal Images

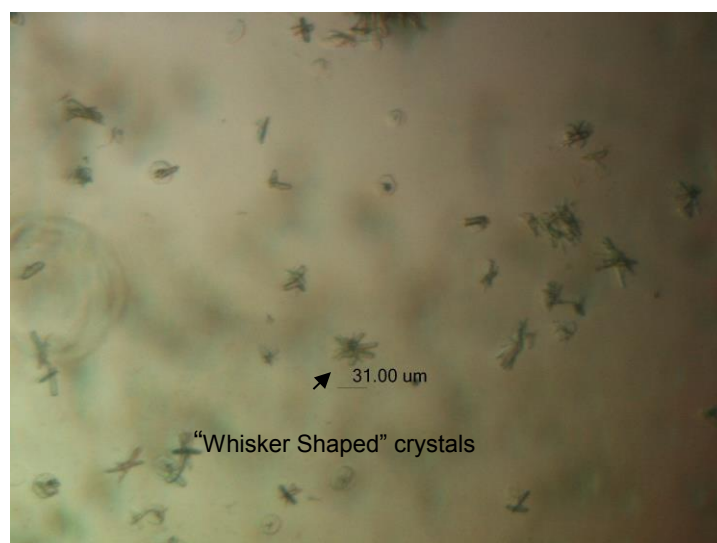


Figure 2.1: p97 ND1 (Without Polyhistidine-tag) Screen 1, Well A3, 11.25x magnification

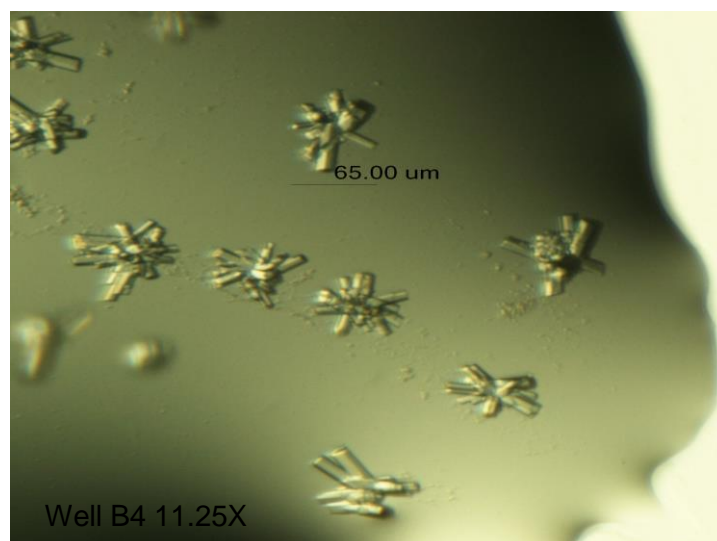


Figure 2.2: p97 ND1 (Without Polyhistidine-tag) Screen 2, Well B4, 11.25x magnification

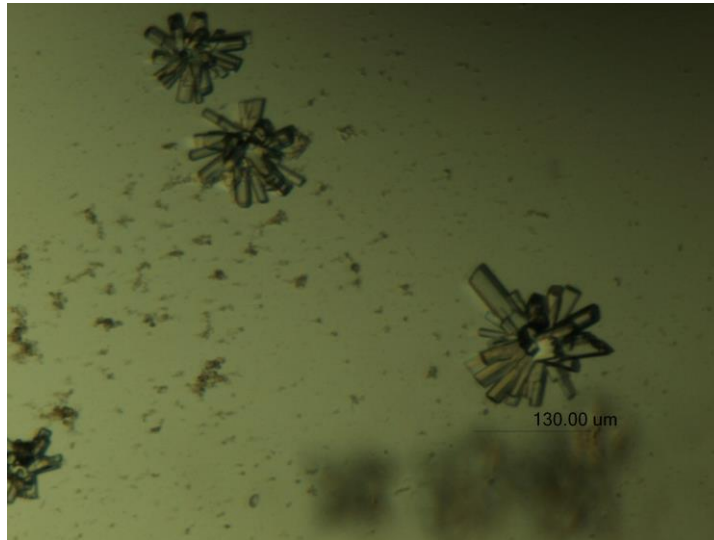


Figure 2.3: p97 ND1 (Without Polyhistidine-tag) Screen 2, Well B8, 11.25x magnification

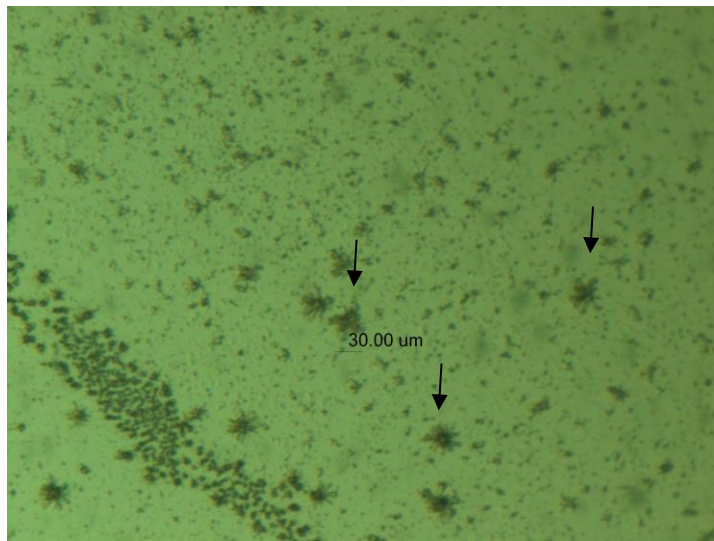


Figure 2.4: p97 ND1 (With Polyhistidine-tag), Well C2, 11.25x magnification

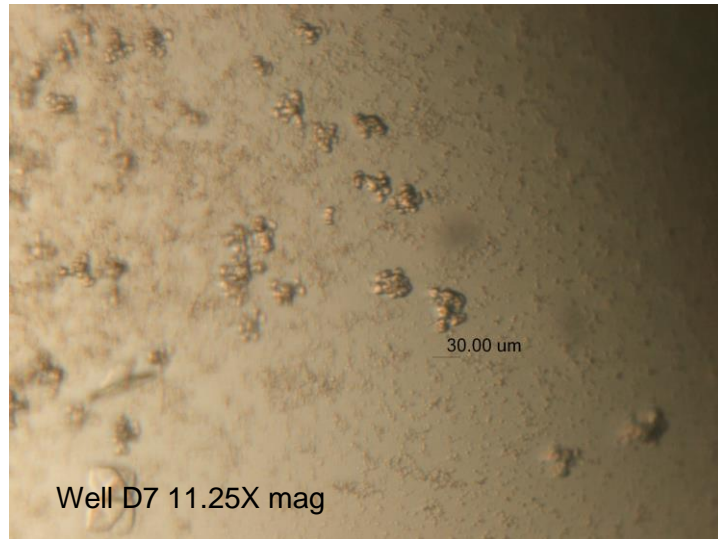


Figure 2.5: p97 ND1 (With Polyhistidine-tag), Well D7, 11.25x magnification

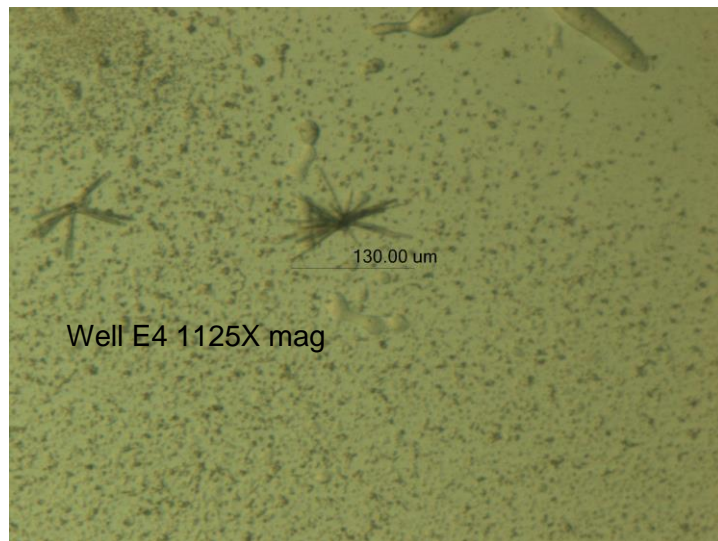


Figure 2.6: p97 ND1 (With Polyhistidine-tag), Well E4, 11.25x magnification

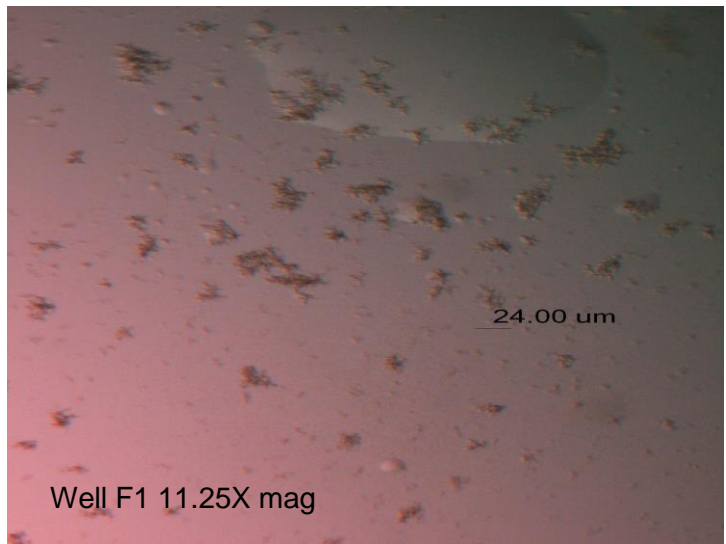


Figure 2.7: p97 Full Length, Well F1, 11.25x magnification

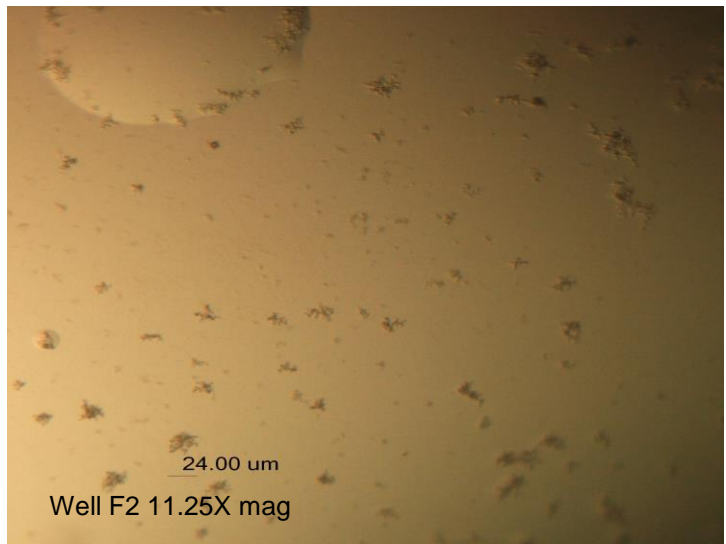


Figure 2.8: p97 Full Length, Well F2, 11.25x magnification

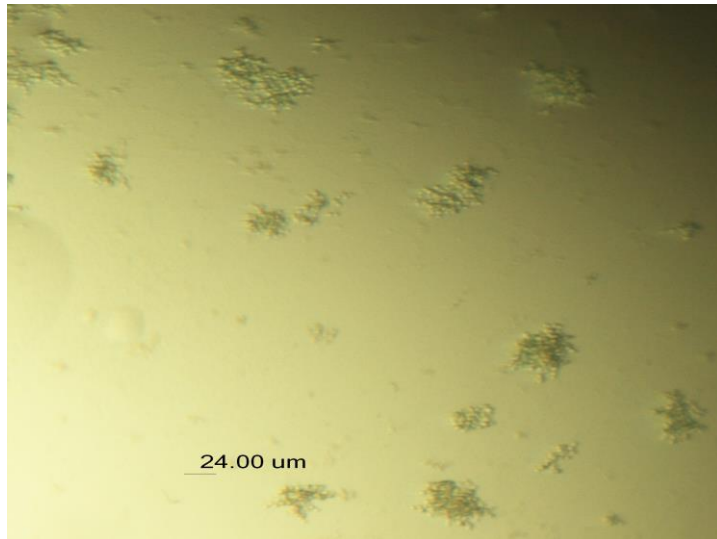


Figure 2.9: p97 Full Length, Well F5, 11.25x magnification

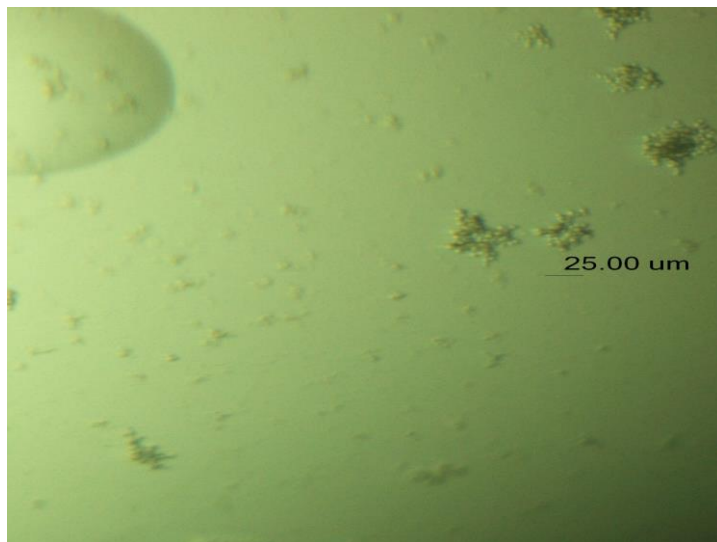


Figure 2.10: p97 Full Length, Well F6, 11.25x magnification

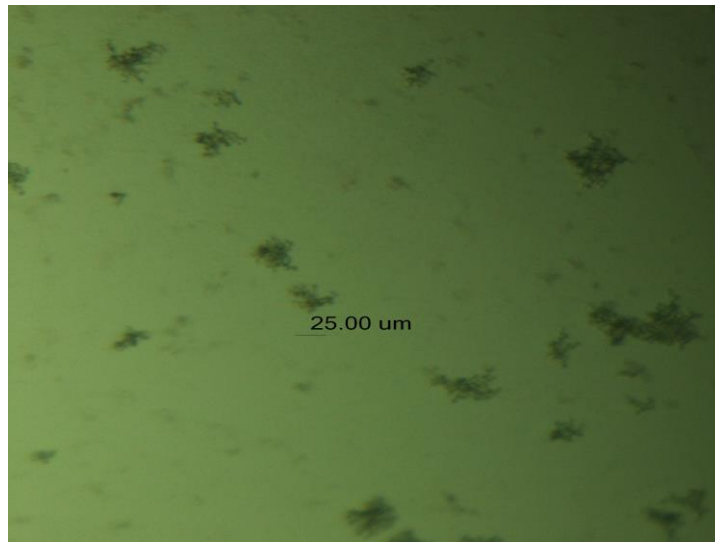


Figure 2.11: p97 Full Length, Well F7, 11.25x magnification

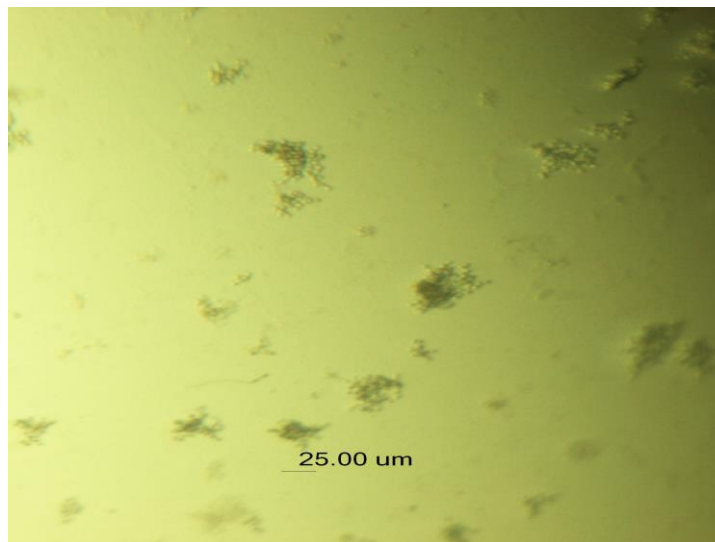


Figure 2.12: p97 Full Length, Well F8, 11.25x magnification

## REFERENCES

1. Tang, W.K., & Xia, D. Role of the D1-D2 Linker of Human VCP/p97 in the Asymmetry and ATPase Activity of the D1-domain. *Scientific Reports*. **6**. 20037-20037 (2016).
2. Tang, W.K., & Xia, D. Altered Intersubunit Communication Is the Molecular Basis for Functional Defects of Pathogenic p97 Mutants. *J.Biol.Chem.* **288**, 36624-36635 (2013).
3. Tang, W.K., Li,D., Li, C.C., Esser, L., Dai, R., Guo, L., Xia, D. A novel ATP-dependent conformation in p97 N-D1 fragment revealed by crystal structures of disease-related mutants. *Embo J.* **29**, 2217-2229 (2010).
4. Zhang, X., Shaw, A., Bates, P.A., Newman, R.H., Gowen, B., Orlova, E., Gorman, M.A., Kondo, H., Dokurno, P., Lally, J., Leonard, G., Meyer, H., Van Heel, M., Freemont, P.S. Structure of the N-Terminal domain and the D1 AAA domain of membrane fusion ATPase p97. *Mol. Cell.* **6**, 1473 (2000).
5. Hanzelmann, P., Schindelin, H. Structural Basis of ATP Hydrolysis and Intersubunit Signaling in the AAA+ ATPase p97. *Structure*. **24**, 127-139 (2016).
6. Davies, J.M., Brunger, A.T., Weis, W.I. Improved Structures of Full-Length p97, an AAA ATPase: Implications for Mechanisms of Nucleotide-Dependent Conformational Change. *Structure*. **16**, 715-726 (2008).

7. Huyton, T., Pye, V.E., Briggs, L.C., Flynn, T.C., Beuron, F., Kondo, H., Ma, J., Zhang, X., Freemont, P.S. The crystal structure of murine p97/VCP at 3.6Å. *J. Struct. Biol.* **144**, 337-348 (2003).
8. Resetar, A. M.; Chalovich, J. M. Adenosine 5'-(gamma-thiotriphosphate): an ATP analog that should be used with caution in muscle contraction studies. *Biochemistry.* **34 (49)**, 16039–16045 (1995).
9. Irving, R.A., Hudson, P.J., & Goding, J.W. Construction, Screening and Expression of Recombinant Antibodies. *Monoclonal Antibodies Third Edition*. Academic Press. 424-464 (1996).

## Chapter 3

### CRYSTALLIZATION OF P97 ND1 IN COMPLEX WITH SELENOPROTEIN S C TERMINAL PEPTIDE

#### 3.1 Introduction

##### 3.1.1 Previous Crystallization Conditions of Complexed p97

As was the case with the crystallization of uncomplexed p97 ND1 and full length p97, it was important to first compile previously published conditions for complexed p97. Those conditions could potentially serve as starting points for crystallization and eliminate the need for a tedious screening process. All available conditions found on PDB for variations of p97 in complex with other peptides were gathered and compared to one another (Table 3.1). This included the protein/peptide complex buffer solutions, pH, and reservoir buffer solutions. These conditions were then compared to previously successful crystallization conditions performed in our own laboratory for the uncomplexed full length p97 and p97 ND1. In Table 3.2, the relative concentrations of protein and peptide are listed, as well as the binding constants for the complexes and molar ratios of the protein to peptide.

The majority of published conditions for a p97 complex used the N-domain only<sup>1-7</sup>, with only one structure using the deleted loop variant of p97<sup>8</sup>, and two complexes of the p97 ND1 domain<sup>9</sup>. Overall, the majority of conditions appeared to follow comparable trends across similar structures, with the reservoir buffer being 100 mM, salt concentrations of approximately 200 mM, and a variation of PEG being used as the precipitant. Compared to previous conditions for uncomplexed p97, the incubation temperatures and buffer pH appeared to vary more, and ratios between protein and peptide also varied between complexes.

Because the initial crystallization trials were going to focus on the p97 ND1 domain in complex with C terminal SELENOS, a combination of successful in-house ND1 conditions and published ND1/peptide conditions were attempted.

### **3.1.2 C Terminal Selenoprotein S Peptide**

For SELENOS, studies show that the C terminus of the protein is the region responsible for binding, as discussed in section 1.2. For the purpose of crystallization, it was decided to focus on the C terminus initially, and attempt to co-crystallize p97 ND1 with a peptide of the SELENOS C terminus, utilizing the minimal binding regions for both proteins. The C terminus SELENOS (CSELENOS) peptide containing residues 173-189 was custom ordered from Genscript with the Sec in position 188 changed to a cysteine. The decision to use the cysteine mutant as opposed to the Sec-containing wild type was made because the two peptides should be structurally similar (the difference being the sulfur in cysteine taking the place of the selenium in selenocysteine), but the cysteine mutant is less likely to oxidize during crystallization.

### **3.1.3 Preparation of p97 ND1 and C Terminal Selenoprotein S Peptide Complex**

The first experiments focused on co-crystallizing the presumed binding region in p97, the ND1 domain, with the binding region in Selenoprotein S, the C-terminus. To do this, p97 ND1 was purified as outlined in Chapter 2.2 and concentrated to 7.0 mg/ml for screen 1 and 10mg/ml for screens 2 and 3. A 10 mg/ml stock solution of CSELENOS peptide 173-189 U188C was made, also in 25 mM Tris (pH 7.5) buffer with 300 mM NaCl, and then added to the protein solution to achieve protein to peptide ratios of either 1:2, 1:3, or 1:5, which were the most common ratios in published reports.

Since previous studies involving p97 and SELENOS binding show the complex forming under reducing conditions, reducing agents TCEP or Tris(hydroxypropyl)-phosphine (THP) were added to the protein/peptide solutions in various concentrations (Section 3.2.1), along with 0.5mM ADP, and the solutions were allowed to incubate at 4°C for one hour prior to laying crystal trays.

### **3.2 Crystallization of p97 ND1 in Complex with C Terminal SELENOS Peptide**

The sections below summarize the crystallization conditions and results in attempts to co-crystallize p97 ND1 with CSELENOS. All crystallization attempts utilized the hanging drop vapor diffusion method, carried out in 48 well crystallization plates. The complex and reservoir buffer mixtures were each suspended as a droplet on the coverslip in various drop ratios, inverted over each well, and sealed using the pre-greased trays. The trays were labelled and incubated at 16°C.

#### **3.2.1 Crystallization Conditions**

To begin, a coarse screening was done to obtain a very general view of how the complex might crystallize in conditions similar to those that were deemed successful in the initial p97 ND1 in-house trials.

For Screen 1 (Table 3.3), the protein/peptide complex was prepared as outlined in 3.1.2 with a protein to peptide ratio of 1:5, the highest peptide concentration deemed appropriate based on published conditions for p97 complexes. TCEP was used as the reducing agent and was added in two concentrations, 2mM and 5mM, to determine which reducing conditions would form the best complex. Then, mimicking conditions for previous ND1 screens, a 100mM sodium citrate buffer (pH 5.8) was used with NaCl concentrations of 0.1-0.3M. PEG 3350 was added as the precipitant in concentrations of

5-20%, the trypsin inhibitor, benzamidine, added in concentrations of 2-6%, and the cryoprotectant, glycerol, between 15-20%. This coarse screen also had a drop ratio of 1:1, meaning that the complex solution and reservoir buffer were added in equal volumes to the drop.

Screen 2 (Table 3.4) was another coarse screen. The protein to peptide molar ratio changed from 1:5 to 1:2 and 1:3, and THP at 2mM used as the reducing agent as opposed to TCEP to compare efficacies. The buffer solution stayed the same as in Screen 1, however the ranges for the PEG 3350 (0-30%), benzamidine (0-6%), and glycerol (0-10%) changed. Another change that was made was the ratio of complex solution and reservoir buffer in the hanging drop, shifting from the usual 1:1 ratio to a 1:2 and 1:3 protein to reservoir solution ratio. As discussed in Chapter 1.5, changes in drop ratios allows one to explore varying initial and final protein concentrations, meaning a wider range of coverage for relative supersaturation<sup>10</sup>.

The third screen, (Table 3.5), was a fine screen around the promising conditions in Screen 2. The protein/peptide solution was in a 1:2 molar ratio, with 2mM THP once again used as the reducing agent. The reservoir buffer was kept consistent, 0.1 M sodium citrate (pH5.8) with 0.1 M NaCl, while the PEG was in concentrations between 15-40%, benzamidine between 2-8%, and glycerol between 0-10%. Each well had a drop ratio of 1:3 protein solution to reservoir solution.

### **3.2.2 Crystal Tray Monitoring**

The crystal trays were monitored using an optical stereoscope, starting at Day 0, to observe the conditions and appearance of each drop at the time of incubation. The drops were monitored every two-three days within the first week, and then at minimum once a

week up to six weeks. Crystals showed no noticeable change between the first week and the third, however, typically after 3 weeks, the drops started to show slightly more amorphous precipitant and oiling. The trays were then checked sporadically for approximately one year. After several months, the crystals were smaller than previously observed and the drops showed an increase in salt crystals and amorphous precipitant. Crystal tracking sheets were utilized to note the appearance of each well and compare changes day to day. Evidence of phase separation, amorphous precipitants, salt crystals, or possible proteins crystals were all noted in the tracking sheets.

### **3.2.3 Results**

For the first coarse screen (Table 3.3), the initial screening at Day 0 already showed light precipitation in some of the wells. Between days 2-9, light to heavy precipitation was observed in all 48 wells, with the higher PEG and benzamidine concentrations also showing signs of phase separation (wells C4 and C8). There were no signs of protein crystals, and no images were taken. The high rate of precipitation could have been due to a number of factors, including high peptide concentrations and high initial protein complex concentrations in the hanging drops.

To try and minimize the appearance of amorphous precipitant, Screen 2 (Table 3.4) lowered the concentration of peptide in the complex from a 1:5 protein/peptide ratio to a 1:2 and 1:3 molar ratio. The drop ratios also changed, each coverslip having two drops per condition, at ratios of 1:2 and 1:3 protein complex solution to buffer solution. This immediately dilutes the protein solution in the drop, which should correspond to less of a chance for immediate precipitation of the protein in solution.

Day 0 observations for Screen 2 did not show immediate precipitation in any of the wells like was observed in Screen 1. Out of these conditions, the 1:3 drop ratio in well

C4 (30% PEG 3350; 0% Glycerol; 6% Benzamidine) showed signs of crystals by day 2, which did not grow in size after an additional several days of growing time. The 1:3 drop for well B7 (20% PEG 3350, 10% glycerol, 3% benzamidine) showed crystals mixed with salt clusters and oiling by day 6. The larger crystals in both drops were approximately 50 $\mu$ M in diameter. No wells for the 1:3 protein/peptide molar ratio conditions (rows D-F) showed signs of protein crystals, the majority of both drop ratios having amorphous precipitant or evidence of oiling, indicating a potentially still high peptide concentration.

Screen 3 was a fine screen around the successful conditions from Screen 2. The complex solution contained a 1:2 ratio of protein to peptide, and the drop ratio for each well was 1:3 protein solution to reservoir buffer, the reservoir buffer conditions as outlined in section 3.2.1 and Table 3.5. As was the case with Screen 2, no immediate precipitation was seen in any of the wells on day 0. By day 2, wells D3 (35% PEG 3350, 0% glycerol, 4% benzamidine) and D5 (25% PEG 3350, 10% glycerol, 4% benzamidine) both showed signs of crystals similar to those seen in screen 2. By day 6, wells B8 (30% PEG 3350, 10% glycerol, 3% benzamidine) and E1 (25% PEG 3350, 0% glycerol, and 6% benzamidine) also yielded crystals. This fine screen did not produce crystals of higher quality than screen 2, but it did show reproducibility amongst similar crystallization conditions, not only for the complex, but when compared to p97 ND1 itself (Table 3.6).

Due to the small size of these crystals, the biggest caveat was that there was no reliable way to definitely prove that the crystals were in fact of the p97 ND1/CSELENOS complex. They did not meet many of the guidelines when choosing an ideal crystal for diffraction, including well defined faces free from defects, opacity, and size. Ideal crystals should be close to 200 or 300 microns in size<sup>11</sup>, and as such, it was decided not to submit for X-Ray Diffraction until higher quality crystals were grown. Nor were the crystals

large enough to fully isolate the crystals and analyze by mass spectrometry. Currently, attempts are being made to better purify p97 ND1 and further study the p97/SELENOS complex independently. With purer protein and through a better understanding of optimal binding conditions for the complex, the hope is that diffraction quality protein crystals can be obtained in the future based on these initial screens.

### **Acknowledgements**

The author would like to thank Zhengqi Zhang, Vicki Wallace, and Rujin Cheng of the Rozovsky group for their work in expressing and purifying the various p97 proteins used for crystallization, as well as Mackenzie Lauro and Brian Bahnson for their assistance in the protein crystallography.

Table 3.1: Published Crystallization Conditions for complexed p97

PDB entry	Protein	Complexed with	Method	Temp (C)	pH	Reservoir Solution			
5EPP <sup>1</sup>	p97 N	RHBDL4 VBM	vapor diffusion sitting drop	20	6.5	0.1 M Bis Tris propane	20% w/v PEG 3350	200 mM sodium acetate	
3TIW <sup>2</sup>	p97 N	VIM of gp78	vapor diffusion hanging drop	25	7	100 mM Tris	18-22% PEG 3000		200 mM NaCl
3QQ8 <sup>3</sup>	p97 N	FAF1 UBX domain	vapor diffusion hanging drop	25	8.5	100 mM Tris	16-20% PEG MME 2000		200 mM trimethyl-N-oxide
5B6C <sup>4</sup>	p97 N	SHP box of Ufd1	vapor diffusion hanging drop	20	7.5	100 mM HEPES	21%(w/v) PEG 8000	10% glycerol	2.9mM n-nonyl- $\beta$ -thiomaltoside
5GLF <sup>5</sup>	p97 N	SHP box of derlin-1	vapor diffusion hanging drop	20	6.5	0.1 M Bis Tris propane	25% w/v PEG 3350		0.2 M MgCl <sub>2</sub>
4KDI <sup>6</sup>	p97 N	OTU1 UBXL	vapor diffusion hanging drop	22	5	10% Tacsimate	20% PEG 2000		
4KDL <sup>6</sup>	p97 N	OTU1 UBXL	vapor diffusion hanging drop	22	5	100 mM sodium acetate	14% PEG 4000		
3QC8 <sup>7</sup>	p97 N	FAF1 UBX domain	vapor diffusion hanging drop	25	8.5		25% w/v PEG 3350		200 mM MgCl <sub>2</sub>
5C1B <sup>8</sup>	p97 $\Delta$ 709-728	UFD1-SHP peptide	vapor diffusion hanging drop	25	5.75	0.4-0.6 M potassium acetate	6.5-7% PEG 4000	100 mM MES	
5IFS <sup>9</sup>	p97 ND1 (1-480)	ASPL- $\Delta$ C	Vapor diffusion sitting drop	4	7	0.1M HEPES/NaOH	23%PEG	0.3M LiSO <sub>4</sub>	
5IFW <sup>9</sup>	p97 ND1 (1-480)	ASPL-C	Vapor diffusion sitting drop	20	5.5	2M ammonium sulfate	0.1M bis-tris		

Table 3.2: Published p97 Complex Binding Conditions

PDB entry	Protein	Complexed with	Concentration of Protein	Concentration of peptide	Molar ratio of protein to peptide	Binding Kd
5EPP <sup>1</sup>	p97 N	RHBDL4 VBM	20 mg/mL; 0.96mM	1.92mM	1:2	N/A
3TIW <sup>2</sup>	p97 N	VIM of gp78	0.5 mM	1mM	1:2	21.3 ± 0.3 nM
3QQ8 <sup>3</sup>	p97 N	FAF1 UBX domain	10 mg/mL complex	Peptide concentration not specified	1:2	1.5 ± 0.42 uM
5B6C <sup>4</sup>	p97 N	SHP box of Ufd1	fusion at 20.7 mg/mL	Fusion used for crystallization	NA	221 ± 36 μM
5GLF <sup>5</sup>	p97 N	SHP box of derlin-1	20 mg/mL; 0.96mM	1.92mM	1:2	N/A
4KDI <sup>6</sup>	p97 N	OTU1 UBXL	20 mg/mL; 0.96mM	2.88mM	1:3	0.71uM
4KDL <sup>6</sup>	p97 N	OTU1 UBXL	20 mg/mL; 0.96mM	2.88mM	1:3	NA
3QC8 <sup>7</sup>	p97 N	FAF1 UBX domain	20 mg/mL; 0.96mM	1.92mM	1:2	11.2 μM
5C1B <sup>8</sup>	p97 Δ709-728	UFD1-SHP peptide	0.1 mM	0.5mM	1:5	3.5 uM
5IFS <sup>9</sup>	p97 ND1 (1-480)	ASPL-ΔC	12.3mg/ml complex	Peptide concentration not specified	1:3	0.2nM
5IFW <sup>9</sup>	p97 ND1 (1-480)	ASPL-C	15.5mg/ml complex	Peptide concentration not specified	1:3	NA

Table 3.3: Crystallization Conditions for Coarse Screen 1 (1:5 protein peptide ratio and 1:1 drop ratio)

Well	Buffer	pH	PEG %	Benzamidine %	Glycerol %
A1	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	5	2	15
A2	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	2	15
A3	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	2	15
A4	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	20	2	15
A5	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	5	2	20
A6	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	2	20
A7	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	2	20
A8	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	20	2	20
B1	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	5	4	15
B2	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	4	15
B3	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	4	15
B4	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	20	4	15
B5	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	5	4	20
B6	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	4	20
B7	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	4	20
B8	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	20	4	20
C1	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	5	6	15
C2	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	6	15
C3	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	6	15
C4	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	20	6	15
C5	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	5	6	20
C6	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	10	6	20
C7	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	15	6	20
C8	100mM Citrate Buffer; 100mM NaCl;7mg/ml protein	5.8	20	6	20
D1	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	5	2	15
D2	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	2	15
D3	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	2	15
D4	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	20	2	15
D5	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	5	2	20
D6	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	2	20
D7	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	2	20
D8	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	20	2	20
E1	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	5	4	15
E2	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	4	15
E3	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	4	15
E4	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	20	4	15

Well	Buffer	pH	PEG %	Benzamidine %	Glycerol %
E5	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	5	4	20
E6	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	4	20
E7	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	4	20
E8	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	20	4	20
F1	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	5	6	15
F2	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	6	15
F3	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	6	15
F4	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	20	6	15
F5	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	5	6	20
F6	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	10	6	20
F7	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	15	6	20
F8	100mM Citrate Buffer; 300mM NaCl;7mg/ml protein	5.8	20	6	20

Table 3.4: Crystallization Conditions for Coarse Screen 2 (10mg/ml protein, 1:2 and 1:3 drop ratio)

Well	Buffer	pH	PEG %	Benzamidine %	Glycerol %
A1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	0	0	0
A2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	10	0	0
A3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	0	0
A4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	0	0
A5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	0	0	10
A6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	10	0	10
A7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	0	10
A8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	0	10
B1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	0	3	0
B2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	10	3	0
B3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	3	0
B4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	3	0
B5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	0	3	10
B6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	10	3	10
B7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	3	10
B8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	3	10
C1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	0	6	0
C2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	10	6	0
C3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	6	0
C4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	6	0
C5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	0	6	10
C6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	10	6	10
C7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	6	10
C8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	6	10
D1	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	0	0	0
D2	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	10	0	0
D3	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	20	0	0
D4	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	30	0	0
D5	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	0	0	10
D6	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	10	0	10
D7	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	20	0	10
D8	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	30	0	10
E1	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	0	3	0
E2	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	10	3	0
E3	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	20	3	0
E4	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	30	3	0

Well	Buffer	pH	PEG %	Benzamidine %	Glycerol %
E5	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	0	3	10
E6	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	10	3	10
E7	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	20	3	10
E8	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	30	3	10
F1	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	0	6	0
F2	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	10	6	0
F3	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	20	6	0
F4	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	30	6	0
F5	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	0	6	10
F6	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	10	6	10
F7	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	20	6	10
F8	100mM Citrate Buffer; 100mM NaCl; 1:3 molar ratio	5.8	30	6	10

Table 3.5: Crystallization Conditions for Coarse Screen 2 (10mg/ml protein, 1:3 drop ratio)

Well	Buffer	pH	PEG %	Benzamidine %	Glycerol %
A1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	15	2	0
A2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	2	0
A3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	2	0
A4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	2	0
A5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	15	2	10
A6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	2	10
A7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	2	10
A8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	2	10
B1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	15	3	0
B2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	3	0
B3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	3	0
B4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	3	0
B5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	15	3	10
B6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	3	10
B7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	3	10
B8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	3	10
C1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	15	4	0
C2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	4	0
C3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	4	0
C4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	4	0
C5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	15	4	10
C6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	20	4	10
C7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	4	10
C8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	4	10
D1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	4	0
D2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	4	0
D3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	35	4	0
D4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	40	4	0
D5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	4	10
D6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	4	10
D7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	35	4	10
D8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	40	4	10
E1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	6	0
E2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	6	0
E3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	35	6	0
E4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	40	6	0

Well	Buffer	pH	PEG %	Benzamidine %	Glycerol %
E5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	6	10
E6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	6	10
E7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	35	6	10
E8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	40	6	10
F1	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	8	0
F2	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	8	0
F3	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	35	8	0
F4	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	40	8	0
F5	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	25	8	10
F6	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	30	8	10
F7	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	35	8	10
F8	100mM Citrate Buffer; 100mM NaCl; 1:2 molar ratio	5.8	40	8	10

Table 3.6: Comparison of successful in-house Crystallization conditions

Crystal	Buffer	Precipitant	Additives	
P97 ND1 (7mg/ml)	100mM Citrate Buffer (5.8) and 100mM NaCl	10% PEG 3350	25% Glycerol	4% Benzamidine
P97 ND1 (7mg/ml)	100mM Citrate Buffer (5.8) and 100mM NaCl	20% PEG 3350	15% Glycerol	2% Benzamidine
P97 ND1 (7mg/ml)	100mM Citrate Buffer (5.8) and 100mM NaCl	20% PEG 3350	20% Glycerol	2% Benzamidine
P97 ND1 + CSELENOS Peptide	100mM Citrate Buffer (5.8) and 100mM NaCl	20% PEG 3350	10% Glycerol	3% Benzamidine
P97 ND1 + CSELENOS Peptide	100mM Citrate Buffer (5.8) and 100mM NaCl	30% PEG 3350	0% Glycerol	6% Benzamidine
P97 ND1 + CSELENOS Peptide	100mM Citrate Buffer (5.8) and 100mM NaCl	30% PEG 3350	10% Glycerol	3% Benzamidine
P97 ND1 + CSELENOS Peptide	100mM Citrate Buffer (5.8) and 100mM NaCl	35% PEG 3350	0% Glycerol	4% Benzamidine
P97 ND1 + CSELENOS Peptide	100mM Citrate Buffer (5.8) and 100mM NaCl	25% PEG 3350	10% Glycerol	4% Benzamidine
P97 ND1 + CSELENOS Peptide	100mM Citrate Buffer (5.8) and 100mM NaCl	25% PEG 3350	0% Glycerol	6% Benzamidine

For P97 ND1 alone, the drop ratio was 1:1. DTT was used as the reducing agent

For p97 ND1 + CSELENOS peptide, the protein concentration was 10mg/ml, and there was a final protein to peptide ratio of 1:2. Drop ratio was 1:3. THP was used as the reducing agent.

### Crystal Images

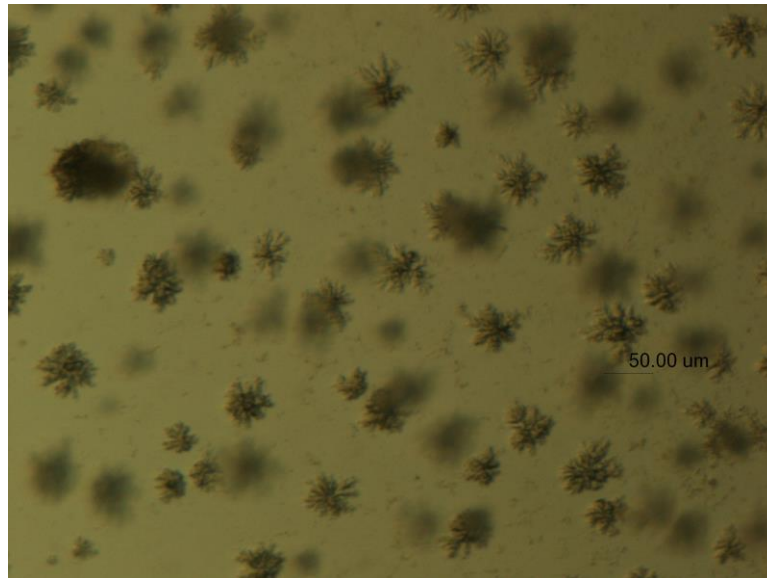


Figure 3.1: p97 ND1/CSELENOS Complex, Screen 2, Well C4, 11.25X Magnification

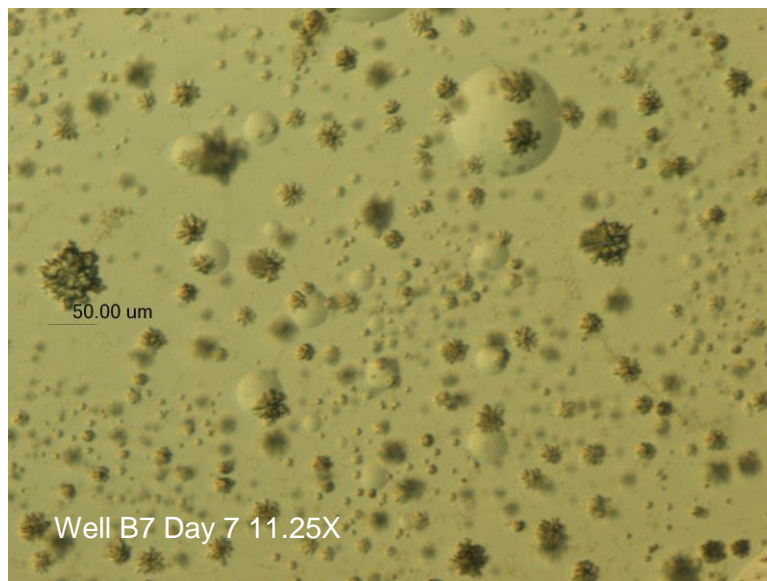


Figure 3.2: p97 ND1/CSELENOS Complex, Screen 2, Well B7, 11.25X Magnification

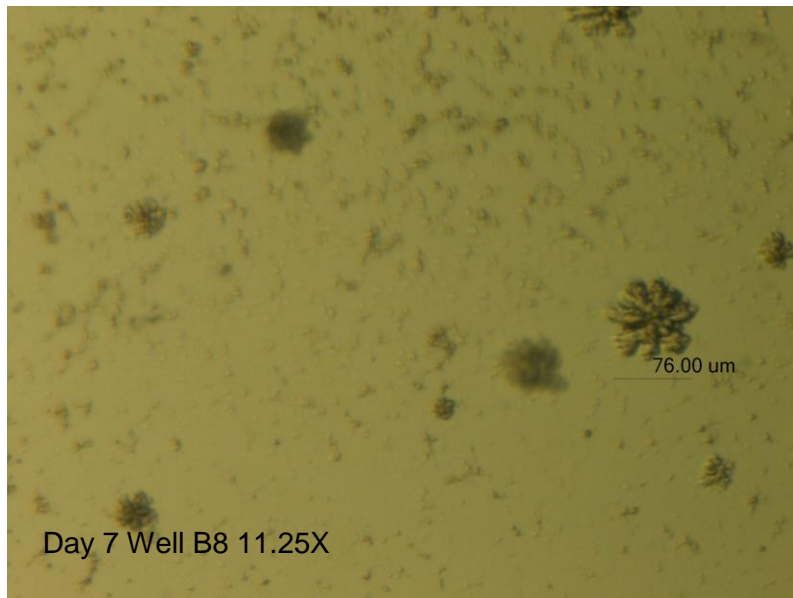


Figure 3.3: p97 ND1/CSELENOS Complex, Screen 3, Well B8, 11.25X Magnification

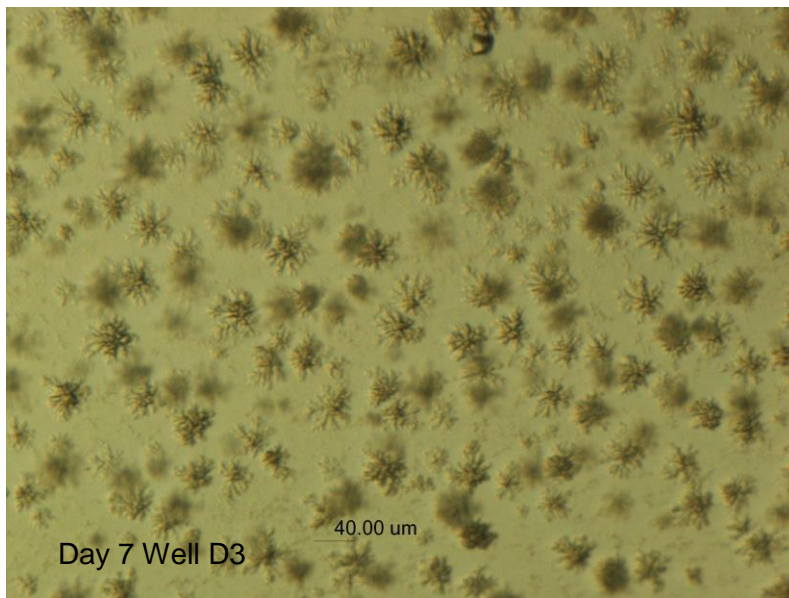


Figure 3.4: p97 ND1/CSELENOS Complex, Screen 3, Well D3, 11.25X Magnification

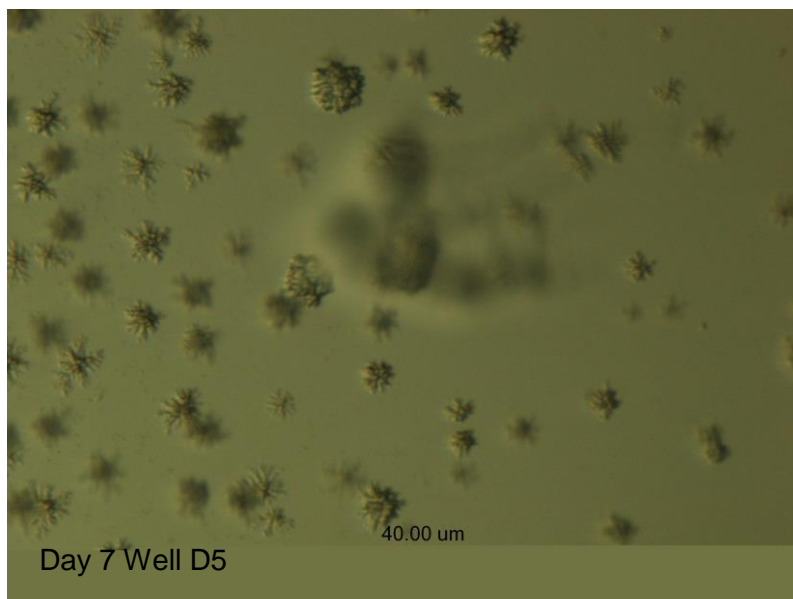


Figure 3.5: p97 ND1/CSELENOS Complex, Screen 3, Well D5, 11.25X Magnification

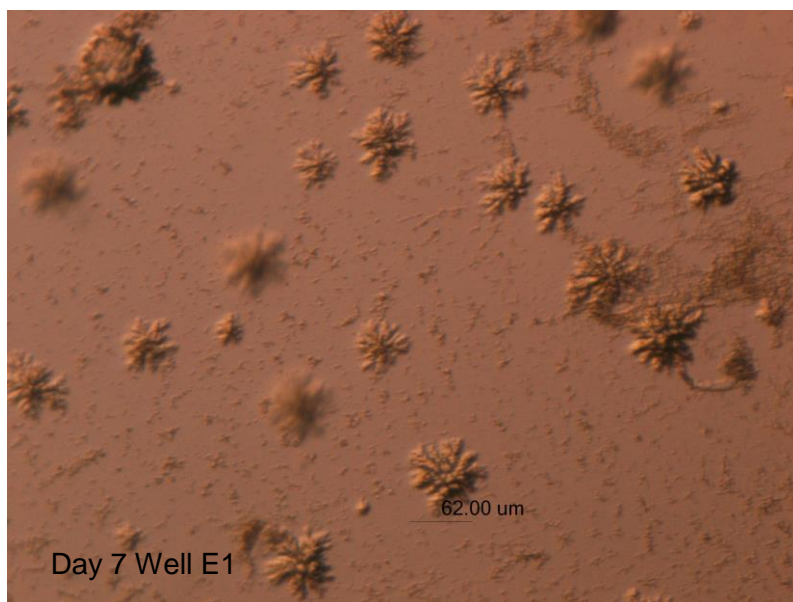


Figure 3.6: p97 ND1/CSELENOS Complex, Screen 3, Well E1, 11.25X Magnification

## REFERENCES

1. Lim, J.J., Lee, Y., Ly, T.T., Kang, J.Y., Lee, J.-G., An, J.Y., Youn, H.-S., Park, K.R., Kim, T.G., Yang, J.K., Jun, Y., Eom, S.H. Structural Insights into the Interaction of p97 N-terminus Domain and VBM Motif in Rhomboid Protease, RHBDL4. *Biochem. J.* **473**, 2863-2880 (2016).
2. Haenzelmann, P., Schindelin, H. Crystal structure of p97N in complex with the C-terminus of gp78. *J. Biol. Chem.* **286**, 38679-38690 (2011).
3. Hanzelmann, P., Buchberger, A., Schindelin, H. Hierarchical Binding of Cofactors to the AAA ATPase p97. *Structure.* **19**, 833-843 (2011).
4. Le, L.T.M., Kang, W., Kim, J.Y., Le, O.T.T., Lee, S.Y., Yang, J.K. Structural Details of Ufd1 Binding to p97 and Their Functional Implications in ER-Associated Degradation. *PLoS ONE.* **11**, e0163394-e0163394 (2016).
5. Lim, J.J., Lee, Y., Yoon, S.Y., Ly, T.T., Kang, J.Y., Youn, H.S., An, J.Y., Lee, J.G., Park, K.R., Kim, T.G., Yang, J.K., Jun, Y., Eom, S.H. Structural insights into the interaction of human p97 N-terminal domain and SHP motif in Derlin-1 rhomboid pseudoprotease. *FEBS Lett.* **590**, 4402-4413 (2016).
6. Kim, S.J., Cho, J., Song, E.J., Kim, S.J., Kim, H.M., Lee, K.E., Suh, S.W., Kim, E.E. Structural Basis for Ovarian Tumor Domain-containing Protein 1 (OTU1) Binding to p97/Valosin-containing Protein (VCP). *J. Biol. Chem.* **289**, 12264-12274 (2014).
7. Kim, K.H., Kang, W., Suh, S.W., Yang, J.K. Crystal Structure of FAF1 UBX Domain in Complex with p97/VCP N Domain Reveals a Conformational Change

- in the Conserved FcISp Touch-Turn Motif of UBX Domain. *Proteins* **79**, 2583-2587 (2011).
8. Hanzelmann, P., Schindelin, H. Characterization of an Additional Binding Surface on the p97 N-Terminal Domain Involved in Bipartite Cofactor Interactions. *Structure* **24**, 140-147 (2016).
  9. Arumughan, A., Roske, Y., Barth, C., Forero, L.L., Bravo-Rodriguez, K., Redel, A., Kostova, S., McShane, E., Opitz, R., Faelber, K., Rau, K., Mielke, T., Daumke, O., Selbach, M., Sanchez-Garcia, E., Rocks, O., Panakova, D., Heinemann, U., Wanker, E.E. Quantitative Interaction Mapping Reveals an Extended UBX Domain in ASPL That Disrupts Functional p97 Hexamers. *Nat. Commun.* **7**, 13047-13047 (2016).
  10. Luft, J.R., Wolfley, J.R., Said, M.I., Nagel, R.M., Lauricella, A.M., Smith, J.L., Thayer, M.H., Veatch, C.K., Snell, E.H., Malkiwski, M.G., DeTitta, G.T. Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature. *Protein Sci.* **16**, 715-722 (2007).
  11. Rhodes, G. *Crystallography Made Clear*, 3rd Edition. Associated Press, Elsevier. (2006).

## Chapter 4

### CRYSTALLIZATION OF GB1 VARIANTS

#### 4.1 Crystallization of GB1 Variants

Protein crystallography was utilized in an attempt to solve the structures of 6 GB1 variants, V29Sem, I6Sem, L5Sem, V39Sem, V34Sem, and V54Sem. Because  $^{77}\text{Se}$  has a nuclear spin of  $\frac{1}{2}$ , its presence in a protein or peptide allows one to observe its chemical environment via NMR spectroscopy<sup>1-2</sup>. This was attempted with the aforementioned GB1 variants, the subsequent crystals not only being useful in solid-state NMR studies, but those of high enough quality could be X-rayed and the diffraction data used to solve these, as of now, unknown structures. The following chapter outlines the conditions that yielded the highest quality GB1 crystals, and highlights which were then submitted for the initial round of x-ray diffraction.

##### 4.1.1 Crystallization Conditions

The screening conditions used for these experiments were based off successful crystallization conditions reported by Gallager<sup>3</sup>, Schmidt<sup>4</sup>, and Frank<sup>5</sup> (Table 4.1), and previously successful in-house crystals. Because prior crystallizations in-house suggested that there were little differences amongst variants, several screening conditions were attempted with multiple protein variants. The initial course screens looked at L5Sem and I6Sem, the proteins in concentrations of 8 mg/ml, 14 mg/ml, and 20 mg/ml in 25 mM sodium acetate buffer, pH 5.5, with 1-2 mM TCEP added as reducing agent for crystals grown under reducing conditions. The reservoir buffer was also a 25mM sodium acetate solution pH 4.5-5.5, with Isopropanol (IPA) and 2-Methyl-2, 4-Pentanediol (MPD) used as precipitants in concentrations ranges of 15-30% and 40-70% respectively (Table 4.2)

Based off of those conditions, we were able to perform finer screens, the most successful conditions being to highest concentration protein (20 mg/ml) in the pH range of 4.6-4.9, with IPA constant at 20% and MPD in the range of 45% -50% (Table 4.3).

All attempts were carried out in 48 well plates using the hanging drop vapor diffusion method. The reservoir buffer and precipitants were added to each well in the outlined concentrations for final volumes of 250uL, and a 1:1 drop ratio of protein solution to peptide (0.5 uL and 0.5 uL) on the cover slide before being suspended over the pre-greased wells. The trays were incubated at 16°C.

#### **4.1.2 Crystal Tray Monitoring**

The crystal trays were monitored using an optical stereoscope, starting at Day 0, to observe the conditions and appearance of each drop at the time of incubation. The drops were monitored every two-three days within the first week, and then at minimum once a week thereafter until harvesting, which was typically done within two weeks. In all cases, crystals appeared and reached their full size within the first week. Crystal tracking sheets were utilized to note the appearance of each well and compare changes day to day. Evidence of phase separation, amorphous precipitants, salt crystals, or possible proteins crystals were all noted in the tracking sheets.

#### **4.1.3 Results**

Initial screens of the variants showed that optimal growth conditions occurred in the pH ranges between 4.5-4.9, MPD ranges between 45-50%, IPA concentration between 20-25 appeared to show appreciable differences, and 20 mg/ml protein concentration yielded the largest crystals. This led to the fine screening conditions

(Table 4.3) used for growing the crystals that would ultimately be frozen and sent for x-Ray diffraction collection. Crystals typically started forming within 2 or 3 days, and were fully grown within one week.

For V29Sem, three separate conditions grew crystals that were frozen for x-ray diffraction, 45% MPD with pH 4.5 buffer solution, 46% MPD with 4.6 pH buffer solution, and 46% MPD at 4.5 pH buffer solution. An interesting phenomenon was observed in some trays where crystals initially grew in a “plate” formation, but after 2 or 3 days, changed to the needle morphology. The needle-like crystals were those ultimately chosen for diffraction.

In the case of I6Sem, crystals were grown under both oxidized (no added reducing agent), and reduced conditions (1-2 mM TCEP), and three crystals were frozen for future analysis: 46% MPD in pH 4.6 buffer under oxidized conditions, 46% MPD in 4.7pH buffer under oxidized conditions, and 48% MPD in 4.7 pH buffer under reduced conditions.

The L5Sem variants were also grown under oxidized and reduced conditions like I6Sem, and there were three conditions deemed the most successful: 46% MPD in pH 4.6 buffer under oxidized conditions, 46% MPD in pH 4.9 buffer under oxidized conditions, and 47% MPD in pH 4.6 buffer with 1mM TCEP.

V39Sem proved to be the most difficult to get high quality crystals, and both reduced and oxidized conditions were attempted. Only two conditions were deemed to grow high enough quality crystals, 45% MPD and 48% MPD in pH 4.8 buffer under oxidized conditions. Similarly, only two crystals were harvested for the A34Sem variant, 47% and 49% MPD in pH 4.9 buffer under reduced conditions. The V54Sem had one

condition with diffraction quality crystals, 46% MPD in pH 4.9 buffer under reduced conditions.

Overall, successful conditions varied only slightly amongst the different GB1 variants, and by using a consistent fine screen of relatively narrow parameters, crystals were able to grow for each variant. While multiple conditions were successful for each protein, these also varied only slightly. The V29Sem appeared to favor the lower MPD concentration and pH ranges, while the I6Sem and L5Sem variants favored 'mid-range' MPD concentrations and pH ranges under both reduced and oxidized conditions. V39Sem was the most successful under oxidized conditions at pH 4.8 exclusively. Similarly, A35Sem and V54Sem favored higher pH 4.9 exclusively with MPD ranges of 46%-49%.

#### **Acknowledgements**

The author would like to thank Qingqing Chen of the Rozovsky group for her work in expressing and purifying the GB1 variants used for crystallization and assistance in laying crystal trays, as well as Mackenzie Lauro, Ming Dong, and Brian Bahnson for their assistance in the protein crystallography and crystal harvesting.

Table 4.1: Previously Published Conditions for GB1

Reference	Protein Conditions		Well Conditions				
	Protein solution	pH	Well solution				pH
<b>Franks<sup>3</sup></b>	13mg/ml protein +10mM Tris- HCl and 25mM NaCl	7.5	4w/v% PEG 4000		100mM NaOAc		4.6
	16mg/ml protein +10mM Tris- HCl and 25mM NaCl	7.5	20w/v% PEG 8000;		100 mM NaOAc	150 mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	5.6
<b>Gallagher<sup>4</sup></b>	10 mg/ml in 50 mM NaOAc	4.5		70%MPD	25 mM NaOAc	300 mM MgSO	4
	10 mg/ml in 50 mM NaOAc	4.5	20% IPA	50% MPD	25 mM NaOAc	50 mM NaCl	4.5
<b>Schmidt<sup>5</sup></b>	10 mg/mL GB1 in 50 mM Na <sub>2</sub> HPO <sub>4</sub>	5.5	6% IPA	50% MPD	25 mM NaOAc	50 mM NaCl	4.5
	10 mg/mL GB1 in 50 mM Na <sub>2</sub> HPO <sub>4</sub>	5.5	20% IPA	50% MPD	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	150 mM NaCl	3.8

Table 4.2: Initial Coarse Screen Conditions for GB1 Variants

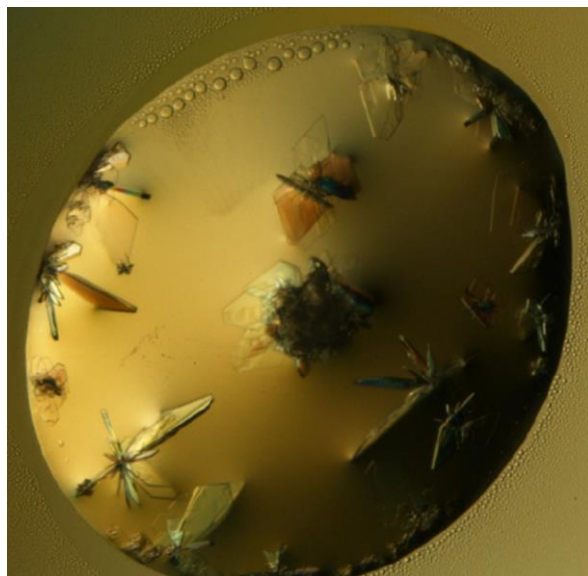
<b>Protein</b>	<b>Protein &amp; Reservoir buffer</b>	<b>pH</b>	<b>% IPA</b>	<b>% MPD</b>
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	15	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	15	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	15	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	15	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	25	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	25	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	25	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	25	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	30	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	30	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	30	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	30	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	15	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	15	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	15	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	15	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	20	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	20	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	20	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	20	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	25	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	25	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	25	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	25	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	30	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	30	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	30	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.0	30	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	5.5	15	40

<b>Protein</b>	<b>Protein &amp; Reservoir buffer</b>	<b>pH</b>	<b>% IPA</b>	<b>% MPD</b>
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	15	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	15	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	15	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	20	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	20	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	20	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	20	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	25	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	25	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	25	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	25	70
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	30	40
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	30	50
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	30	60
8, 14, or 20 mg/ml protein (1:1 drop ratio)	25 mM NaCH <sub>3</sub> CO <sub>2</sub>	5.5	30	70

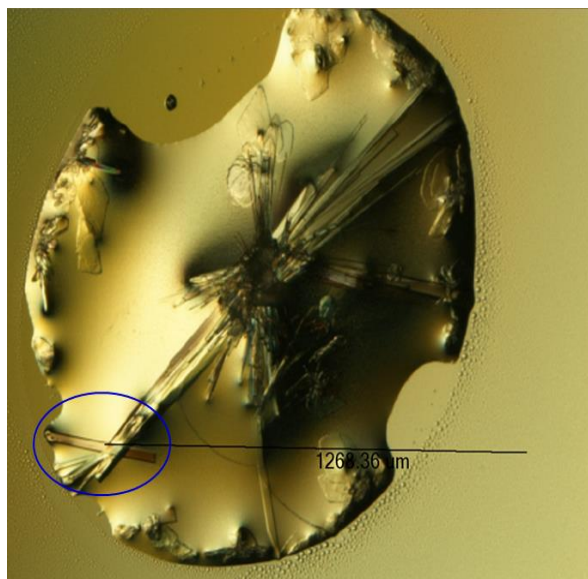
Table 4.3: Fine Screen Conditions used for All GB1 Variants

<b>Protein</b>	<b>Protein &amp; Reservoir buffer</b>	<b>pH</b>	<b>% IPA</b>	<b>% MPD</b>
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	45
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	46
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	47
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	48
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	49
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.5	20	50
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.6	20	45
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.6	20	46
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.6	20	47
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.6	20	48
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.6	20	49
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.6	20	50
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.7	20	45
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.7	20	46
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.7	20	47
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.7	20	48
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.7	20	49
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.7	20	50
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.8	20	45
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.8	20	46
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.8	20	47
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.8	20	48
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.8	20	49
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.8	20	50
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.9	20	45
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.9	20	46
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.9	20	47
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.9	20	48
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.9	20	49
20 mg/ml protein (1:1 drop ratio)	25 mM NaCH3CO2	4.9	20	50

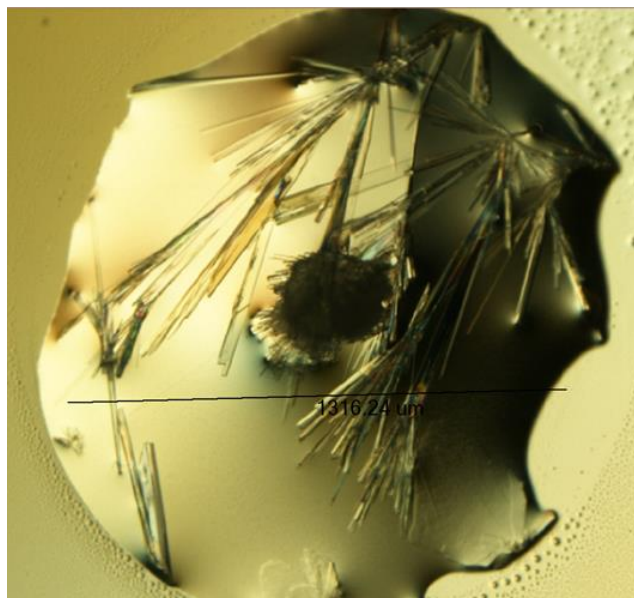
## Images of Select Crystals Harvested for Diffraction



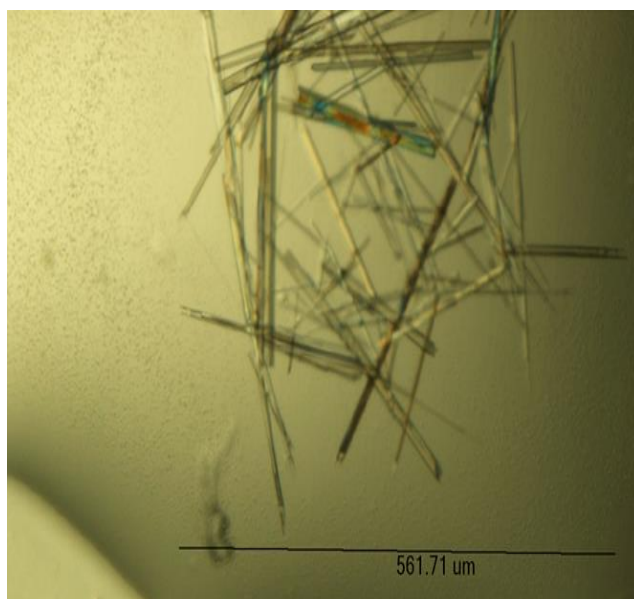
**Figure 4.1:** V29Sem Plate Morphology (45% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.5, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, 2 mM TCEP)



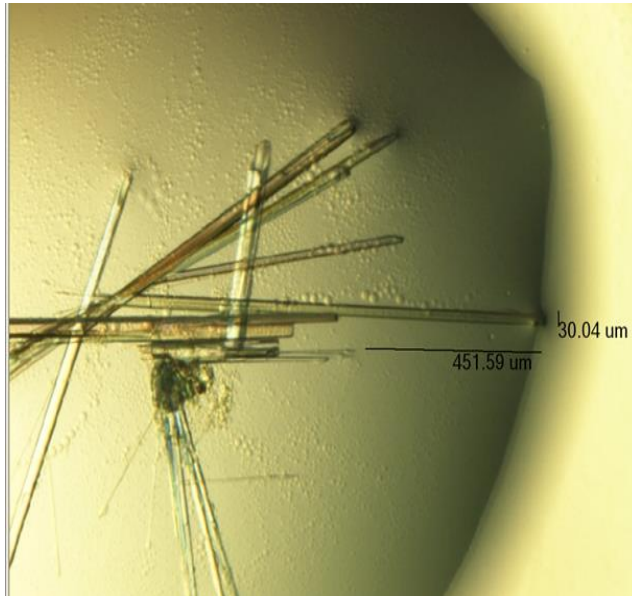
**Figure 4.2:** V29Sem Needle Morphology (45% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.5, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, 2 mM TCEP)



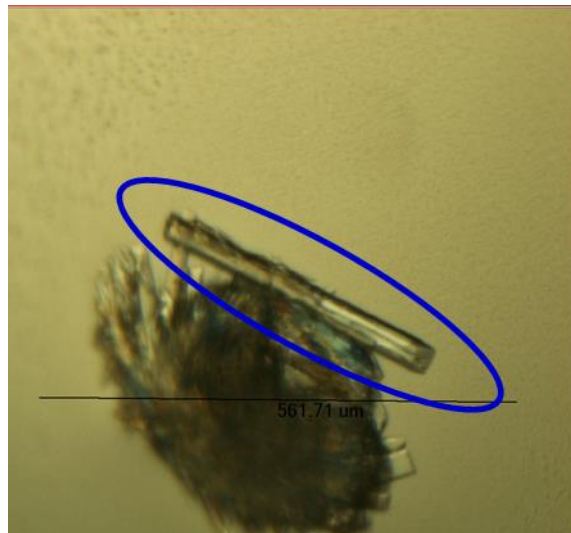
**Figure 4.3:** V29Sem (46% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.5, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, 2 mM TCEP)



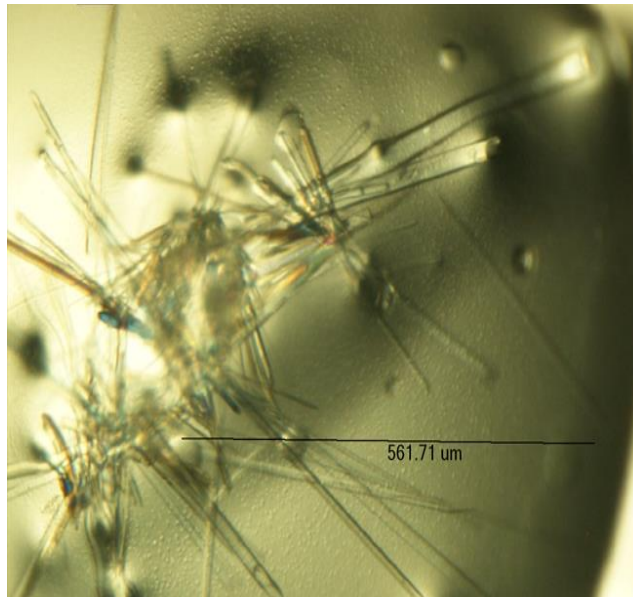
**Figure 4.4** I6Sem (46% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.7, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent)



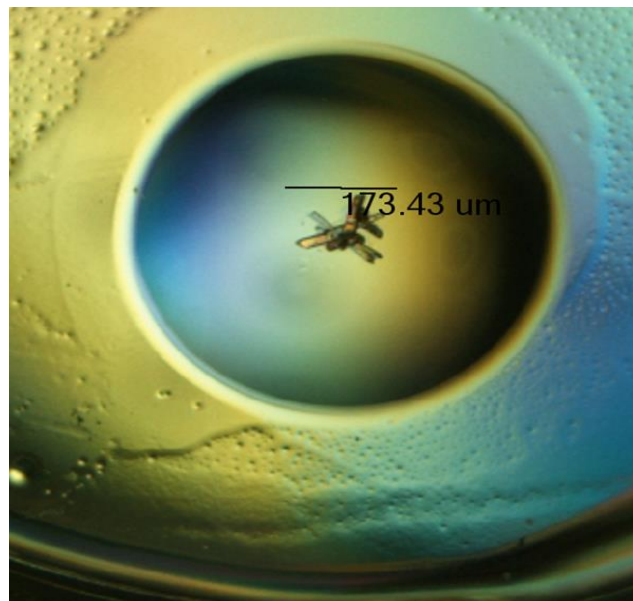
**Figure 4.5:** I6Sem (48% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.7, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, 1 mM TCEP)



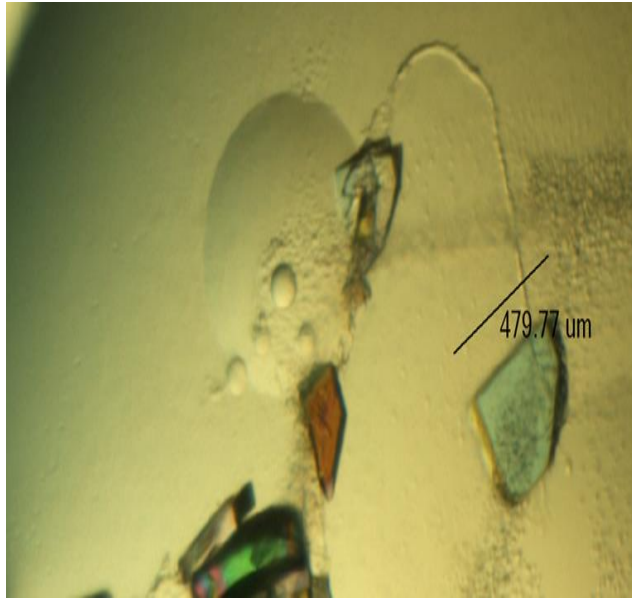
**Figure 4.6:** L5Sem (46% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.6, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent)



**Figure 4.7:** L5Sem (46% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.9, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent)



**Figure 4.8:** L5Sem (47% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.6, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent)



**Figure 4.9:** Representative Image of A34Sem (47% and 49% MPD, 20% IPA, 25 mM sodium acetate buffer pH 4.9, 20 mg/ml protein in 25 mM sodium acetate buffer, pH 5.5, No Reducing Agent)

## REFERENCES

1. Rozovsky, S., *<sup>77</sup>Se NMR Spectroscopy of Selenoproteins*. *Biochalcogen Chemistry: The Biological Chemistry of Sulfur, Selenium, and Tellurium*, Bayse, C. A.; Brumaghim, J. L., Eds. ACS Press: 127-142 (2013).
2. Liu, J., Rozovsky, S., *<sup>77</sup>Se NMR spectroscopy of selenoproteins*. *In Selenium Its molecular Biology and Role in Human Health*, Hatfield, D. L.; Gladyshev, V. N.; Schweizer, U.; Tsuji, P. A., Eds. Springer, 187-198 (2016).
3. Frank, M.K., Dyda, F., Dobrodumoc, A., Gronenborn, A.M. Core Mutations Switch Monomeric Protein GB1 into an Intertwined Tetramer. *Nature Str. Biol.* **11**, 877-885 (2002).
4. Gallagher, T., Alexander, P., Bryan, P., Gilliland, G.L. Two Crystal Structures of the B 1 Immunoglobulin-Binding Domain of Streptococcal Protein G and Comparison with NMR. *Biochem.* **33**, 4721-4729 (1994).
5. Frericks-Schmidt, H.L., Sperling, L.J., Gao, Y.G., Wylie, B.J., Boettcher, J.M., Wilson, S.R., Rienstra, C.M. Crystal Polymorphism of Protein GB1 Examined by Solid-State NMR Spectroscopy and X-ray Diffraction. *J. Phys. Chem.* **111**, 14362-14369 (2007).

## Chapter 5

### CONCLUSION AND FUTURE WORK

#### 5.1 Crystallization of p97/SELENOS Complex

This work laid out the groundwork for the use of protein crystallization in the structural determination of the p97/SELENOS complex. As a starting point, the desire was to complex the known binding regions of each component, the ND1 domain of p97, and the C terminus for SELENOS. For p97 ND1 in complex with the C terminus of Selenoprotein S, successful in-house conditions were first determined for both the full length p97 protein, as well as the ND1 domain independently. Based off of these findings, conditions were then carried over into the crystallization attempts of the ND1 domain in complex with CSELENOS.

The ND1 crystals were attempted first, both with the histidine affinity tag and without. For the Polyhistidine-tag free protein, the most successful conditions used a protein concentration of 7mg/ml as opposed to 14 mg/ml. 100 mM citrate buffer solution at pH 5.8 with 100 mM NaCl had fewer salt crystals than the 300mM salt conditions, and also had a higher rate of success in producing protein crystal. Using the fine screen conditions of PEG 3350 in concentrations between 10-20%, benzamidine between 1-3%, and glycerol between 15-30%, crystals were able to form. The best quality crystals formed were with the conditions of 20% PEG 3350, 15% glycerol, and 2% benzamidine, and the conditions of 20% PEG 3350, 20% glycerol, and 2% benzamidine. These are identical conditions save for the glycerol concentration, and since the glycerol serves primarily as a cryoprotectant, it is not surprising that slight variances in glycerol concentration had little to no impact on the crystal growth.

Because the ultimate goal of this work was to crystallize the ND1/CSELENOS complex, crystals of ND1 with the Histidine tag still intact were also attempted. This would prove beneficial if there were purification issues with the complex. If an additional step was needed to ensure purification of the complex, binding the SELENOS to a polyhistidine-tagged p97 would allow for purification of the complex through the metal affinity column. Knowing that the complex could be crystallized with the affinity tag still intact would keep this purification step as a possibility since the affinity tag cleavage step could be detrimental to the complex. The similar initial screens were performed on ND1 with the polyhistidine-tag as were done without, with the same buffer solution and salt ranges. PEG varied between 5-20%, benzamidine between 2-6%, and glycerol between 15-20%. Three distinct conditions produced crystals, and each varied in crystal shape. The conditions using 100 mM citrate buffer (pH 5.8), 100 mM NaCl, 10% PEG 3350, 15% glycerol and 6% benzamidine gave crystals of similar size and appearance as the ND1 crystals grown without an intact tag. The condition with higher precipitant and salt concentrations but lower benzamidine (100 mM citrate buffer (pH 5.8), 300 mM NaCl, 15% PEG 3350, 20% glycerol, 2% benzamidine) saw crystal clusters, while another condition very similar (100mM citrate buffer (pH 5.8), 300mM NaCl, 20% PEG 3350, 15% glycerol, 4% benzamidine), yielded needle-like crystals. Despite some differences in appearance for some crystals, this initial work shows that ND1 crystals are able to grow in relatively similar conditions whether the affinity tag is present or not.

Experiments to optimize conditions for the full length p97 were also performed in the event that the full length p97/SELENOS complex was more stable than using only the ND1 domain, or if it turns out that binding of SELENOS occurs somewhere other than in the ND1 domain of p97. Several trials were performed for the full length p97 on its own,

and these experiments demonstrated how crucial protein purity plays a role in ideal crystallization conditions. Trials 1 and 2 used identical conditions save for the protein concentration and purity, but crystals only grew in trial 2. Overall, there were 6 successful conditions in this trial, all using the buffer solution of 100 mM MES with 200 mM ammonium sulfate at pH 7 and PEG 4000 concentrations of 11, 12, 15, 16, 18, and 18%.

By combining previously published data and screening for p97 ND1 crystallization conditions, we were able to successfully grow crystals of what are believed to be the p97 ND1/CSELENOS complex. It was determined that optimal conditions for the protein/peptide complex were a protein/peptide molar ratio of 1:2, with a starting p97 ND1 concentration of 10 mg/ml, all under reducing conditions. These experiments also demonstrated how varying drop ratios in vapor diffusion crystallography can lead to crystal growth, with the initial complex concentration being diluted in a drop ratio of 1:3 protein complex solution and reservoir buffer. Overall, these initial experiments grew crystals under six separate conditions. The buffer solution remained constant (100 mM Citrate Buffer (5.8) and 100 mM NaCl), while the PEG 3350, glycerol, and benzamidine concentrations varied between 20-30%, 0-10%, and 3-6% respectively.

There were several challenges with these initial experiments, one being that the crystals grown were too small to be safely isolated and frozen for analysis. This meant that it was not possible to determine for sure that the protein crystals were indeed those of the p97 ND1/CSELENOS complex, or simply crystals of either protein or peptide grown independently. There were also questions of protein purity. It's been discussed at length how impurities can negatively impact optimal crystal growth, so attempts are still being made to better purify the protein. The p97/SELENOS complex is also still being studied

by our lab to better understand optimal binding conditions and form a protein/peptide complex ideal for crystallization.

Future work will build off of these initial crystallization conditions, and once a purer, more stable complex is made, more fine tuning around these successful parameters can begin. Once diffraction quality crystals are grown, they will be harvested, flash frozen in liquid nitrogen, and submitted for X-ray diffraction. The diffraction data will then, ideally, not only show whether or not the crystal is the complex, but through molecular replacement and structural refinement, the structure of the p97 ND1/CSELENOS complex can be determined.

## **5.2 Crystallization of Six GB1 Selenomethionine Variants**

Previous work by our lab has shown the successful incorporation of selenomethionine in various locations throughout GB1 to serve as probes of their local environment through spectroscopic studies. This also presented an opportunity to grow diffraction quality crystals to submit for X-ray analysis and solve the 3-dimensional structures for these previously unreported six variants.

There were only subtle differences in the crystallization conditions for the six variants, allowing for the same fine screen template to be applied across all variants successfully. All variants were equilibrated against 25mM acetate buffer between the pH ranges of 4.5-4.9, MPD in concentrations between 45-50%, and IPA at 20% consistently. Some variants, I6Sem and L5Sem, were able to grow under both reduced and oxidized conditions, while V39Sem only grew in oxidized conditions. Crystals of each variant deemed suitable for X-ray diffraction were harvested, flash frozen, and submitted for analysis.

Future work on this project will include analyzing the diffraction data, first to determine if the crystals were high enough quality, and if they are, to solve the crystal structures. For crystals not of high enough quality, these conditions will once again be used as screens to grow new crystals for analysis.