

**FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF NEUTRAL  
CHOLESTEROL ESTER HYDROLASE 1**

by

Meghan May Klems

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

Summer 2016

© 2016 Meghan May Klems  
All Rights Reserved

ProQuest Number: 10192368

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10192368

Published by ProQuest LLC (2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346

**FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF NEUTRAL  
CHOLESTEROL ESTER HYDROLASE 1**

by

Meghan May Klems

Approved: \_\_\_\_\_  
Murray V. Johnston, 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

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

Signed: \_\_\_\_\_  
Brian J. Bahnson, Ph.D.  
Professor in charge of dissertation

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

Signed: \_\_\_\_\_  
Sharon Rozovsky, Ph.D.  
Member of dissertation committee

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

Signed: \_\_\_\_\_  
John Koh, Ph.D.  
Member of dissertation committee

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

Signed: \_\_\_\_\_  
Wilfred Chen, Ph.D.  
Member of dissertation committee

## ACKNOWLEDGEMENTS

Obtaining a graduate degree is a process not only taken upon by the individual but by those that surround them. I would like to take a moment to acknowledge some of the individuals who have helped me achieve this goal.

To begin, I would like to recognize my advisor, Dr. Brian Bahnsen. You have provided encouragement and support, thank you for giving me the freedom to explore while providing the safety net to my inabilities. I would also like to specifically recognize the generosity and tremendous help of Dr. Sharon Rozovsky, who not only advised me in biochemistry but provided a mentor in every aspect of my graduate career far and beyond what anyone could ask from a committee member. There were also a number of UD faculty that were exceptionally helpful throughout my time at the university. My committee members, Dr. John Koh and Wilfred Chen, provided ideas and feedback critical to this dissertation. Dr. Thorpe, Dr. Grimes, and Dr. Zhuang also provided valuable experience, insights, and time that was both greatly appreciated and thought provoking.

Dr. Paul Silver has also been one of my most valued friends and mentors within the University of Delaware. His enthusiasms for teaching is contagious, and I could never truly express my gratitude for all that he has done for me.

I am greatly appreciative of my labmates, both past and present, including Dr. Prabha Srinivasan, Dr. Elizabeth Monillas, Dr. Stephanie Scott, Dr. Anastasia Thevenin, Dr. Ming Dong, Dr. Tara Gonzalez, Dr. Lynn Cadieux, Dr. Benjamin Israel, Dr. Stephanie Ramadan, Dr. Lisa Steinberg, Mackenzie Lauro, Michael Wilburn, Shilin Xie, and Walter Drake.

I have also been extraordinarily blessed by family and friends. This is not a process that can be completed alone and I would not be who I am today without the love and influence of my brothers Timothy and Thomas Shields, or their wives Christina and Antoinette. Thank you for your love and support, and most of all for picking up the phone when I needed you. To my parents, my success is your success. I owe everything I have to you. From a young age you encouraged me to pursue my dreams, and I have never wanted something so much as to make you proud.

To my husband, Joe, words cannot even begin to describe how much your love and support has meant to me. You are the reason I am here, the reason I stayed, and the reason I got up after each fall. Thank you for always being exactly what I needed, and for never losing the confidence that I lacked in myself. This is as much of your accomplishment as it is my own.

This manuscript is dedicated to my children

Jacob Alexander and Elizabeth Mae

## TABLE OF CONTENTS

LIST OF TABLES .....	ix
LIST OF FIGURES .....	x
LIST OF ABBREVIATIONS .....	xii
ABSTRACT .....	xiv

### Chapter

1	INTRODUCTION .....	1
1.1	The Role of Neutral Cholesterol Ester Hydrolase 1 in Cancer and Atherosclerosis .....	1
1.2	Inhibition of NCEH1 by Organophosphorus Compounds .....	4
1.3	Known Structural Components of NCEH1 .....	5
1.4	Summary and Thesis Direction .....	9
	REFERENCES .....	12
2	PURIFICATION OF NEUTRAL CHOLESTEROL ESTER HYDROLASE 1 FROM <i>ESCHERICHIA COLI</i> .....	15
2.1	Introduction .....	15
2.2	Materials and Methods .....	16
2.2.1	PCR Primers Utilized for <i>E. coli</i> Expression .....	16
2.2.2	Cloning into pET100/D-TOPO, pET20b+, and Expression of NCEH1 .....	17
2.2.3	Cloning into pGEX-4T3, Expression and Purification of NCEH1 .....	20
2.2.4	Cloning into pMAL-C5X, Expression and Purification of NCEH1 .....	22
2.2.5	Activity Assays of Neutral Cholesterol Ester Hydrolase 1 .....	23
2.2.6	Protein Crystallization Screening of NCEH1 .....	24
2.2.7	Random Mutagenesis Protocol .....	24
2.3	Results and Discussion .....	25
2.3.1	pET100/D-topo and pET20b+ Expressions .....	25
2.3.2	Expression and Purification of pGEX-4T3 .....	27
2.3.3	Expression and Purification of pMAL-C5x .....	31
2.3.4	Functional Characterization of NCEH1 .....	33
2.3.5	Protein Crystallization Screening of NCEH1 .....	37
2.3.6	Random Mutagenesis .....	38

2.4	Conclusions .....	41
	REFERENCES .....	44
3	PURIFICATION OF NCEH1 FROM <i>PICHA PASTORIS</i> .....	46
3.1	Introduction .....	46
3.2	Materials and Methods .....	47
3.2.1	Preparation of Competent <i>Pichia Pastoris</i> Cells.....	47
3.2.2	Cloning of NCEH1 into pPICZ $\alpha$ and pPICZ.....	47
3.2.3	Glycosylation Mutagenesis in pPICZ $\alpha$ .....	49
3.2.4	Expression of NCEH1 in <i>Pichia pastoris</i> .....	50
3.2.5	Purification of NCEH1 in <i>Pichia pastoris</i> .....	51
3.2.5.1	pPICZ .....	51
3.2.5.2	pICZ $\alpha$ .....	52
3.2.6	Functional Characterization of NCEH1 .....	52
3.2.7	Structural Characterization of NCEH1 .....	53
3.3	Results and Discussion .....	54
3.3.1	Expression Optimization of NCEH1 .....	54
3.3.2	Purification Optimization of NCEH1 from pPICZ $\alpha$ .....	57
3.3.3	Functional Characterization of NCEH1 from pPICZ.....	64
3.3.4	Crystallization Screening of NCEH1 from pPICZ.....	66
3.3.5	Glycosylation Mutagenesis of NCEH1 in pPICZ $\alpha$ .....	70
3.4	Conclusions .....	75
	REFERENCES .....	79
4	CONCLUSIONS AND FUTURE DIRECTIONS .....	80
4.1	NCEH1 Expression and Purification from <i>E. Coli</i> .....	80
4.2	<i>Pichia Pastoris</i> .....	82
4.3	Future Directions .....	84
4.3.1	<i>Escherichia coli</i> Expression .....	84
4.3.2	<i>Pichia Pastoris</i> .....	85
4.3.3	Other Expression Systems.....	87
	REFERENCES .....	89

Appendix

A	RANDOM MUTAGENESIS OF PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE .....	90
A.1	Introduction .....	90
A.2	Materials and Methods .....	93
A.2.1	Random Mutagenesis .....	93
A.2.2	Subcloning of PAFAH into the pINCOP Vector. ....	93
A.2.3	Screening of Mutagenic PAFAH.....	94
A.2.4	Subcloning of Mutagenic DNA into pGEX-6P1 .....	95
A.3	Results and Discussion .....	96
A.4	Conclusions and Future Directions .....	102
	REFERENCES .....	103
B	LICENSE AGREEMENTS.....	105

## LIST OF TABLES

<b>Table 2.1:</b> The PCR primers utilized in the amplification of NCEH1 for <i>E. coli</i> . .....	17
<b>Table 3.1:</b> Primers utilized in the cloning of NCEH1 into <i>P. pastoris</i> . .....	48
<b>Table 3.2:</b> Table showing the PCR mutagenic primers for the N to Q mutagenesis... 50	
<b>Table 3.3:</b> Table of crystal conditions used for the optimization of crystallization of NCEH1. ....	53
<b>Table 3.4:</b> Crystallography solutions yielding microcrystal .....	66

## LIST OF FIGURES

<b>Figure 1.1:</b> The ester hydrolysis reaction catalyzed by NCEH1. ....	2
<b>Figure 1.2:</b> The sequence alignment of NCEH1.....	6
<b>Figure 1.3:</b> Transmembrane prediction of NCEH1. ....	8
<b>Figure 1.4:</b> Helical wheel projection of the second hydrophobic sequence in NCEH1. ....	9
<b>Figure 2.1:</b> Summary of the NCEH1 constructs created for recombinant expression.....	26
<b>Figure 2.2:</b> SDS PAGE analysis of the purification of GST fused NCEH1 truncation 56-440. ....	30
<b>Figure 2.3:</b> The NCEH1 catalyzed reaction of the modified natural substrate 2-thio-acetyl MAGE with water. ....	34
<b>Figure 2.4:</b> NCEH1 catalyzed ester cleavage of para-nitrophenyl acetate to para-nitrophenol and acetic acid.....	35
<b>Figure 2.5:</b> Example of the functional characterization of NCEH1 on the substrate PNPA.....	36
<b>Figure 2.6:</b> MBP_115-440 NCEH1 fusion microcrystals.. ....	38
<b>Figure 2.7:</b> Hydrolysis activity of PNPA in mAbs/min by NCEH1.....	40
<b>Figure 3.1:</b> NCEH1 constructs developed for expression in <i>Pichia pastoris</i> .....	54
<b>Figure 3.2:</b> Time course expression of <i>Pichia pastoris</i> . ....	57
<b>Figure 3.3:</b> SDS PAGE analysis of a purification of NCEH1 from pPICZ $\alpha$ . ....	59
<b>Figure 3.4:</b> SDS analysis of a purification of 111-440 NCEH1 from pPICZ $\alpha$ .....	60
<b>Figure 3.5:</b> SDS PAGE analysis of purified NCEH1 from the pPICZ expression vector.....	62
<b>Figure 3.6:</b> SDS PAGE analysis of the purification of NCEH1 using Ni-sepharose, cobalt, and Ni-NTA resins.....	63
<b>Figure 3.7:</b> The structure of the inhibitor JW480.. ....	65

<b>Figure 3.8:</b> SDS PAGE gel representation of the NCEH1 sample used to lay hanging drop crystal screens 1 and 2. ....	68
<b>Figure 3.9:</b> Crystal resulting from 0.2 M sodium acetate trihydrate, 0.1 Tris hydrochloride pH 8.5, 30% w/v PEG 4,000, 2.5 mM ATP.....	69
<b>Figure 3.10:</b> SDS PAGE analysis of the purifications of NCEH1 glycosylation mutagenesis from pPICZ $\alpha$ . ....	73
<b>Figure 3.11:</b> SDS PAGE analysis of the NCEH1 triple mutant.. ....	75
<b>Figure A.1:</b> The hydrolysis of PAF by PAFAH. ....	90
<b>Figure A.2:</b> Serine hydrolase reaction with organophosphates. ....	91
<b>Figure A.3:</b> The colorimetric reaction of PAFAH.....	100

## LIST OF ABBREVIATIONS

AADACL1	arylacetamide deacetylase-like 1
ABPP	Activity based protein profiling
ATP	Adenosine-5'-triphosphate
BCA	Bicinchoninic acid
CE	Cholesterol Esters
DTNB	5'5-dithiobis(2-nitrobenzoate)
EDTA	Ethylenediamine-N, N, N', N'-tetraacetic acid
FPLC	Fast protein liquid chromatography
GST	Glutathione S-transferase
INP	Ice nucleation protein
IPTG	Isopropyl- $\beta$ -d-thiogalactopyranoside
LB	Lysogeny Broth
MAGE	Monoalkylglycerol ether
NCEH1	Neutral Cholesterol Ester Hydrolase 1
Ni-NTA	Nickel-Nitrotriactic acid
OP	Organophosphorus compound
OPH	Organophosphate hydrolase
PAF	Platelet-activating factor,
PAF-AH	Platelet-activating factor acetylhydrolase
PCR	Polymerase chain reaction
PKC	Protein Kinase CF
PNPA	4-nitrophenyl acetate

rSAP        Shrimp alkaline phosphatase

SDS-PAGE Sodium dodecyl sulfate–polyacrylamide gel electrophoresis

TEV        Tobacco Etch Virus

## ABSTRACT

Neutral cholesterol ester hydrolase 1 (NCEH1), also known as KIAA1363 and arylacetamide deacetylase-like 1 (AADACL1), has been found to be upregulated in several invasive cancers. The ability of this enzyme to act as a biomarker and possible therapeutic target increased the desire for research that elucidated the endogenous substrate of NCEH1, which was initially demonstrated to be 2-acetyl monoalkylglycerol ether (2-acetyl MAGE), a *de novo* precursor to platelet-activating factor. Later work also demonstrated a controversial role of NCEH1 in the hydrolysis of cholesterol esters in macrophages, preventing the rupture of foam cells and thereby preventing atherosclerosis. In recent years, research on NCEH1 has continued to portray the enzyme as a pro-tumorigenic, anti-atherosclerotic, regulatory enzyme. However, its role has also expanded to include participation in the regulation of platelet aggregation, and it has surfaced as a potential protein engineering target due to its ability to hydrolyze toxic organophosphate compounds. The structural knowledge of NCEH1 is limited to structures of homologous serine hydrolases, the existence of an N-terminal transmembrane helix and posttranslational glycosylation predicted to occur at residues 270, 287, and 389.

A variety of fusion *E. coli* constructs were created using a truncated NCEH1 which removed the transmembrane helix. Further truncation removing a second hydrophobic region was found to have no effect on solubility. It was determined that NCEH1 requires a large fusion protein, like MBP, to help with the stability of the enzyme, and cleavage of MBP fusion resulted in degradation of the protein. Functional characterization of the NCEH1 fusion protein was carried out with para-nitrophenyl acetate and a modified natural substrate 2-thio-acetyl MAGE. Also

reported within this work are the first reported  $K_M$  and  $k_{cat}$  values for the hydrolysis of PNPA by NCEH1, determined to be 5.3 mM and 4 s<sup>-1</sup>, respectively.

In an effort to isolate purified enzyme lacking a fusion protein partner, constructs of the 111-440 NCEH1 truncation were developed for expression in the eukaryotic yeast system *Pichia pastoris*. Both intracellular and secreted expression protocols were developed and generated a stable construct with a small histidine tag. The construct was further modified to include a TEV protease cleavage site to allow for the removal of the histidine tag after purification and both expression protocols yielded purified stable NCEH1. The enzyme was recovered in high purity and adequate concentration for the screening of optimal crystallization conditions providing an important first step toward the structural elucidation of NCEH1.

Previous research has indicated glycosylation for functional activity; however the work of this thesis provides an alternative view of the need for glycosylation. Instability of NCEH1 in an *E. coli* expression system supports a more structural role for glycosylation over functional. Mutation of the three glycosylation residues in the *Pichia pastoris* secretory expression system, without a loss in functional ability, also demonstrates the possible need for glycosylation in folding and stability.

The work compiled in this thesis advances our knowledge of NCEH1. The development of an expression and purification protocol generating a purified enzyme advances the ability to functionally characterize NCEH1. In addition, preliminary crystallization screens provided three potential conditions for future protein crystal optimization. The NCEH1 structure will ultimately further our knowledge and ability to elevate the treatments of diseases directly correlated to its function.

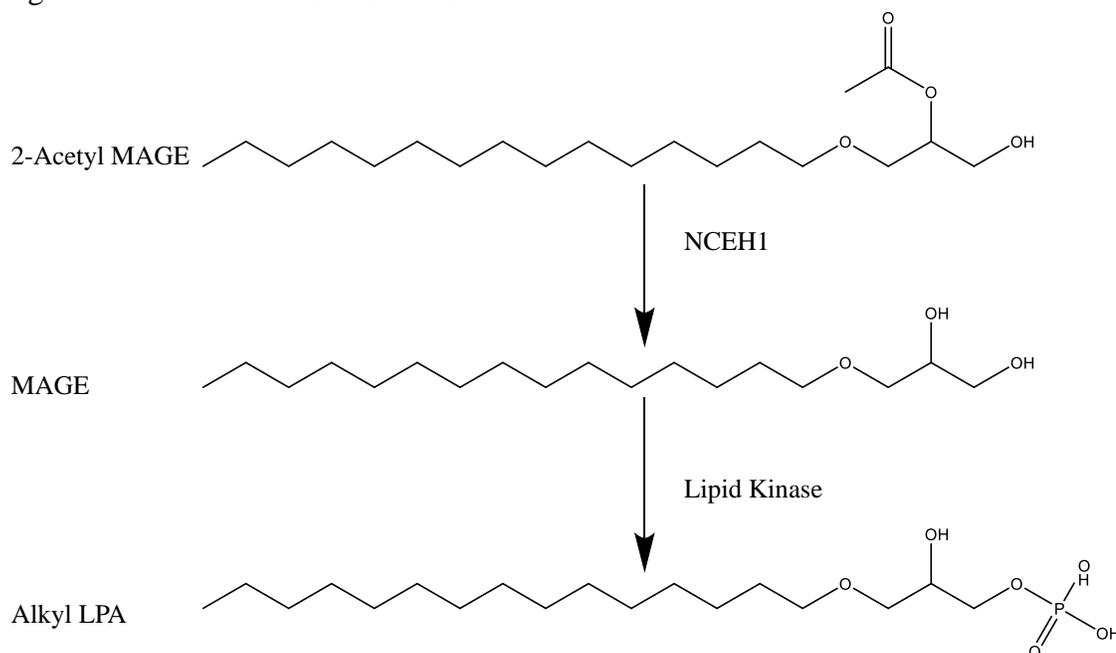
## Chapter 1

### INTRODUCTION

#### 1.1 The Role of Neutral Cholesterol Ester Hydrolase 1 in Cancer and Atherosclerosis

Neutral cholesterol ester hydrolase 1 (NCEH1), also known as KIAA1363 and arylacetamide deacetylase-like 1 (AADACL1), was first discovered by Dr. Benjamin Cravatt via activity based protein profiling (ABPP). The purpose of ABPP is to develop probes that bind to the active site of a protein. ABPP probes can be engineered to be general, and recognize either large classes of proteins like serine hydrolases, or they can be highly specific and recognize more individual targets such as NCEH1. In addition to a reactive region, each probe also contains a region that can be used to detect its presence, in the form of a fluorophore or biotin.<sup>1</sup> Jessani *et al.* used ABPP to visualize serine hydrolase activity in several different human cancer cell lines. Their results indicated an upregulation of NCEH1 in human breast, melanoma, and ovarian cancers.<sup>2</sup> Additional studies linked NCEH1 to aggressive forms of brain, pancreatic, and prostate cancers, making the elucidation of its natural substrate critical to understanding the enzyme's link to tumorigenesis.<sup>2-5</sup> Chiang *et al.* developed a selective inhibitor using ABPP which aided in the discovery that 2-acetyl monoalkylglycerol ether (2-acetyl MAGE), a precursor to the de novo synthesis of platelet activating factor, was the endogenous substrate of NCEH1.<sup>3</sup> The predicted reaction, shown in Figure 1.1, cleaves 2-acetyl MAGE to produce a C16 MAGE. Figure 1.1 also illustrates how the MAGE product of the NCEH1 reaction can be

further modified by phosphorylation to create alkyl lysophosphatidic acid (Alkyl LPA), which is a known modulator of tumor invasiveness.<sup>6</sup> Disruption of this reaction reduced cancer cell migration and tumor growth in vivo, making NCEH1 a therapeutic target of interest for the treatment of several invasive cancers.



**Figure 1.1:** The ester hydrolysis reaction catalyzed by NCEH1. NCEH1 mediated cleavage of 2-acetyl MAGE to MAGE which can then be phosphorylated into alkyl-LPA a known modulator of tumor invasiveness.

As shown in Figure 1.1, up-regulation of NCEH1 leads to the formation of the pro-tumorigenic compound alkyl LPA. While the down-regulation of NCEH1 would prevent this reaction and cause an accumulation of 2-acetyl MAGE. The metabolite 2-acetyl MAGE is an immediate precursor to platelet activating factor (PAF), a potent lipid mediator that causes inflammation and leads to atherosclerosis, a leading cause of heart disease and stroke.<sup>3, 5, 7</sup> The significant point of this observation is that NCEH1 promotes tumor invasiveness, while simultaneously preventing atherosclerosis.

Further work performed in macrophages and megakaryocytes showed that NCEH1 was indeed involved in the prevention of atherosclerosis.<sup>7-10</sup> However, while the vast majority of the literature is in agreement that NCEH1 appears to have an anti-atherosclerotic effect, the mechanism of prevention remains controversial.

Atherosclerosis develops when foam cells, generated by the unlimited uptake of modified lipid droplets containing esterified cholesterol, rupture creating an immune response that leads to plaques within the arteries.<sup>10</sup> To combat this process, cholesterol esters (CE) must be hydrolyzed to form free cholesterol, which can be safely secreted from the cell. In 2008, Okazaki *et al.* determined that NCEH1 plays a role in the hydrolysis of cholesterol esters to free cholesterol. This work is where the name NCEH1 (neutral cholesterol ester hydrolase 1) was acquired.<sup>10, 11</sup>

In 2010, the results of the Okazaki study were challenged by Buchebner *et al.* when they were unable to replicate their results in mice.<sup>7</sup> However, recent studies have concluded that there are differences in the hydrolysis of CE among macrophages from different species leading to the conclusion that mice are not an appropriate model for human cholesterol esterase activity.<sup>8, 9, 12</sup> In 2011, Igarashi *et al.* utilized human monocyte-derived macrophages to study NCEH1 and concluded that it was responsible for the majority of the neutral CE hydrolase activity, reaffirming its role as a neutral cholesterol ester hydrolase.<sup>9</sup>

To further complicate the role NCEH1 plays in human biology, its natural substrate, 2-acetyl MAGE, is also a known component of platelets. In 2012, Holly *et al.* suggested that the accumulation of 2-acetyl MAGE could inhibit the signaling cascade for integrin  $\alpha$ IIb $\beta$ 3 activation implying a regulatory role of NCEH1 in platelet aggregation.<sup>8</sup> In this instance, 2-acetyl MAGE competes with a separate upregulated

component that binds to Protein Kinase C (PKC), the responsible enzyme for activating the integrin  $\alpha$ IIb $\beta$ 3. The binding of 2-acetyl MAGE to PKC prevents the activation of the integrin, thus preventing the aggregation of platelets. The up-regulation of NCEH1 would lead to a decrease in the concentration of 2-acetyl MAGE and result in an increase in platelet aggregation. This increase, may lead to clot formation responsible for heart attacks or stroke. The regulatory role of NCEH1 in platelets, its anti-atherosclerotic behavior in macrophages, and its promotion of tumor invasiveness all strongly justify the need to further characterize its structure and function.

## **1.2 Inhibition of NCEH1 by Organophosphorus Compounds**

Organophosphorus (OPs) compounds are a major class of insecticides as well as potent chemical warfare agents.<sup>13</sup> In addition to being shown to hydrolyze acetylated MAGE and cholesterol esters, NCEH1 has also displayed the ability to detoxify common organophosphorus compounds such as chlorpyrifos oxon (CPO), parathion, and paraoxon.<sup>5, 6, 13, 14</sup> Nomura *et al.* demonstrated the defense of nerve tissue against the dangerous effects of OPs by comparing cholinergic symptoms such as tremoring and mortality of wild type mice verse NCEH1 knockout mice. Their results showed a drastic increase in both tremoring and mortality in the knockout mice when treated with insecticides CPF and parathion.<sup>5</sup> NCEH1 not only demonstrated a protective role, but also displays an unusual ability to spontaneously reactivate from phosphorylation preventing its deactivation, creating a therapeutic interest for a soluble form of NCEH1 be obtained.<sup>5</sup>

Structural determination of NCEH1 could provide the possibility of directed mutagenesis to improve the catalytic efficiency of NCEH1 against dangerous nerve

agents. The ability to obtain soluble protein could also provide the ability to create a bioscavenger capable of hydrolyzing OPs in the blood plasma before they came in contact with acetylcholinesterase.

### **1.3 Known Structural Components of NCEH1**

NCEH1 is a single membrane spanning type II membrane protein. It has three domains: an N-terminal domain predicted to be required for localization to the ER membrane, a catalytic domain, and a C-terminal lipid binding domain.<sup>15</sup> There are three predicted human isoforms of NCEH1, varying in length between 448, 440, and 275 amino acids. Much of the literature has focused on the mouse form of the protein with very few studies focused on the human isoforms. The sequence alignment between the three human isoforms and the mouse isoform of NCEH1 is shown in Figure 1.2.

Mouse	1	-----MRSSCVLI <del>AA</del> LAALALAAAYVYIPLPSAVS
IsoformB	1	MSSCRGQKVAGGLRVVSPFPLCQPAGEPSQGMRSSCVLLTALVALAAAYVYIPLPGSVS
IsoformA	1	MSSCRGQKVAGGLRVVSPFPLCQPAGEPSQGMRSSCVLLTALVALAAAYVYIPLPGSVS
IsoformC	1	-----
Mouse	29	DPWKLMLLDATFRGAQQVSNLIHSLGLNHHLTALNFIITTSFGKQ <del>SARSS</del> PKVKVTD <del>DFD</del>
IsoformB	61	DPWKLMLLDATFRGAQQVSNLIHYLGLSHHLLALNFIIVSFGKKS <del>AWSSA</del> QVKVTD <del>DFD</del>
IsoformA	61	DPWKLMLLDATFRGAQQVSNLIHYLGLSHHLLALNFIIVSFGKKS <del>AWSSA</del> QVKVTD <del>DFD</del>
IsoformC	1	-----
Mouse	89	GVEVRVFE <del>GS</del> PKPEEPLR <del>RSV</del> YI <del>THGGG</del> WALASA-----KISYYD <del>QLCT</del> TMAEELNA
IsoformB	121	GVEVRVFE <del>GP</del> PKPEEPLK <del>RSV</del> YI <del>THGGG</del> WALASA-----KIRYYDELCTAMAEELNA
IsoformA	121	GVEVRVFE <del>GP</del> PKPEEPLK <del>RSV</del> YI <del>THGGG</del> WALASASASWSPSDEIRYYDELCTAMAEELNA
IsoformC	1	-----MAEELNA
Mouse	141	VIVSIEYRLV <del>PQ</del> VYFPEQIHDVVRATKYFL <del>Q</del> PEVLDK <del>YK</del> VDPGRV <del>GIS</del> SGDS <del>S</del> AGGNLAAAL
IsoformB	173	VIVSIEYRLV <del>PKV</del> YFPEQIHDVVRATKYFLKPEVLQKYMVDPGRICISGDS <del>S</del> AGGNLAAAL
IsoformA	181	VIVSIEYRLV <del>PKV</del> YFPEQIHDVVRATKYFLKPEVLQKYMVDPGRICISGDS <del>S</del> AGGNLAAAL
IsoformC	8	VIVSIEYRLV <del>PKV</del> YFPEQIHDVVRATKYFLKPEVLQKYMVDPGRICISGDS <del>S</del> AGGNLAAAL
Mouse	201	GQOFTYV <del>ASL</del> KNK <del>LK</del> LQALVYPVLOALDFNTPSYQCSM <del>NT</del> PILPRHVMVRYWLDYFKGNY
IsoformB	233	GQOFTQDASLKNK <del>LK</del> LQALYIPVLOALDFNTPSYQQNVNTPILPRYVMVKYVVDYFKGNY
IsoformA	241	GQOFTQDASLKNK <del>LK</del> LQALYIPVLOALDFNTPSYQQNVNTPILPRYVMVKYVVDYFKGNY
IsoformC	68	GQOFTQDASLKNK <del>LK</del> LQALYIPVLOALDFNTPSYQQNVNTPILPRYVMVKYVVDYFKGNY
Mouse	261	DFV <del>E</del> AMIVN <del>NHT</del> SLDVER <del>AAAL</del> RARLD <del>WTS</del> L <del>LPS</del> SIK <del>K</del> NYKPI <del>M</del> QTTGNARIVQEL <del>P</del> QLL
IsoformB	293	DFVQAMIVN <del>NHT</del> SLDVEEAAAVRARLN <del>NWTS</del> L <del>L</del> PASFTK <del>N</del> YKPVVQTTGNARIVQEL <del>P</del> QLL
IsoformA	301	DFVQAMIVN <del>NHT</del> SLDVEEAAAVRARLN <del>NWTS</del> L <del>L</del> PASFTK <del>N</del> YKPVVQTTGNARIVQEL <del>P</del> QLL
IsoformC	128	DFVQAMIVN <del>NHT</del> SLDVEEAAAVRARLN <del>NWTS</del> L <del>L</del> PASFTK <del>N</del> YKPVVQTTGNARIVQEL <del>P</del> QLL
Mouse	321	DA <del>A</del> ASPLIAD <del>Q</del> AVL <del>Q</del> LLPKTYILTCEH <del>D</del> VLRDDGIMYAKRLESAGV <del>N</del> VTLDHFEDGF <del>H</del> GC
IsoformB	353	DARSAPLIADQAVLQ <del>L</del> LPKTYILTCEH <del>D</del> VLRDDGIMYAKRLESAGVEVTLDFEDGF <del>H</del> GC
IsoformA	361	DARSAPLIADQAVLQ <del>L</del> LPKTYILTCEH <del>D</del> VLRDDGIMYAKRLESAGVEVTLDFEDGF <del>H</del> GC
IsoformC	188	DARSAPLIADQAVLQ <del>L</del> LPKTYILTCEH <del>D</del> VLRDDGIMYAKRLESAGVEVTLDFEDGF <del>H</del> GC
Mouse	381	MIFTSWPTNFSVGIRTRNSYIKWLDQNL
IsoformB	413	MIFTSWPTNFSVGIRTRNSYIKWLDQNL
IsoformA	421	MIFTSWPTNFSVGIRTRNSYIKWLDQNL
IsoformC	248	MIFTSWPTNFSVGIRTRNSYIKWLDQNL

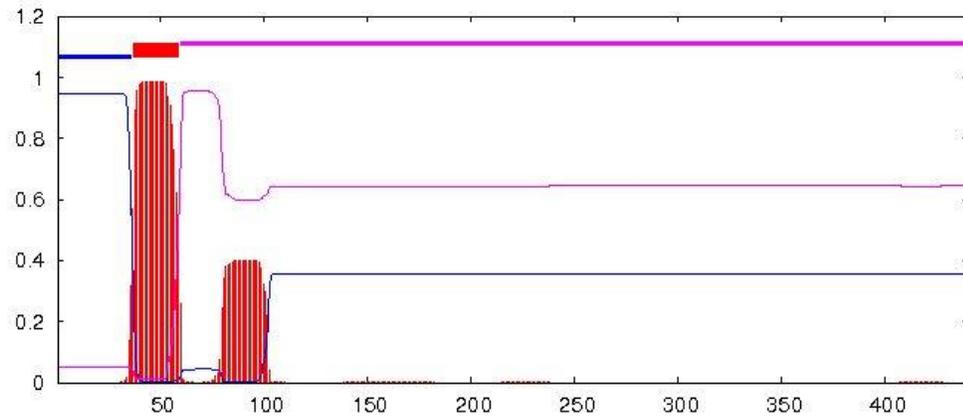
**Figure 1.2:** The sequence alignment of NCEH1. Alignment shows the three human isoforms and the mouse isoform of NCEH1. The characteristic oxyanion hole is highlighted in yellow, while the active site residues at S191, D348, and H378 are highlighted in red. The three predicted human glycosylation sites are dictated by the red arrows at N270, N287, and N389. The human and mouse form share Asn-glycosylation sites at N270 and N389, however the mouse form lacks the asparagine at position 287 and has been shown to be glycosylated at position 367 demonstrated by the red astricks.

The alignment shows human isoforms A and B only differ in the addition of a short amino acid sequence approximately one hundred and sixty amino acids into the protein. The purpose of this sequence has not been previously studied. Isoform C is shortened relative to the other two isoforms due to a start codon downstream of the previous two. This isoform also has a stronger Kozak sequence. Functional differences of this shortened isoform have not been studied experimentally. The mouse form of the protein is approximately 32 amino acids shorter than human isoform B, giving rise to a protein 408 amino acids in length. Comparing the sequences of human isoform B and the mouse protein yields a protein sequence identity of 88%, however this identity value does not include the mouse form's lack of the first 32 amino acids present in the human form of the protein.

Both human and mouse have three predicted glycosylation sites. The mouse glycosylation sites, verified by Igarashi *et al.*, are at amino acid 270, 367, and 389.<sup>15</sup> The human form shares glycosylation sites at the amino acid positions 270 and 389, however it is important to note that the human form of NCEH1 does not contain an asparagine at location 367, but has a third predicted glycosylation site at position 287. Previous research has demonstrated an importance of glycosylation as it is predicted to be required for catalytic activity of NCEH1.<sup>15</sup>

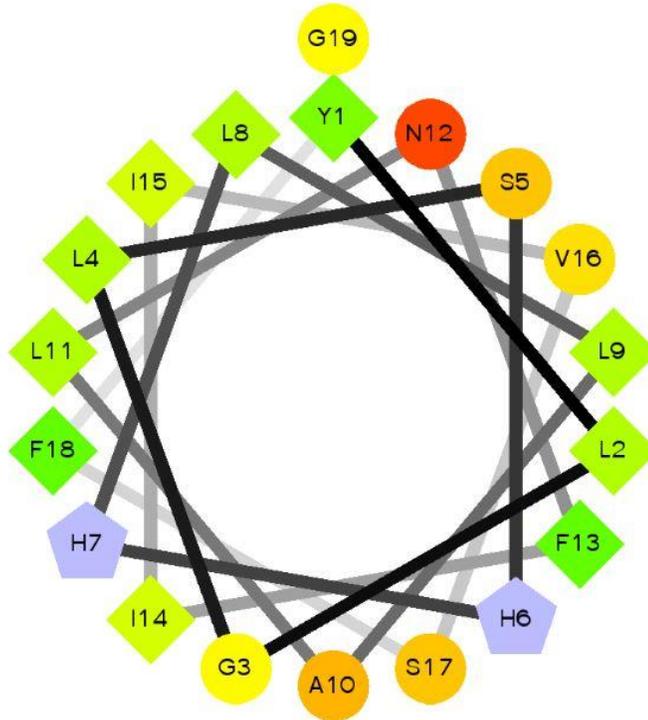
Figure 1.3 demonstrates the predicted transmembrane region comprising the first 50 amino acids of NCEH1 isoform b. The transmembrane helix is predicted to

functionally localize the enzyme to the ER membrane and the catalytic core of the protein faces the lumen.<sup>15</sup>



**Figure 1.3:** Transmembrane prediction of NCEH1. Prediction generated from the transmembrane prediction software <http://www.cbs.dtu.dk/services/TMHMM/> created by Krogh *et al.* The transmembrane spans the ER membrane and the catalytic domain faces the lumen.<sup>16</sup>

A second hydrophobic region can be seen immediately after the transmembrane helix. The sequence corresponding to this second region was inserted into a helical wheel projection model and is shown in Figure 1.4. The helical wheel demonstrates a hydrophobic region, shown as diamonds within the model, with flanking hydrophilic residues, shown as circles, indicating this to be an amphipathic sequence, which would likely be peripherally associated to the membrane of the ER.



**Figure 1.4:** Helical wheel projection of the second hydrophobic sequence in NCEH1. The projection was completed using the program from <http://rzlab.ucr.edu> created by Don Armstrong and Raphael. Zidovetzki. Version: Id: wheel.pl,v 1.4 2009-10-20 21:23:36 don Exp. Hydrophobic residues are displayed as diamonds with an increasing gradient of green to demonstrate varying levels of hydrophobicity. The neutral and hydrophilic residues are displayed as circles with an increasing gradient of red to show an increase in hydrophilic nature. The two blue residues represent histidine, which would be mostly neutral in the native ER environment with a pH of 7.3. The sequence "YLGLSHLLALNFIIVSFG" corresponding to amino acids 84 to 102, shows a hydrophobic side flanked by hydrophilic residues. This alpha helical structure indicates an association to the ER membrane.

#### 1.4 Summary and Thesis Direction

NCEH1 has been linked to the top three killers of the United States: cancer, heart attack, and stroke. It has been shown to have an invasive tumorigenic effect in

cancer<sup>2-4</sup>, aid in the prevention of atherosclerosis<sup>9-11</sup>, regulate platelet aggregation<sup>8</sup>, and detoxify OP compounds<sup>5, 17</sup>. Multiple sources have indicated that NCEH1 is a possible therapeutic drug target,<sup>2, 5, 8</sup> yet its structure has not been elucidated. In addition, the association of NCEH1 to a membrane has prevented purification for functional and structural characterization. As a result, the majority of the work completed on NCEH1 has been in mammalian expression or mouse models, neither of which provided an adequate yield or the ability to purify NCEH1 for structural studies. Previous functional studies of NCEH1 have been completed with the use of crude lysates, macrophages, platelets, or from cellular tissue.<sup>3, 13, 15, 18-20</sup> However, true determinations of intrinsic kinetic values such as  $K_M$  and  $k_{cat}$  are ideally carried out using purified protein. The following describes the work that has led to the successful expression, purification and preliminary crystallization of NCEH1. In chapter 2, the expression and purification of NCEH1 from the recombinant *Escherichia coli* system will be discussed. Limitations of the system due to prokaryotic folding pathways lead to the inability to express the enzyme without a large fusion tag. This tag was unable to be cleaved effectively and functional characterization was carried out with the fusion of maltose binding protein still attached. Functional characterization discussed on the basis of activity toward para-nitrophenyl acetate (PNPA) and a modified natural substrate 2-thio-acetyl MAGE resulted in the direct measurement of  $K_M$  and  $k_{cat}$  and contrasted the previously reported data<sup>15</sup> suggesting glycosylation was needed for catalytic ability. Chapter 3 discusses the expression and purification of NCEH1 from *Pichia pastoris* and demonstrates the advantages of this eukaryotic expression system. Here we report the first successful isolation of a stable and purified NCEH1, without a fusion tag, from both an intracellular and secretion based expression protocol. In

addition, the purification of NCEH1 from *Pichia pastoris* lead to the recovery of concentrations adequate enough for extensive protein crystal screening to begin. The conclusions of this work will be stated in chapter 4 along with the future goals and directions of the NCEH1 project.

## REFERENCES

- [1] Cravatt, B. F., Wright, A. T., and Kozarich, J. W. (2008) Activity-based protein profiling: From enzyme chemistry, *Annual Review of Biochemistry* 77, 383-414.
- [2] Jessani, N., Liu, Y. S., Humphrey, M., and Cravatt, B. F. (2002) Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness, *Proceedings of the National Academy of Sciences of the United States of America* 99, 10335-10340.
- [3] Chiang, K. P., Niessen, S., Saghatelian, A., and Cravatt, B. F. (2006) An enzyme that regulates ether lipid signaling pathways in cancer annotated by multidimensional profiling, *Chemistry & Biology* 13, 1041-1050.
- [4] Chang, J. W., Nomura, D. K., and Cravatt, B. F. (2011) A Potent and Selective Inhibitor of KIAA1363/AADACL1 that Impairs Prostate Cancer Pathogenesis, *Chemistry & Biology* 18, 476-484.
- [5] Nomura, D. K., Fujioka, K., Issa, R. S., Ward, A. M., Cravatt, B. F., and Casida, J. E. (2008) Dual roles of brain serine hydrolase KIAAL1363 in ether lipid metabolism and organophosphate detoxification, *Toxicology and Applied Pharmacology* 228, 42-48.
- [6] Casida, J. E., Nomura, D. K., Vose, S. C., and Fujioka, K. (2008) Organophosphate-sensitive lipases modulate brain lysophospholipids, ether lipids and endocannabinoids, *Chemico-Biological Interactions* 175, 355-364.
- [7] Buchebner, M., Pfeifer, T., Rathke, N., Chandak, P. G., Lass, A., Schreiber, R., Kratzer, A., Zimmermann, R., Sattler, W., Koefeler, H., Froehlich, E., Kostner, G. M., Birner-Gruenberger, R., Chiang, K. P., Haemmerle, G., Zechner, R., Levak-Frank, S., Cravatt, B., and Kratky, D. (2010) Cholesteryl ester hydrolase activity is abolished in HSL<sup>-/-</sup> macrophages but unchanged in macrophages lacking KIAA1363, *Journal of Lipid Research* 51, 2896-2908.
- [8] Holly, S. P., Chang, J. W., Niessen, S., Boulaftali, Y., Smith, M. C., Phillips, R. M., Black, J. L., Piatt, R., Li, W., Weyrich, A. S., Bergmeier, W., Cravatt, B. F., and Parise, L. V. (2012) Identification of AADACL1 As a Novel Regulator of Human Platelets Via Chemoproteomics, *Blood* 120.
- [9] Igarashi, M., Osuga, J.-i., Uozaki, H., Sekiya, M., Nagashima, S., Takahashi, M., Takase, S., Takanashi, M., Li, Y., Ohta, K., Kumagai, M., Nishi, M., Hosokawa, M., Fledelius, C., Jacobsen, P., Yagyu, H., Fukayama, M., Nagai, R., Kadowaki, T., Ohashi, K., and Ishibashi, S. (2010) The Critical Role of

Neutral Cholesterol Ester Hydrolase 1 in Cholesterol Removal From Human Macrophages, *Circulation Research* 107, 1387-1395.

- [10] Okazaki, H., Igarashi, M., Nishi, M., Sekiya, M., Tajima, M., Takase, S., Takanashi, M., Ohta, K., Tamura, Y., Okazaki, S., Yahagi, N., Ohashi, K., Amemiya-Kudo, M., Nakagawa, Y., Nagai, R., Kadowaki, T., Osuga, J.-i., and Ishibashi, S. (2008) Identification of Neutral Cholesterol Ester Hydrolase, a Key Enzyme Removing Cholesterol from Macrophages, *Journal of Biological Chemistry* 283, 33357-33364.
- [11] Sekiya, M., Osuga, J.-i., Nagashima, S., Ohshiro, T., Igarashi, M., Okazaki, H., Takahashi, M., Tazoe, F., Wada, T., Ohta, K., Takanashi, M., Kumagai, M., Nishi, M., Takase, S., Yahagi, N., Yagyu, H., Ohashi, K., Nagai, R., Kadowaki, T., Furukawa, Y., and Ishibashi, S. (2009) Ablation of Neutral Cholesterol Ester Hydrolase 1 Accelerates Atherosclerosis, *Cell Metabolism* 10, 219-228.
- [12] Contreras, J. A., and Lasuncion, M. A. (1994) Essential differences in cholesteryl ester metabolism between human monocyte-derived and J774 macrophages-evidence against the presence of hormone sensitive lipase in human macrophages, *Arteriosclerosis and Thrombosis* 14, 443-452.
- [13] Nomura, D. K., Leung, D., Chiang, K. P., Quistad, G. B., Cravatt, B. F., and Casida, J. E. (2005) A brain detoxifying enzyme for organophosphorus nerve poisons, *Proceedings of the National Academy of Sciences of the United States of America* 102, 6195-6200.
- [14] Ross, M. K., Pluta, K., Bittles, V., Borazjani, A., and Crow, J. A. (2016) Interaction of the serine hydrolase KIAA1363 with organophosphorus agents: Evaluation of potency and kinetics, *Archives of Biochemistry and Biophysics* 590, 72-81.
- [15] Igarashi, M., Osuga, J.-i., Isshiki, M., Sekiya, M., Okazaki, H., Takase, S., Takanashi, M., Ohta, K., Kumagai, M., Nishi, M., Fujita, T., Nagai, R., Kadowaki, T., and Ishibashi, S. (2010) Targeting of neutral cholesterol ester hydrolase to the endoplasmic reticulum via its N-terminal sequence, *Journal of Lipid Research* 51, 274-285.
- [16] Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes, *Journal of Molecular Biology* 305, 567-580.
- [17] Nomura, D. K., Durkin, K. A., Chiang, K. P., Quistad, G. B., Cravatt, B. F., and Casida, J. E. (2006) Serine hydrolase KIAA1363: Toxicological and structural features with emphasis on organophosphate interactions, *Chemical Research in Toxicology* 19, 1142-1150.

- [18] Sekiya, M., Osuga, J.-i., Igarashi, M., Okazaki, H., and Ishibashi, S. (2011) The Role of Neutral Cholesterol Ester Hydrolysis in Macrophage Foam Cells, *Journal of Atherosclerosis and Thrombosis* 18, 359-364.
- [19] Chang, J. W., Moellering, R. E., and Cravatt, B. F. (2012) An Activity-Based Imaging Probe for the Integral Membrane Hydrolase KIAA1363, *Angewandte Chemie-International Edition* 51, 966-970.
- [20] Holly, S. P., Chang, J. W., Li, W. W., Niessen, S., Phillips, R. M., Piatt, R., Black, J. L., Smith, M. C., Boulaftali, Y., Weyrich, A. S., Bergmeier, W., Cravatt, B. F., and Parise, L. V. (2013) Chemoproteomic Discovery of AADACL1 as a Regulator of Human Platelet Activation, *Chemistry & Biology* 20, 1125-1134.

## Chapter 2

### PURIFICATION OF NEUTRAL CHOLESTEROL ESTER HYDROLASE 1 FROM *ESCHERICHIA COLI*

#### 2.1 Introduction

Neutral cholesterol ester hydrolase 1 (NCEH1) was first identified as a protein of interest in cancer cell invasiveness by Jessani *et al.* in 2002. With the use of activity based protein profiling (ABPP) NCEH1, known as KIAA1363, was identified as being up-regulated in highly aggressive cancerous cell lines.<sup>1</sup> NCEH1 was shown to be associated with a membrane component of the cell and posttranslationally modified via glycosylation. Research has suggested that this glycosylation is required for the activity of the protein.<sup>1,2</sup> While the drawbacks of being up-regulated in cancerous cell lines are obvious, other research has identified NCEH1 as serving a critical role in the prevention of atherosclerosis. This work has suggested that NCEH1 plays a critical protective role, by acting as a cholesterol esterase. NCEH1 helps to prevent the formation of the macrophage foam cells that lead to atherosclerosis by cleaving the cholesterol esters.<sup>3</sup> While this work remains controversial,<sup>4</sup> our current understanding of NCEH1 indicates that it is both pro-cancerous and anti-atherosclerotic. Thus, an analysis of the structure of NCEH1 and an in-depth characterization of its functional components remain critical to our understanding of this crucial enzyme.

NCEH1 has been predicted to have a single transmembrane (TM) helix, thereby making it a transmembrane protein<sup>2</sup> and significantly increasing the complexity of expression and purification. As a consequence of this complexity, researchers have generally utilized localization and functional studies from

mammalian cell lysates, knock-out mice, or adenovirus expressions to study NCEH1<sup>1-3, 5-7</sup>. The primary advantage of this form of experimentation is the ability to test the enzyme's activity in the presence of possible cofactors and small molecule modulators; however, the lack of purified protein precludes the ability to calculate critical kinetic parameters such as  $k_{cat}$  and  $K_M$ .<sup>5</sup> In order to measure the kinetic parameters that describe the functional activity of NCEH1, this work pursued recombinant over expression methods. Utilizing a recombinant expression system has many advantages over mammalian expression protocols. These advantages include ease of use, cost efficiency, and most importantly the yield and purity of the production of a protein of interest. With a recombinant approach, large amounts of soluble proteins can be purified easily and efficiently.

The work presented in this chapter will explore several expression and purification strategies and evaluate the ability of NCEH1 to be expressed and purified from *E. coli*. Furthermore, insights will be provided into the influence of post-translational modification on the activity of NCEH1. *E. coli* lack the machinery to perform many mammalian post-translational modifications and are therefore particularly well suited for this purpose.

## **2.2 Materials and Methods**

### **2.2.1 PCR Primers Utilized for *E. coli* Expression**

The coding DNA for the NCEH1 gene was obtained from Dr. Benjamin F. Cravatt (The Scripps Research Institute, La Jolla, CA) with the first 55 amino acids truncated, which removed the N-terminal the TM helix, from this construct. The DNA

was amplified for the insertion of the NCEH1 construct into a variety of different expression vectors, with the goal of optimizing the expression and purification conditions. The primers utilized for the insertion of NCEH1 into the *E. coli* expression vectors are listed in Table 2.1. The primers are identified by name, primer sequence in the 5'-3' direction, and their corresponding restriction enzyme. The primers corresponding to the insertion of NCEH1 into the pET100/D-topo expression vector lack restriction enzymes. As will be discussed later, the topo expression system does not require the use of restriction enzymes to subclone the gene into the vector.

Primer Name	Primer Sequence 5' – 3'	Restriction Site
KIAA-forward	AAAAAAGAATTCCTGTCGACCCCTGGAAGCTG	ECORI
KIAA-reverse	AGACAGCGTCGACTTATTACAGGTTTTGATCTAGCCAC	SALI
Topo-forward	CACCTTCCGTGTCGACCCCTGGAAGCTG	
Topo-reverse	TTACAGGTTTTGATCTAGCCGTGTCGACCCCTGGAAGCTG	
His-forward	AAAAAAAACATATGTTCCGTGTCGACCCCTGGAAGCTG	NDEI
His-reverse	AAAAAAGTCGACCAGGTTTTGATCTAGCCACTTG	SALI
Trunc35	AAAAAAGAATTCACTGGCACTGAATTTATCATTG	ECORI
Trunc55	AAAAAAGAATTC AAGGGAGACACAGACTTTGATGG	ECORI
PmaI56Forward	AAAAAACATATGTCGTCGACCCCTGGAAGCTG	NDEI
PmaI15Forward	AAAAAACATATGACCGACACAGACTTTGATGG	NDEI
For275	AAAAACATATGGCTGAGGAATTGAATGC	NDEI

**Table 2.1:** The PCR primers utilized in the amplification of NCEH1 for *E. coli*. The primers are listed by name, sequence in the 5'-3' direction, and the associated restriction enzymes required for insertion into the vector of interest. The Topo forward and reverse primers did not require restriction enzymes in their ligation protocol.

### 2.2.2 Cloning into pET100/D-TOPO, pET20b+, and Expression of NCEH1

The NCEH1 gene was amplified for insertion into the pET100/D-topo expression vector using the Topo-forward and Topo-Reverse primers listed in Table

2.1. The underlined sequence in the forward primer corresponds to a short overhang allowing the gene to be inserted into the expression vector. The reverse primer lacks an overhanging sequence and is ligated into the vector with a blunt end ligation technique. The sequence was ligated into the expression vector with PCR blunt Topo, provided by Invitrogen, and transformed into chemically competent One Shot Top10 *E. coli* cells with the use of the Champion pET100 Directional TOPO cloning reaction mix. The mixture was then heat shocked before rocking at 37 °C with 1 mL Lysogeny Broth (LB). Transformation products were plated on LB agar containing 0.2 mg/mL ampicillin. Single colonies were chosen to be grown-up individually overnight and the resulting vector contained an N-terminal histidine tag confirmed by Genewiz, inc sequencing.

Multiple constructs of truncated 56-440 NCEH1 were amplified for insertion into the pET20b+ expression vector by varying the primers used to amplify the sequence. KIAA-forward, His-forward, KIAA-reverse, and His-reverse primers, listed in Table 2.1 were used to generate a variety of unique constructs that demonstrated different features including histidine tags and a periplasmic localization sequence. The amplified DNA was PCR purified using a Qiagen preparation kit, and separated by gel electrophoresis. The band of interest was then excised and purified a second time with a Qiagen preparation kit. Purified DNA, as well as purified empty vector, were double digested with the appropriate restriction enzymes for approximately 3 h at 37 °C before being separated and purified by gel electrophoresis. The resulting purified, digested DNA and plasmid were ligated overnight at 14 °C with T4 DNA ligase (New

England Biolabs). The ligated sample was dialyzed against pure distilled and deionized H<sub>2</sub>O (ddH<sub>2</sub>O) for 20 min before a small portion was mixed with 20  $\mu$ L of DH5 $\alpha$  electrocompetent cells and stored on ice for 5 min. The cells were placed in an electroporation chamber and pulsed at 25 kV. The electroporated sample was mixed with antibiotic free LB, and grown at 37 °C for 1 h. Aliquots of the mixture were plated on LB agar containing 0.2 mg/mL ampicillin and incubated overnight at 37 °C. Single colonies were selected and grown overnight in LB, allowing for the plasmid purification of each colony. Sequencing via Genewiz Inc. confirmed the presence of the appropriate insert and the plasmids containing the NCEH1 sequence were transformed into BL21-CodonPlus(DE3)-RIPL cells (Agilent). Small aliquots from the single colony sample grown overnight were retained for further experimentation and stored at -80 °C until required.

The stored aliquots of cells containing the NCEH1 expression vector were used to inoculate overnight cultures for a small scale expression of the various NCEH1 constructs. Overnights were used to inoculate 50 mL of LB broth containing 0.2 mg/mL ampicillin and the cells were grown for approximately 3 h until an OD<sub>600</sub> of 1 was achieved. Expression was induced with 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside, IPTG, and the cells were grown at 30 °C for another 6 h before harvesting the pellet. The pellets were resuspended in 50 mM sodium phosphate buffer containing 150 mM NaCl, pH 7.4. Each resuspension was sonicated 3 times for 30 s with 5 min of rest on ice between each cycle. After the suspensions darkened in color, they were incubated with 1% tween 20, and rocked for 20 min at 4 °C. The

suspensions were then centrifuged for twenty min at 23000 x g to remove the membrane fraction and the resulting supernatant was decanted and further analyzed by SDS page.

### **2.2.3 Cloning into pGEX-4T3, Expression and Purification of NCEH1**

The NCEH1 gene was amplified for ligation into pGEX-4T3 using the forward primers KIAA-forward, Trunc35, and Trunc55, listed in Table 2.1 and all corresponding to separate truncations of the NCEH1 gene. All of the DNA that was amplified utilized the same reverse primer, listed as KIAA-reverse in Table 2.1. Restriction digest was performed with NDEI and SALI restriction enzymes (New England Biolabs) and the DNA was ligated overnight at 14 °C with T4 DNA ligase. The ligated vector was electroporated with electrocompetent DH5 $\alpha$  cells as previously described, and plated overnight for single colonies. After verification of the sequence, individual colonies were transformed into BL21-CodonPlus(DE3)-RIPL cells.

Small scale expressions were completed as previously described; however pellets were resuspended in 50 mM Tris buffer, 150 mM NaCl, 1 mM EDTA, 1  $\mu$ M pepstatin, pH 7.8. The pellets were lysed by sonication, rocked at 4 °C with 1% tween 20 for 20 min, and centrifuged at 23000 x g for 20 min. The supernatant was utilized for verification of overexpression with SDS PAGE analysis.

Once expression was verified, colonies were selected for large scale expression. Six 7 L flasks containing 1 L of LB with 0.2 mg/mL ampicillin each were inoculated at a ratio of 1/100 from an overnight stock solution. The cells were grown for approximately 4 h at 37 °C until an OD<sub>600</sub> of 1 was reached. The cells were induced with 1 mM IPTG, expressed overnight at room temperature and harvested by

centrifugation at 8,200 x g. The pellets were stored at -80 °C for future experimentation.

Pellets were resuspended in 50 mM Tris buffer, 150 mM NaCl, 1 mM EDTA, 1 μM pepstatin, pH 7.8. They were sonicated for 1 min intervals with 50% power three times. There were 5 min of rest on ice between each sonication to maintain the temperature of solution. After sonication, the lysed cells were supplemented with 1% tween 20 and rocked at 4 °C for twenty min before being centrifuged at 23000 x g for an additional twenty min. To eliminate chaperone proteins that have been identified as purifying with GST the supernatant was shaken at 37 °C for twenty min with 5 mM ATP and 20 mM MgCl<sub>2</sub> before being centrifuged at 23000 x g for a second twenty min spin. The supernatant was incubated, rocking at 4 °C in a gravity flow chromatography column supplemented with glutathione sepharose beads (GE healthcare) for 1 h. The supernatant was drained and the column was washed with 50 mM Tris buffer, 150 mM NaCl, 1 mM EDTA, pH 7.8. Protein was eluted from the glutathione beads with sequential 10 mL elutions of 50 mM Tris buffer, 150 mM NaCl, 1 mM EDTA, 25 mM reduced glutathione, pH 7.8. Each elution rocked for 10 min at room temperature before being drained from the column. The protein elutions were combined and concentrated to 3 mL with an Amicon centrifugation column possessing a molecular cutoff of 30 kDa. The reduced glutathione was then removed by overnight dialysis into a 50 mM Tris buffer, containing 150 mM NaCl, 1 mM EDTA, pH 7.8. After overnight dialysis, the sample was incubated rocking at 4 °C for 4 h with thrombin protease to cleave the GST tag. The resulting sample was incubated, and rocked at 4°C on a glutathione sepharose gravity column for 1 h to remove the cleaved GST from the sample. The cleaved NCEH1 eluate was then loaded onto a

gravity flow column containing benzamidine sepharose (GE healthcare) and rocked for 1 h at 4 °C to remove the thrombin protease from the purified sample. The unretained protein was then concentrated to 500 µL before dialyzing overnight into 10 mM HEPES, 10 mM NaCl, pH 7.4. The dialyzed sample was either used or frozen immediately. Retained samples were kept at -80 °C until needed for further experimentation.

#### **2.2.4 Cloning into pMAL-C5X, Expression and Purification of NCEH1**

The DNA for the insertion of NCEH1 into the pMAL-C5X vector was amplified using the forward primers pmal56forward, pmal115forward, and For275 whose sequences are displayed in Table 2.1. These primers represent the 56-440 truncation removing the transmembrane helix, the 115-440 truncation removes the transmembrane helix as well as a second hydrophobic region, and the final 165-440 truncation mimics the third human isoform of NCEH1. The reverse primer, KIAA-reverse, remained the same. The amplified DNA was digested with SALI and NDEI restriction enzymes (New England Biolabs) and followed the same ligation and transformation protocol previously described for the cloning of NCEH1 into pGEX-4T3. Once the sequence was verified, the samples were transformed into BL21-CodonPlus(DE3)-RIPL cells, and expressed following the same protocol as the pGEX vector.

Pelleted cells containing the NCEH1 gene were resuspended in 50 mM sodium phosphate buffer, 300 mM NaCl, pH 7.4. The cells were then lysed by sonicating three times for 1 min intervals at 50% power with a 5 min rest on ice between each sonication. Tween 20 was added to a final concentration of 1% and the suspension was rocked at 4 °C for twenty min. The membrane portion of the sample was then removed by two rounds of centrifugation for twenty min at 23000 x g. After the first spin, the supernatant was added to a second tube and spun again. The supernatant was then incubated on a gravity flow column containing Ni-sepharose beads (GE healthcare) for 1 h. The unretained portion of the supernatant was drained, and the column was washed with 50 mM sodium phosphate buffer, 300 mM NaCl, pH 7.4. A second wash with 50 mM sodium phosphate buffer, 300 mM NaCl, 100 mM imidazole, pH 7.4 was completed before the protein was eluted in 10 mL elutions of 50 mM sodium phosphate buffer, 300 mM NaCl, 300 mM imidazole, pH 7.2. The elutions were combined and concentrated to 3 mL before being dialyzed overnight into 10 mM HEPES, 10 mM NaCl, pH 7.4.

### **2.2.5 Activity Assays of Neutral Cholesterol Ester Hydrolase 1**

NCEH1 was functionally characterized with the use of Para-Nitrophenylacetate (Sigma), abbreviated PNPA. Various amounts of PNPA were added to a 1 mL cuvette containing 50 mM Tris-base buffer, pH 7.8. NCEH1 was added and mixed thoroughly to start the reaction. Each sample was monitored for 30 s at 405 nm.

NCEH1 activity was also detected with a modified natural substrate 2-thio-acetyl MAGE (Caymen Chemical). Various amounts of 2-thio-acetyl MAGE were included in a 1 mL total volume assay containing 50 mM sodium phosphate buffer, pH

7.4, 10 mM 5,5'-dithiobis(2-nitrobenzoic acid) or DTNB (Sigma), and the reaction was started by thoroughly mixing NCEH1 into solution. Each sample was monitored for 300 s at 412 nm.

### **2.2.6 Protein Crystallization Screening of NCEH1**

Protein samples utilized for structural characterization were purified from the pMAL expression system and remained as a MBP fusion protein. Samples demonstrating a concentration higher than 0.5 mg/mL and a purity greater than 90% were utilized for initial hanging drop screening after being forced through a syringe filter. Crystal screens 1 and 2 were obtained from Hampton Research and stored at room temperature after preparation. Hanging drops were prepared by mixing 1  $\mu$ L of sample with 1  $\mu$ L of well solution.

### **2.2.7 Random Mutagenesis Protocol**

Random mutagenesis was carried out on the DNA of truncation 56-440 NCEH1 using the Pmal56forward primer and KIAA-reverse primer listed in Table 2.1 for insertion into the pMAL\_c5x vector. The initial PCR to create a library of mutagenic DNA utilized a low fidelity taq polymerase and 10 mM mutagenic dNTPs 8-oxo-2'-deoxyguanosine-5'-triphosphate and 2'-deoxy-P-nucleoside-5'-triphosphate (Trilink Biotechnologies). The PCR reaction was carried out in duplicate 50  $\mu$ L samples for 20 cycles with 1  $\mu$ L of fresh template added to each tube after the tenth cycle. The two prepared reactions were combined and PCR purified using Qiagen spin columns to create one library. The purified sample was used as a template for a second PCR procedure, following standard PCR protocols for the amplification of the DNA library. Amplified mutagenic DNA was inserted into the pMAL\_c5X expression

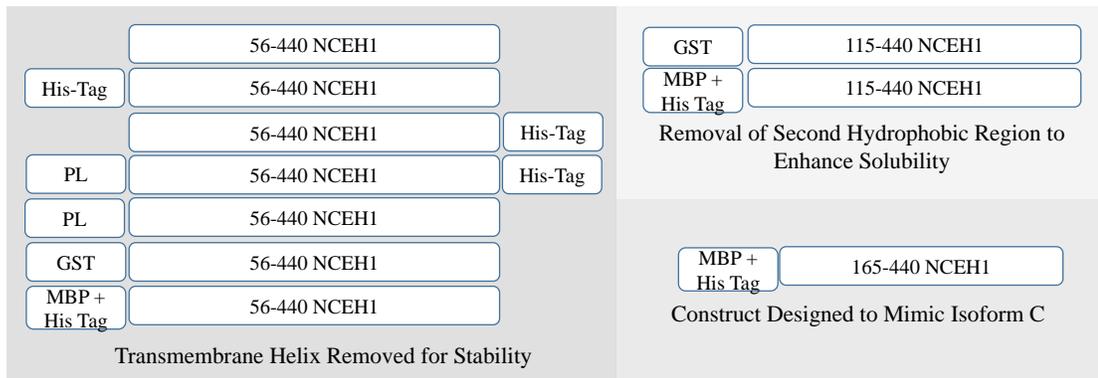
vector and transformed directly into BL21-CodonPlus(DE3)-RIPL cells for expression screening. Small scale expressions were carried out shaking in 5 mL volumes at 37 °C. These expressions were induced with 0.3 mM IPTG and expressed for 2 h at 37 °C. Each of the small scale expressions were then spun down for 10 min at 3,200 x g before being resuspended in 2 mL of 50 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 7.8. Each colony was lysed by sonication 3 times in 15 s bursts with a resting time on ice between each cycle. A 1 mL aliquot from each sample was retrieved and microcentrifuged at max speed for 10 min before the supernatant was removed and tested for the hydrolysis of PNPA.

## **2.3 Results and Discussion**

### **2.3.1 pET100/D-topo and pET20b+ Expressions**

The pET100/D-topo expression vector includes a short N-terminal histidine tag. The pET20b+ expression vector possesses a C-terminal histidine tag, and the vector contains a pelB secretion sequence, which allows the possibility to secrete the protein. Several constructs were designed to evaluate possible fusion combinations and determine if any of these combinations would offer optimal expression of NCEH1. The different combinations of the tag and localization sequences present are described in Figure 2.1.

### NCEH1 *Escherichia coli* Constructs



**Figure 2.1:** Summary of the NCEH1 constructs created for recombinant expression. Each of the constructs varied either the location or identity of the fusion tag used to improve the localization, solubility, or purification of the NCEH1 protein. The first five constructs listed on the left side of the figure correspond to the pET100/D-topo and pET20b+ expression vectors. All constructs listed with a GST tag associated to NCEH1 were ligated into the pGEX-4T3 expression vector and all constructs associated with MBP are referred to as pMAL-c5x.

The advantage of the *pelB* localization sequence is that it presents the opportunity to increase the stability of the cloned gene by localizing it to the periplasm. Meanwhile, the histidine tag facilitates protein purification on a nickel affinity column. The constructs were designed to optimize expression by screening several conditions. As shown in Figure 2.1, the structure of the constructs range from containing no fusion to including both a periplasmic localization sequence and a small histidine tag. Constructs were prepared with a histidine tag on either side of the inserted gene. This was done to compensate for the possibility that the addition of a small fusion tag would interfere with the protein's folding pattern, causing it to become unstable. The N-terminal periplasmic localization sequence was added in an

attempt to increase the solubility of the NCEH1 gene if it was secreted into the periplasm of the cell. However, as discussed with the histidine tag, these additional amino acids may destabilize the protein. Therefore, additional constructs were created that excluded this localization sequence. The construct lacking a tag altogether was created in case an addition to either terminus of the protein resulted in its instability or loss of function.

Each of the constructs above were cloned, verified with DNA sequencing, transformed into BL21-CodonPlus(DE3)-RIPL cells and expressed from single colonies on a small scale. After 5 h of induced expression there was no detectable activity in the cells containing the NCEH1 gene when compared to the empty vectors. Each of the samples were analyzed via SDS PAGE, and no protein band of the appropriate molecular weight was detected that would indicate an overexpression of NCEH1. Based on this data, it was concluded that the pET vectors and our approach were not successful to express the NCEH1 gene. The truncation of the first 55 amino acids may limit proper folding or destabilize the protein, as suggested in Igarashi *et al.*<sup>2</sup> The results above support the conclusion of Igarashi *et al.* that this TM domain is important for stability. Consequently, small fusion tags may not compensate for the loss of this region, and something larger may be required to stabilize NCEH1.

### **2.3.2 Expression and Purification of pGEX-4T3**

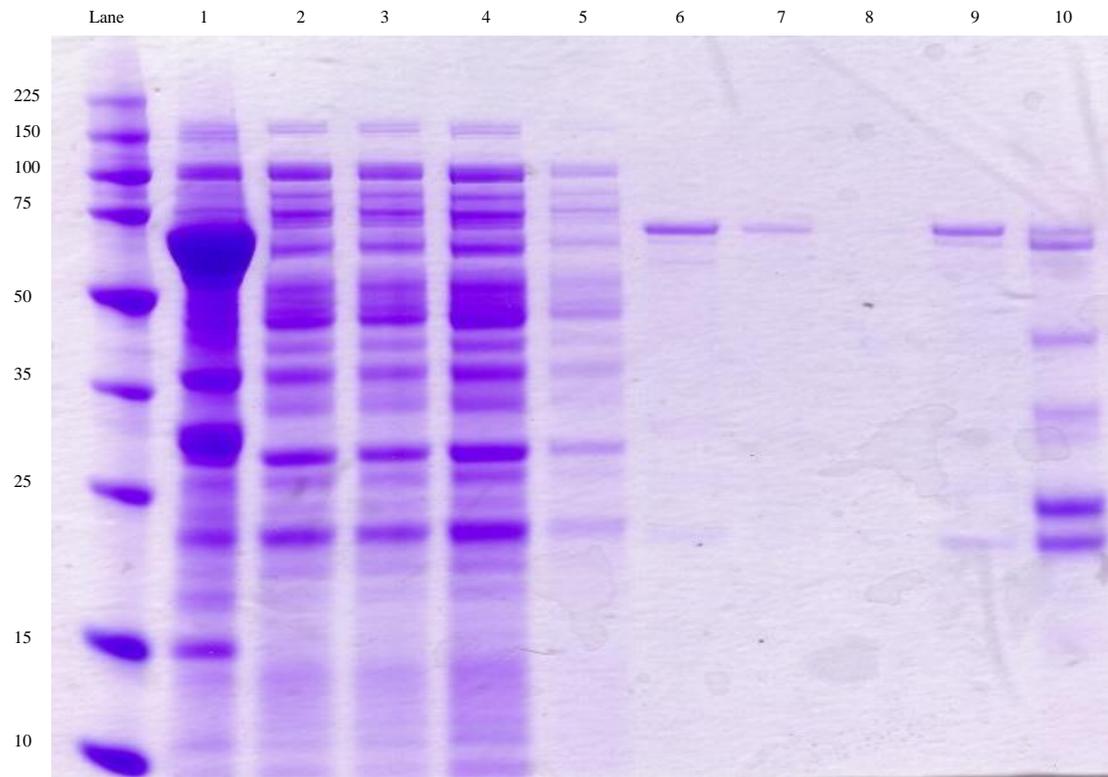
The pGEX expression vectors contain an N-terminal glutathione S-transferase (GST) fusion tag with a thrombin cleavage signal. The GST fusion tag, approximately

25 kDa in molecular weight, is a large fusion tag that could stabilize the fusion protein as well as simplify purification via affinity chromatography. As discussed above, a major predicted TM helix, amino acids 1-55, was truncated prior to the start of this work. However, when the full sequence of NCEH1 is inserted into the TM helix prediction model a second fairly hydrophobic region is predicted to span amino acids 56-100. In an effort to remove any possibility of instability due to a hydrophobicity, an additional truncation was made to remove the second hydrophobic region, leaving amino acids 115-440. Both constructs, also shown in Figure 2.1, were cloned, verified by DNA sequencing, and transformed into BL21-CodonPlus(DE3)-RIPL cells. Unlike the previous experimentation with the two pET expression vectors, the pGEX expression constructs demonstrated expression in small scale analysis. The presence of protein was verified via SDS PAGE and activity was confirmed within the crude lysate when compared to empty vector.

Figure 2.2 displays a gel showing the different stages of a typical sample purification. As mentioned above, NCEH1 displayed expression with the pGEX-4T3 vector, representing a significant step forward relative to the pET vector. However, a substantial amount of the expressed protein remained in the insoluble fraction. Exhaustive experimentation utilizing multiple detergents, detergent incubation times, expression temperatures, IPTG induction concentration, and the construction of the 115-440 NCEH1 truncation all failed to significantly increase the concentration of NCEH1 retrieved from the soluble fraction. Ultimately, purification of all 6 L of the large scale expression was necessary to retrieve even a small amount of the fusion protein. The sample was loaded onto the glutathione beads, and washed as detailed in the methods section above. Samples of each step were collected and analyzed for

fusion protein. In Figure 2.2, the unretained portion and two wash samples show very little fusion protein present in these samples. The NCEH1 was then eluted from the column, over the course of three sequential incubations with reduced glutathione and show relatively pure fusion protein at an apparent molecular weight of 70 kDa. Figure 2.2 also demonstrates the instability of the enzyme after incubation with thrombin protease. The protease was unable to efficiently cleave the entirety of the fusion protein, as evident with the remaining 70 kDa band, and the formation of several smaller bands, as well as a precipitate, indicates instability of the cleaved protein. In an attempt to further purify NCEH1 the sample was further incubated with both glutathione sepharose and benzamidine sepharose to remove free GST and the thrombin protease. Unfortunately as a serine hydrolase, NCEH1 also bound to the benzamidine column making it an inefficient purification strategy.

To improve the yield of the purification, the dialyzed sample containing the free GST, NCEH1 and thrombin was instead loaded onto an anion exchange column and separated utilizing FPLC. The three proteins were able to be separated utilizing this strategy; however, the protein yield remained too low for this to serve as a practical method.



**Figure 2.2:** SDS PAGE analysis of the purification of GST fused NCEH1 truncation 56-440. The protein molecular weight ladder is labeled to the left of the gel while each of the lanes are identified at the top of the gel. Lane 1 and 2 represents the solubility of NCEH1 as the pellet and supernatant of the cellular purification. Lanes 3-5 represent the unretained supernatant and washes respectively and lanes 6-8 represent the elutions of the NCEH1 fusion protein. Lane 9 represents the protein before the thrombin cleavage while lane 10 shows the breakdown of the protein after thrombin cleavage.

The conclusion reached for the pGEX-4T3 purification trials was that despite a dramatic increase in the expression of NCEH1 relative to the pET vector, pGEX was also not an optimal vector for expression. Furthermore, the thrombin protease was unable to completely cleave the produced fusion protein and the protein that was cleaved was unable to maintain stability in solution. Therefore, the pMAL-c5x vector was pursued as a viable alternative option to the pGEX vector.

### **2.3.3 Expression and Purification of pMAL-C5x**

The pMAL expression system includes an N-terminal maltose binding protein (MBP) fusion tag with a Factor Xa cleavage signal. The vector was generously offered from Dr. Sharon Rozovosky (University of Delaware) and had been modified to include a histidine tag at the N-terminus of the MBP tag. Studies have suggested that MBP has the ability to increase the solubility of an otherwise insoluble protein<sup>8</sup> and at 42 kDa still provided a large, stable and soluble fusion protein to promote the expression and stability of NCEH1. Three truncations 56-440, 115-440, and 165-440 NCEH1 constructs were cloned, verified with sequencing, and transformed into BL21-CodonPlus(DE3)-RIPL cells. Truncations 56-440 and 115-440 were both designed to remove hydrophobic regions. Truncation 165-440 was designed to mimic the third possible isoform of NCEH1. Truncations 56-440 and 115-440 both showed expression at the projected molecular weight, approximately 80 and 75 kDa respectively, in small scale expression trials. However 165-440 never showed any useful level of expression; thus, suggesting that too much of the N-terminus had been removed even with the MBP tag for stability. The large scale expressions showed a small increase in the solubility of NCEH1, but the largest fraction of NCEH1 produced remained in the insoluble fraction. The protein was purified utilizing a Ni-sepharose gravity flow

column and fusion protein was eluted from the column with high concentrations of imidazole. The resulting solution was concentrated and dialyzed into a buffer for Factor Xa cleavage containing 20 mM Tris-Cl, 100 mM NaCl, and 2 mM CaCl<sub>2</sub>, pH 8.0. The protein was incubated with Factor Xa, and once again displayed incomplete cleavage and instability once cleaved from the MBP fusion protein.

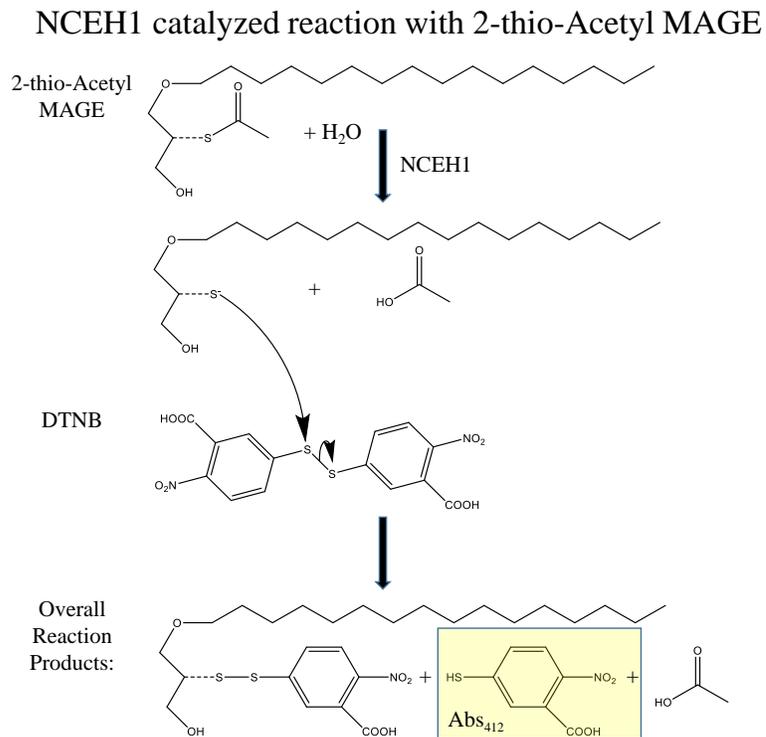
The inability to effectively purify NCEH1 which had been cleaved from its MBP fusion protein led to the decision to pursue characterization with the full MBP-NCEH1 fusion. Aspects of expression and purification were modified in an effort to increase the yield of the target NCEH1 fusion protein. An immediate increase in the yield of NCEH1 was realized by switching from Ni-sepharose to Ni-NTA beads. Ni-NTA also provided the additional advantage of decreasing the nonspecific binding within the purification process, thus increasing the purity of the sample. However, the yield of the MBP-NCEH1 fusion protein remained low, with concentrations typically around 0.08 mg/L of expression. This low yield led to an exploration of which variables within the expression protocol could be further optimized.

The major limitation on the yield of NCEH1 remained to be the expression of the mammalian protein within a recombinant system leading to an accumulation of NCEH1 in the insoluble fraction. Oganessian *et al.* evaluated the effect of osmotic stress and heat shock in recombinant protein overexpression, and observed that it led to an increase in the solubility of previously insoluble proteins. Their findings suggested that the addition of a short 30 min heat shock at 47 °C increases the *E. coli* heat shock protein content, effectively increasing the availability of chaperone proteins to facilitate stable folding of the induced protein. In addition to this finding, they noted that increasing the NaCl concentration from 170 to 500 mM in LB also increased the

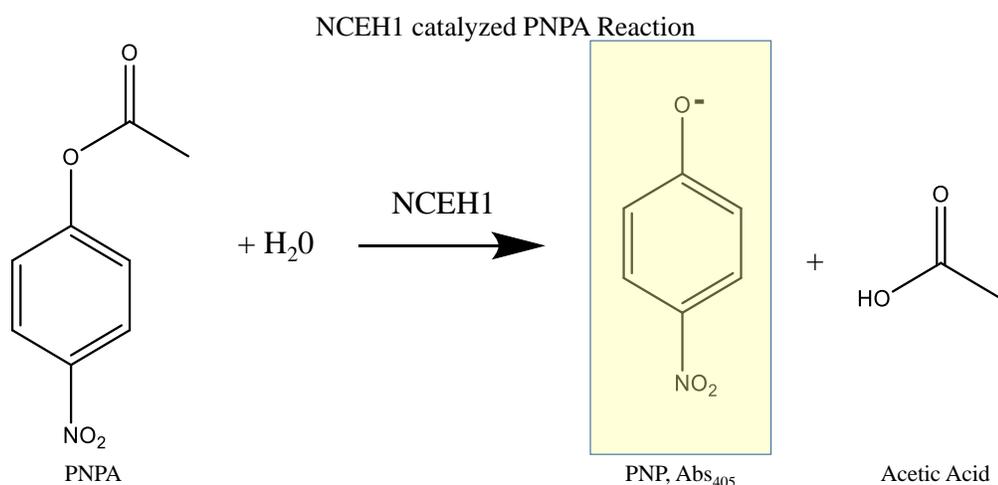
solubility of previously insoluble proteins<sup>9</sup>. Therefore, both approaches were adopted into the expression protocol in an attempt to increase the solubility of NCEH1. The large scale flasks were still grown to an OD<sub>600</sub> of approximately 1 at 37°C. However, once an OD<sub>600</sub> of 1 was obtained, the cells were incubated for 30 min at 47°C before being incubated on ice for 1 h. The cells were induced with IPTG and incubated shaking overnight at room temperature. These modifications increased the typical purified expression yields from an average of 0.08 mg/L of expression to 0.15 mg/L of expression. While still low, these yields were more suitable for the functional characterization of the NCEH1 fusion protein.

#### **2.3.4 Functional Characterization of NCEH1**

It has been suggested that NCEH1 requires glycosylation for functionality<sup>2</sup>, therefore NCEH1 originating from a recombinant system lacking the ability to glycosylate should show no activity. To test this hypothesis, NCEH1 was evaluated with a modified natural substrate, 2-thio-acetyl MAGE. This molecule, and the relevant reaction chemistry, is shown in Figure 2.3. Activity was also tested using the common serine esterase substrate para-nitrophenyl acetate (PNPA), whose structure and chemistry is shown below in Figure 2.4. As mentioned above, 2-thio-acetyl MAGE is a substrate specific to the NCEH1 protein. Therefore, the assay utilizing this substrate was employed as a primary method to determine whether the protein purified from the affinity column was active NCEH1. When the sample was tested, activity was observed, data not shown, and further studies to evaluate  $K_M$ ,  $V_{max}$ , and turnover were begun with the PNPA substrate.



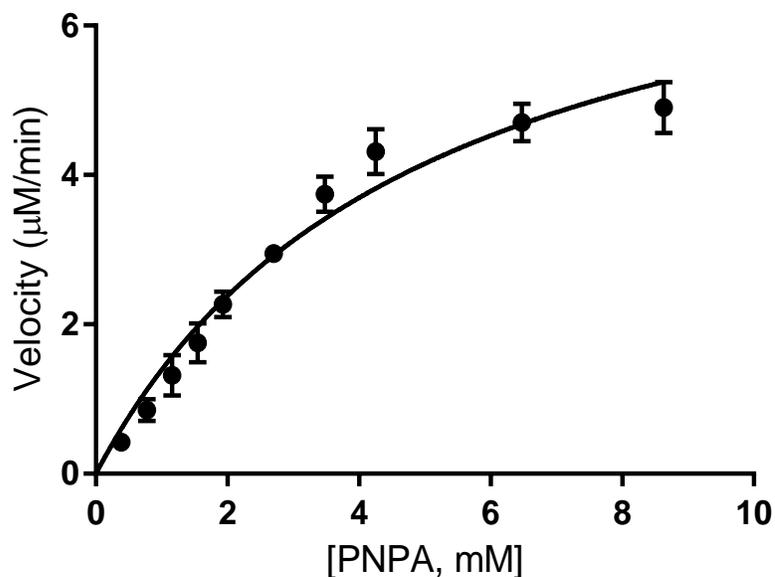
**Figure 2.3:** The NCEH1 catalyzed reaction of the modified natural substrate 2-thio-acetyl MAGE with water. The natural substrate, acetyl MAGE simply has an oxygen in the position of the sulfur creating an ester bond, this modification is demonstrated with a dashed line. The reaction involves the hydrolysis of the thiol bond in 2-thio-acetyl MAGE to create a free thiol and acetic acid. The free thiol can then attack the disulfide bond of DTNB resulting in a product that can be spectroscopically monitored at 412 nm.



**Figure 2.4:** NCEH1 catalyzed ester cleavage of para-nitrophenyl acetate to para-nitrophenol and acetic acid. Formation of para-nitrophenol can be spectroscopically monitored for an increase in absorbance at 405 nm.

The assays were carried out at room temperature in 1 mL volumes of 50 mM Tris buffer, pH 7.8 with known volumes of 100 mM PNPA. The assay was started with the addition of purified NCEH1 fusion protein to solution and the change in absorbance was monitored at 405 nm for 30 s. The resulting data was analyzed with GraphPad prism 7 and fitted with a nonlinear graphing function to obtain predicted values for  $K_M$  and  $V_{max}$ . The  $V_{max}$  was utilized to calculate the predicted  $k_{cat}$  of NCEH1 for the conversion of PNPA to PNP. An example of the nonlinear fit and the calculated values for  $K_M$ ,  $V_{max}$  and  $k_{cat}$  are shown below in Figure 2.5.

## Michaelis-Menten data



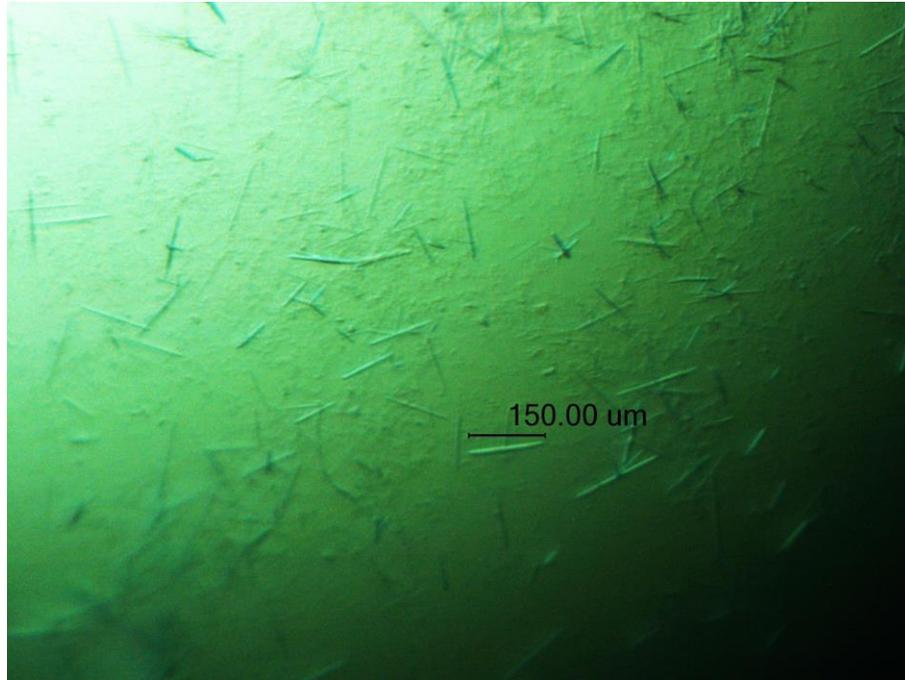
	$K_M$ (mM)	$V_{max}$ (µM/min)	$k_{cat}$ (s <sup>-1</sup> )
Average	5.3	6.6	4.02
Standard Deviation	0.6	0.9	0.12

**Figure 2.5:** Example of the functional characterization of NCEH1 on the substrate PNPA. The predicted  $K_M$  and  $V_{max}$  values were calculated from GraphPad Prism 7, while the  $k_{cat}$  value is calculated from the respective  $V_{max}$  and enzyme concentration used within the assay. The results are the average of three separate assays performed from three separate MBP\_56-440 NCEH1 fusion protein purifications.

The results of the functional characterization of the fusion MBP\_56-440 NCEH1 on PNPA showed a predicted  $K_M$  of approximately 5.3 mM,  $V_{max}$  of 6.6 µM/min, and a turnover of 4.02 s<sup>-1</sup>. The characterization was also carried out on the truncated 115-440 NCEH1 protein and no significant changes were noted (data not shown).

### **2.3.5 Protein Crystallization Screening of NCEH1**

Protein samples of the MBP fusion for both 56-440 and 115-440 NCEH1 truncations obtained at a concentration higher than 0.5 mg/mL and with a purity greater than 90% were used to lay crystal screens. Hanging drop crystal screens were laid with the Hampton Research screen kits 1 and 2 with 1  $\mu$ L of purified sample mixed with 1  $\mu$ L of each individual well's solution. Each crystal tray was incubated at room temperature and evaluated for the presence of crystals regularly. Within two weeks the MBP\_115-440 NCEH1 truncation yielded very small microcrystals, shown in Figure 2.6, in the condition corresponding to 0.2 M lithium sulfate monohydrate, 0.1M Tris hydrochloride pH 8.5, and 30% polyethylene glycol 4,000.



**Figure 2.6:** MBP\_115-440 NCEH1 fusion microcrystals. Crystals visualized after two weeks in condition 17 of Hampton Research screen 1. The solution contained 0.2 M lithium sulfate monohydrate, 0.1 M Tris hydrochloride pH 8.5, and 30% polyethylene glycol 4,000.

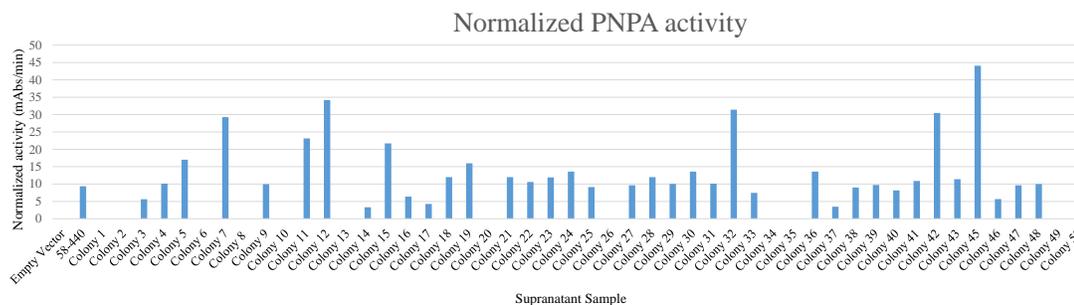
Regretfully, screening around the conditions of the well solution yielding the microcrystals failed to yield any crystals large enough to be of interest. As a result, further screening of the fusion protein for crystallography purposes was suspended until a higher protein yield could be obtained.

### **2.3.6 Random Mutagenesis**

Previous experimentation designed to increase the solubility and yield of NCEH1 included the manipulation of the expression and purification protocols. However, mutation of the NCEH1 protein itself has not been explored. Mutations to

the NCEH1 gene yield the possibility of increasing the protein's stability and solubility within the recombinant system. Site directed mutagenesis was not pursued because of the lack of a reliable structure for NCEH1; however, random mutagenesis provided the possibility of randomly inserting mutations that may aide in the stability or solubility of the protein.

Random mutagenesis was carried out using the truncation 56-440 DNA template of NCEH1. PCR primers pMAL56forward and KIAA-reverse from Table 2.1 were utilized for the amplification of DNA. To induce mutagenesis, a low fidelity taq polymerase was utilized as well as supplementation with 10 mM mutagenic dNTPs 8-oxo-2'-deoxyguanosine-5'-triphosphate and 2'-deoxy-P-nucleoside-5'-triphosphate. The mutagenic dNTPS introduce mutations at an approximate amount of 6-19% for 10-30 cycles of PCR<sup>10</sup>. Two separate tubes were prepared for PCR amplification with fresh template being added after 10 cycles to maximize the mutation library. After 20 cycles the two tubes were combined to create one library and this sample was used moving forward following standard PCR protocols for the amplification of the DNA library. Amplified mutagenic DNA was inserted into the pMAL\_c5X expression vector and transformed directly into BL21-CodonPlus(DE3)-RIPL cells for expression screening. Individual colonies were selected for small scale expression and resulting supernatants were tested for PNPA activity compared to wild type 56-440 NCEH1 and empty vector as shown in Figure 2.7.



**Figure 2.7:** Hydrolysis activity of PNPA in mAbs/min by NCEH1. Normalizes activity present in small scale expression supernatants normalized to empty pMAL\_c5x. The figure shows empty vector in the first column, followed by wild type 56-440 NCEH1 fusion protein, and 50 colonies chosen at random.

The three colonies with the highest activity for PNPA were selected, expressed, and purified for NCEH1. The corresponding retrieved yields were approximately 0.47 mg/L of expression, 0.43 mg/L of expression, and 0.52 mg/L of expression. Each of these purifications offered some confidence that an increase in either solubility or stability was achieved. Each of these colonies produced a significantly higher yield than the 0.15 mg/L of expression that was previously being purified. The increase in yield further progresses our ability to obtain a high enough concentration of pure protein for protein crystallization trials. However, there is a great deal of work that remains to be done in order to fully characterize the results of this technique. For instance, at this point in time it is unknown if the increase in protein yield is in direct correlation to a mutation(s) or because the result of uncontrolled variation in the expression conditions within the cells themselves. The possibility of the measured increase may also be a misrepresented artifact and not a result of random mutagenesis. Additionally, if the individual colonies do contain

mutations that lead to either increased stability or solubility of NCEH1, it is unknown if the mutations are the same mutation or separate entirely.

Significant work still needs to be completed within this portion of the project. Each of the plasmids needs to be prepared and re-inserted into DH5 $\alpha$  cells for propagation and sequencing analysis. The resulting sequences will be checked for mutations and all mutations present will be evaluated against a homology model for information regarding their possible effects on the NCEH1 gene.

## **2.4 Conclusions**

Igarashi *et al.* predicted that the N-terminus amino acids 1-33 of NCEH1 are required for its stability.<sup>2</sup> In the work described above, this hypothesis was supported when multiple constructs of the NCEH1 truncated protein were unable to be expressed when inserted into the pET expression vectors. The addition of the small histidine tag was not a significant enough stabilizing force to assist in the folding of the NCEH1 protein, and the protein could not be obtained in a functional form using this construct. The protein was able to be expressed within the pGEX and pMAL expression systems; however, several limitations arose. The protein benefitted from the presence of the glutathione transferase and maltose binding proteins fusion tags, but a large portion of the expressed protein remained insoluble. Expression conditions were modified to include a small heat shock response and change the osmolarity of the system with the goal of increasing the solubility of NCEH1. These modifications increased the soluble yield of protein by 80% from 0.08 mg/L of expression to approximately 0.15 mg/L of expression. While this marks a significant increase, this value is still low to effectively screen protein crystallization conditions to obtain suitable crystals for X-ray crystallography structural work and only microcrystals were observed. The thrombin

and Factor Xa proteases were also shown to be ineffective at quantitatively cleaving the fusion protein. Furthermore, this work showed that the cleaved protein becomes unstable and begins to fragment. Due to the physical limitations faced in trying to purify NCEH1 cleaved from its fusion partner, functional characterization of MBP\_56-440 NCEH1 fusion protein was pursued. This work yielded kinetic characterization of NCEH1 with a  $K_M$  value of 5.3 mM,  $V_{max}$  of 6.6  $\mu\text{M}/\text{min}$ , and a  $k_{cat}$  value of 4  $\text{s}^{-1}$ . The activity did not appear to be affected by further truncation of the N-terminus to amino acid 115. In addition, the measured activity of recombinant NCEH1 that is not glycosylated with both PNPA and 2-thio-acetyl MAGE substrates is in direct contrast to the work that states that glycosylation is required for activity of NCEH1<sup>1, 2</sup>. Based on these results, it seems more likely that glycosylation is required for folding and subsequent stability of NCEH1, and not necessarily functionality itself.

The main goal of the neutral cholesterol ester hydrolase project continues to be crystallization of the protein to obtain its structure. In an effort to further increase the yield of NCEH1, random mutagenesis was performed on the DNA of the 56-440 truncation. Mutagenic DNA was transformed directly into BL21-CodonPlus(DE3)-RIPL cells to screen multiple colonies for an increase in solubility or stability. Three colonies were selected from this data for a large scale expression and purification. On average, these colonies yielded approximately 0.47 mg/L of expression, significantly higher than the 0.15 mg/L of expression obtained from previous work using unmutated DNA. However, the apparent increase in NCEH1 solubility was not confirmed to be directly linked to the mutagenesis.

Future work for the recombinant portion of the NCEH1 project will require inserting the mutagenic NCEH1\_pMAL-C5X expression vector into DH5 $\alpha$  cells for

sequencing analysis. At that point any mutations present can then be compared to a homology model of NCEH1 to evaluate the cause of their effectiveness. The overall protein yield of NCEH1 will need to be much higher for successful crystallization from a recombinant system. At this point in time, the inability to retrieve high levels of soluble NCEH1 reduces the effectiveness of recombinant expression for the pursuit of crystallographic structure determination. In an effort to increase the overall yield of NCEH1 protein, the next chapter will explore NCEH1 expression in a *Pichia pastoris* system.

## REFERENCES

- [1] Jessani, N., Liu, Y. S., Humphrey, M., and Cravatt, B. F. (2002) Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness, *Proceedings of the National Academy of Sciences of the United States of America* 99, 10335-10340.
- [2] Igarashi, M., Osuga, J.-i., Isshiki, M., Sekiya, M., Okazaki, H., Takase, S., Takanashi, M., Ohta, K., Kumagai, M., Nishi, M., Fujita, T., Nagai, R., Kadowaki, T., and Ishibashi, S. (2010) Targeting of neutral cholesterol ester hydrolase to the endoplasmic reticulum via its N-terminal sequence, *Journal of Lipid Research* 51, 274-285.
- [3] Okazaki, H., Igarashi, M., Nishi, M., Sekiya, M., Tajima, M., Takase, S., Takanashi, M., Ohta, K., Tamura, Y., Okazaki, S., Yahagi, N., Ohashi, K., Amemiya-Kudo, M., Nakagawa, Y., Nagai, R., Kadowaki, T., Osuga, J.-i., and Ishibashi, S. (2008) Identification of Neutral Cholesterol Ester Hydrolase, a Key Enzyme Removing Cholesterol from Macrophages, *Journal of Biological Chemistry* 283, 33357-33364.
- [4] Buchebner, M., Pfeifer, T., Rathke, N., Chandak, P. G., Lass, A., Schreiber, R., Kratzer, A., Zimmermann, R., Sattler, W., Koefeler, H., Froehlich, E., Kostner, G. M., Birner-Gruenberger, R., Chiang, K. P., Haemmerle, G., Zechner, R., Levak-Frank, S., Cravatt, B., and Kratky, D. (2010) Cholesteryl ester hydrolase activity is abolished in HSL<sup>-/-</sup> macrophages but unchanged in macrophages lacking KIAA1363, *Journal of Lipid Research* 51, 2896-2908.
- [5] Ross, M. K., Pluta, K., Bittles, V., Borazjani, A., and Crow, J. A. (2016) Interaction of the serine hydrolase KIAA1363 with organophosphorus agents: Evaluation of potency and kinetics, *Archives of Biochemistry and Biophysics* 590, 72-81.
- [6] Sekiya, M., Osuga, J.-i., Nagashima, S., Ohshiro, T., Igarashi, M., Okazaki, H., Takahashi, M., Tazoe, F., Wada, T., Ohta, K., Takanashi, M., Kumagai, M., Nishi, M., Takase, S., Yahagi, N., Yagyu, H., Ohashi, K., Nagai, R., Kadowaki, T., Furukawa, Y., and Ishibashi, S. (2009) Ablation of Neutral Cholesterol Ester Hydrolase 1 Accelerates Atherosclerosis, *Cell Metabolism* 10, 219-228.
- [7] Igarashi, M., Osuga, J.-i., Uozaki, H., Sekiya, M., Nagashima, S., Takahashi, M., Takase, S., Takanashi, M., Li, Y., Ohta, K., Kumagai, M., Nishi, M., Hosokawa, M., Fledelius, C., Jacobsen, P., Yagyu, H., Fukayama, M., Nagai,

- R., Kadowaki, T., Ohashi, K., and Ishibashi, S. (2010) The Critical Role of Neutral Cholesterol Ester Hydrolase 1 in Cholesterol Removal From Human Macrophages, *Circulation Research* 107, 1387-1395.
- [8] Weis, W., Benhar, I., Oberoi, P., Jabulowsky, R., and Dalken, B. (2010) Maltose-Binding Protein Enhances Secretion of Recombinant Human Granzyme B Accompanied by *In Vivo* Processing of a Precursor MBP Fusion Protein., PLoS ONE.
- [9] Oganessian, N., Ankoudinova, I., Kim, S. H., and Kim, R. (2007) Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization, *Protein Expression and Purification* 52, 280-285.
- [10] Zaccolo, M., Williams, D. M., Brown, D. M., and Gherardi, E. (1996) An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues, *Journal of Molecular Biology* 255, 589-603.

## Chapter 3

### PURIFICATION OF NCEH1 FROM *PICCHIA PASTORIS*

#### 3.1 Introduction

*Escherichia coli* overexpression approaches have been used in the production of proteins of interest for decades. However, broad application of this system with eukaryotic genes tends to be problematic. This is due to the lack of machinery for both posttranslational modification and eukaryotic protein folding. Therefore, the attempt to express eukaryotic proteins in *E. coli* frequently results in the expression of the protein of interest into inclusion bodies.<sup>1</sup> As noted in the previous chapter, expression of NCEH1 in an *E. coli* expression system yielded a significant amount of protein in the insoluble fraction. Thus, the *E. coli* model system was found to be inadequate, as the main goal of this project was to obtain pure and high yields of NCEH1 for crystallization and structural characterization.

Leuking *et al.* explored the comparison of 29 separate cDNA clones in *E. coli* and *Pichia pastoris*. Their work showed a strong increase in the soluble expression in *P. pastoris* when compared to the *E. coli* expression system. All 29 clones were successfully expressed into the soluble fraction with *P. pastoris* while only 9 were detected in the soluble component of the *E. coli* expression.<sup>2</sup>

As discussed in chapter 2, the *E. coli* expression of NCEH1 was not only heavily biased toward the insoluble fraction, but the soluble NCEH1 also became highly unstable without the presence of a large protein fusion tag. This observation led to the conclusion that the fusion tag, and possibly posttranslational modifications, stabilize the protein. For this reason *P. pastoris* was investigated as an expression system for NCEH1. The purpose of this chapter is to highlight the benefits of

expression and purification of NCEH1 within the *P. pastoris* system. This chapter will highlight the creation of a successful expression and purification strategy, while exploring further structural and functional characterization of NCEH1.

## **3.2 Materials and Methods**

### **3.2.1 Preparation of Competent *Pichia Pastoris* Cells**

Competent cells for the transformation of *P. pastoris* need to be freshly prepared before each transformation. Each of the three cell lines, X-33, GS115, and KM71H, were prepared following the following protocol. The cells were grown up overnight in 5 mL of yeast peptide dextrose (YPD) at 30 °C. The overnight was used to inoculate 500 mL of fresh YPD, and the fresh cells were grown to an OD<sub>600</sub> of 1.3-1.5 before being centrifuged at 3,000 rpm for 5 min at 4 °C. The pellet was resuspended in 500 mL of ice-cold, sterile water and centrifuged a second time. The pellet was then re-suspended in 250 mL of ice-cold, sterile water, centrifuged again, and re-suspended in 20 mL of ice-cold 1 M sorbitol. The process was completed with a final centrifugation at 3,600 rpm and the pellet was resuspended into 1 mL of ice cold sorbitol.

### **3.2.2 Cloning of NCEH1 into pPICZ $\alpha$ and pPICZ**

The pPICZ $\alpha$  and pPICZ expression vectors were generously obtained from Dr. Colin Thorpe (University of Delaware, Newark, DE) and originated from the Invitrogen *Pichia pastoris* expression kit. The primers used for the amplification of NCEH1 are listed in Table 3.1 with the corresponding sequence and restriction enzymes, and the primers utilized in the cloning of NCEH1 into *Pichia pastoris*. The forward primers PiczAfor 58, PiczAfor111, PiczAlpha56, PiczAlpha11 were each

used for insertion into their corresponding pPICZ/ $\alpha$  expression vector with an N-terminal truncation at either amino acid 56 or 111. The pPICZ expression vector lacks an initiation start codon; therefore, the forward primers PiczAfor58 and PiczAfor111 each had to contain a yeast consensus sequence. This consensus sequence is bolded and underlined in Table 3.1. PiczANot1Rev was utilized as a reverse primer to create a construct without a protease cleavage site, leaving the short c-myc tag and histidine tag intact. A TEV protease site was also engineered into both vectors with the use of the KiaaRevTev1, PicZalphaTev2, and PiczANot1Tev2 reverse primers.

Primer Name	Primer Sequence 5'-3'	Restriction Site
PiczAfor58	AAAAAGAATTCA <b><u>AAAATAATGTCT</u></b> TCCGTGTCCGACCCC	ECORI
PiczAfor111	AAAAAGAATTCA <b><u>AAAATAATGTCT</u></b> TAAGGTGACCGACACAG	ECORI
PiczArev	AAAAAGGGCCCCAGGTTTTGATCTAGCCACTTG	APAI
PiczAlpha56	AAAAAGAATTCCCTGGCTCCGTGTCCGACCCC	ECORI
PiczAlpha111	AAAAAGAATTCCAAGTGAAGGTGACCGACACAG	ECORI
KiaaRevTev1	GTAGAGGTTCTCCAGGTTTTGATCTAGCCACTTG	
PiczAlphaTev2	TTTTTTCGGCCGCGGACTGGAAGTAGAGGTTCTC	NOTI
PiczANot1Tev2	TTTTTTCGGCCGCGGGACTGGAAGTAGAGGTTCTC	NOTI
PiczANot1Rev	AAAAAGCGCCGCGCAGGTTTTGTTGATCTAGCCACTTG	NOTI

**Table 3.1:** Primers utilized in the cloning of NCEH1 into *P. pastoris*. Forward primers PiczAfor 58, PiczAfor111, PiczAlpha56, PiczAlpha11 were each used as the forward primers for insertion into their corresponding pPICZ/ $\alpha$  expression vector with either a truncation at amino acid 56 or 111. PiczANot1Rev was utilized as a reverse primer to create a construct with no protease cleavage site, leaving the short c-myc tag and histidine tag intact. A tev protease site was also engineered into both vectors with the use of the KiaaRevTev1, PicZalphaTev2, and PiczANot1Tev2 reverse primers.

The DNA was amplified using PCR, digested with the corresponding restriction enzymes, gel purified, and ligated overnight at 14 °C into the corresponding vector. The ligated vector containing NCEH1 was transformed into DH5 $\alpha$  *E. coli* cells by electroporation for amplification purposes. Single colonies containing the plasmid of interest were grown overnight on LB agar plates containing 25  $\mu$ M Zeocin

(Invitrogen). Plasmids isolated from single colonies were sequenced by Genewiz, Inc., and purified plasmids verified to contain NCEH1 were linearized with a single restriction enzyme digestion with PME1 (New England Biolabs) for 3 h at 37 °C. The linearized DNA was purified with a Qiagen Mini prep kit prior to being mixed with one of the following *Pichia pastoris* competent cell strains: X-33, KM71H, GS115, and stored on ice for 5 min. The cellular mixture was electroporated with a single 1.5 kV pulse, mixed with 1 mL of ice-cold sorbitol and incubated at 30 °C, without shaking, for 2 h. The cells were plated on YPD agar containing 100 µM of zeocin, incubated at 30 °C for 3-5 days until the appearance of single yeast colonies was observed. Colonies were chosen and grown up at 30 °C. Each colony was incubated with zymolase (Sigma-Aldrich) to cleave the yeast cell wall and NCEH1 was amplified utilizing primers specific for the AOX gene within the expression vector. The PCR products were used for the verification of the NCEH1 sequence by Genewiz, Inc.

### **3.2.3 Glycosylation Mutagenesis in pPICZ $\alpha$**

NCEH1 has three predicted Asn glycosylation sites located at amino acids 270, 287, and 389. Mutagenesis was performed at each of these sites by changing asparagine to glutamine. Glutamine is similar in structure and charge to asparagine, but lacks the ability to be glycosylated. Site-directed mutagenesis was carried out with the PCR primers listed in Table 3.2. Template DNA from NCEH1, already ligated into pPICZ $\alpha$ , was utilized with a PCR elongation time of 12 min per run to achieve complete replication of the pPICZ $\alpha$  plasmid. The mutagenesis was completed over 15 cycles, and the resulting amplified DNA was subjected to DPN1 digestion for

2 h at 37 °C to digest the original wild type DNA. The sample was then gel purified and the purified DNA was mixed with chemically competent DH5a *E. coli* cells. The cell mixture was heat shocked for 45 s at 42 °C and incubated on ice for 5 min before 1 mL of LB was added to the cells and incubated, shaking at 37 °C for 1 h. The cells were plated on LB agar containing 25 µM Zeocin overnight at 37 °C. Single colonies were selected for sequencing analysis before being inserted into the *Pichia pastoris* KM71H strain as previously described.

Primer Name	Primer Sequence 5'-3'	Original Sequence
ForwardK270Q	CGTTAAC <b>CA</b> ACACTTCACTTG	AAT
ReverseK270Q	CAAGTGAAGTGTG <b>TTGG</b> TTAACG	
ForwardK287Q	GGGCCCGTCTA <b>CAAT</b> TGGACATCC	AAC
ReverseK287Q	GATGTCC <b>ATTG</b> TAGACGGGCC	
ForwardK389Q	CTGGCCACCC <b>CAAT</b> TCTCAGTGG	AAC
ReverseK389Q	CCACTGAGAA <b>TTGG</b> TGGGCCAG	

**Table 3.2:** Table showing the PCR mutagenic primers for the N to Q mutagenesis. Asparagine residues mutated located at amino acids 270, 287, and 389. The table displays the primer name, sequence in the 5' to 3' direction, and the original amino acid sequence of the DNA. The mutagenic DNA is bolded and underlined within the primer sequence.

### 3.2.4 Expression of NCEH1 in *Pichia pastoris*

Overnights of the NCEH1 constructs in both pPICZ and pICZα were grown up at 30 °C and used to inoculate six 4 L flasks containing 1 L of buffered minimum glycerol media 100 mM potassium phosphate, pH 6.0, 1.3% yeast nitrogen base, 2.5 µM Biotin, 1% glycerol. The 4 L flasks were grown up for 24 h at 30 °C before being pelleted by centrifugation. Each pellet was resuspended in 200 mL of buffered minimum methanol media, 100 mM potassium phosphate, pH 6.0, 1.34% yeast

nitrogen base, 2.5  $\mu$ M Biotin, 0.5% methanol. The expression was then continued at 30 °C. Expression of pPICZ $\alpha$  was supplemented with 100 mM sorbitol at 12 h intervals during the expression. The pellets of the constructs within the pPICZ vector were harvested by centrifugation and stored at -80 °C for further experimentation. The various constructs from the pPICZ $\alpha$  had their supernatant stored at -80 °C for further experimentation.

### **3.2.5 Purification of NCEH1 in *Pichia pastoris***

#### **3.2.5.1 pPICZ**

Once expressed pPICZ pellets containing NCEH1 were re-suspended in 50 mM sodium phosphate buffer, 300 mM NaCl, pH 7.4. The cells were lysed by sonication using three 1 min cycles at 50% power with a 5 min rest on ice between each cycle. Tween 20 was added to a final concentration of 1% and the suspension was rocked at 4 °C for 20 min. The membrane portion was then removed by two rounds of centrifugation for 20 min at 23000 x g. After the first spin, the supernatant was added to a fresh tube and spun again. To bind the His-tag of NCEH1, the supernatant was then incubated on a gravity chromatography column containing Ni-sepharose beads (GE Healthcare) for 1 h. The unretained supernatant was removed and the column was washed with a solution containing 50 mM sodium phosphate buffer and 300 mM NaCl at pH 7.4. A second wash was performed using a solution containing 50 mM sodium phosphate buffer, 300 mM NaCl and 100 mM imidazole at pH 7.4. The protein was eluted from the column using three 10 mL aliquots of a solution containing 50 mM sodium phosphate buffer, 300 mM NaCl and 300 mM imidazole at pH 7.2. The eluted protein samples were combined and concentrated to 3

mL before being dialyzed overnight into 10 mM HEPES containing 10 mM NaCl at pH 7.4.

### **3.2.5.2 pICZ $\alpha$**

The supernatant of pICZ $\alpha$  containing the expressed NCEH1 constructs was dialyzed against 50 mM sodium phosphate buffer, pH 7.4, containing 300 mM NaCl. The dialyzed supernatant was then applied to a 5 mL nickel sepharose FPLC column (GE Healthcare). The column was washed with 50 mM sodium phosphate buffer, pH 7.4, with 300 mM NaCl and eluted with a 0-300 mM gradient of imidazole. Fractions containing protein were evaluated for activity using the 2-thio-acetyl MAGE assay described in Chapter 2, and the fractions containing NCEH1 were pooled and concentrated. The purified NCEH1 was dialyzed overnight into 10 mM HEPES containing 10 mM NaCl at pH 7.4.

### **3.2.6 Functional Characterization of NCEH1**

The concentration of the purified NCEH1 was determined by a BCA assay (Pierce) and its purity was determined via SDS PAGE coupled to ImageJ software.<sup>3</sup> The activity of the purified protein was determined by analysis with PNPA and 2-thio-acetyl MAGE. PNPA assays were performed in 1 mL total volume in 50 mM Tris buffer, pH 7.8, with approximately 10 mM PNPA. The assay was monitored at 405 nm for 30 s following the addition of NCEH1. The hydrolysis of 2-thio-acetyl MAGE is followed by monitoring a coupled reaction with DTNB, described previously in chapter 2, monitored for 300 s at 412 nm. The assay was completed in 1 mL volumes containing 50 mM sodium phosphate buffer pH 7.8, 10 mM DTNB, and various amounts of substrate.

### 3.2.7 Structural Characterization of NCEH1

Purified protein obtained at a concentration higher than 0.5 mg/mL and greater than 90% purity, as determined by ImageJ, was utilized for crystal screening trials. Initial hanging drop screening was completed with crystal screens 1 and 2 from Hampton Research and stored at room temperature. To prepare protein samples for crystal screening each sample was forced through a syringe filter before hanging drops were prepared by mixing a 1  $\mu$ L of protein sample with a 1  $\mu$ L of well solution. When necessary, certain conditions were further screened with the additive screen kit also provided by Hampton Research. All screens were closely monitored for the formation of crystals and all solutions yielding crystals were recreated in the laboratory for further screening. The crystallography solutions recreated in the laboratory are listed in Table 3.3, along with their corresponding Hampton Research kit well solution number. Sodium acetate trihydrate, calcium acetate hydrate, sodium cacodylate, and polyethylene glycol (PEG) 4,000 and 8,000 were each from Sigma-Aldrich. The other chemicals utilized for crystallography, including ammonium sulfate and Tris hydrochloride, were purchased from Fisher Scientific.

Hampton Research Screen	Well Number	Condition
1	45	.2M Zinc acetate dihydrate, .1M sodium cacodylate trihydrate pH 6.5, 18% PEG 8,000
1	46	.2M Calcium acetate trihydrate, .1M sodium cacodylate pH 6.5, 15% PEG 8,000
2	22	.2M Sodium acetate trihydrate, .1M Tris hydrochloride pH 8.5, 30% PEG 4,000
2	30	.2M Ammonium Sulfate, 30% PEG 8,000

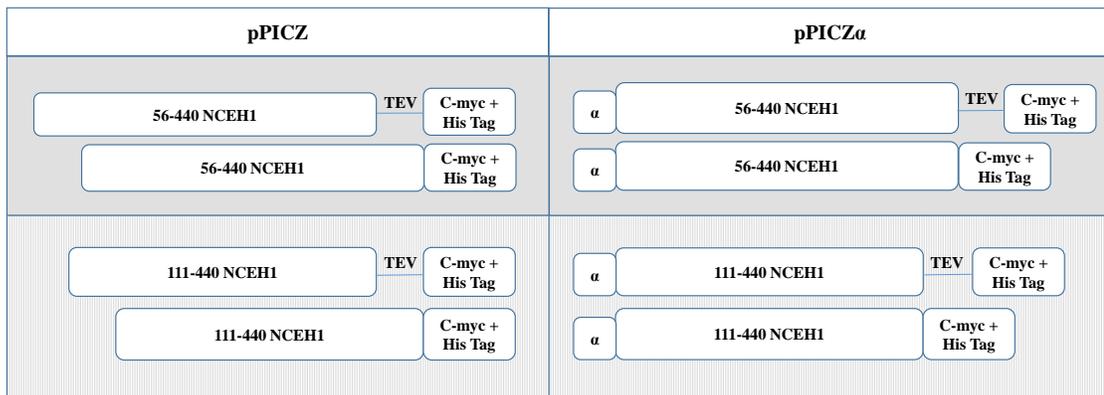
**Table 3.3:** Table of crystal conditions used for the optimization of crystallization of NCEH1.

### 3.3 Results and Discussion

#### 3.3.1 Expression Optimization of NCEH1

Heterologous expression of NCEH1 in *Pichia pastoris* could be achieved either by intracellular expression using pPICZ, or by secreted expression, using pPICZ $\alpha$ . DNA of the truncated 56-440 or 111-440 NCEH1 was amplified, digested, and ligated into both of the expression vectors for the purpose of defining optimal expression conditions. Figure 3.1 demonstrates the various constructs designed for the expression of NCEH1 in *Pichia pastoris*.

**NCEH1 *Pichia pastoris* Expression Constructs**



**Figure 3.1:** NCEH1 constructs developed for expression in *Pichia pastoris*. Constructs developed for intracellular expression with pPICZ are labeled to the left, while constructs developed for secretion with pPICZ $\alpha$  are labeled to the right. Each of the constructs are grouped by truncation length and separated on the presence of a tev protease site that was engineered into the expression vectors. The pPICZ $\alpha$  vector contains an N-terminal  $\alpha$ -factor secretion sequence for the secretion of NCEH1 into the media during expression.

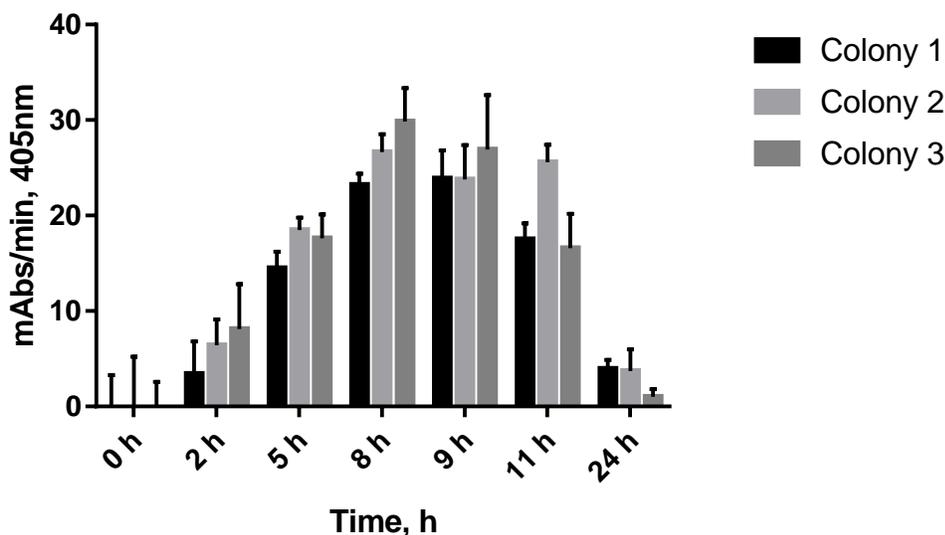
Insertion of a TEV protease site into both pPICZ and pPICZ $\alpha$  vectors was completed with the use of multiple rounds of PCR. The first round of PCR utilized

the KiaaRevTev1 reverse primer, while the second round utilized the PiczANot1Tev2 reverse primer. The protease site was successfully added before digestion and insertion into the expression vectors. Expression vectors containing each of the constructs listed within Figure 3.1 were transformed into *E. coli* DH5 $\alpha$  cells for propagation and sequencing. Once their sequence was confirmed, the constructs were linearized and inserted into one of three *Pichia pastoris* strains: X-33, GS115 or KM71H. The X-33 strain has a wild type genotype. GS115 has a mutation that prevents the synthesis of histidine, thus requiring GS115 to be supplemented with histidine in minimal media. KM71H has an insertion that disrupts the AOX1 gene, thereby disrupting the utilization of methanol, and creating a slower expression of the cloned gene of interest. Small scale expression trials on multiple colonies of each strain were conducted with both NCEH1 truncations 56-440 and 111-440 in all three *P. pastoris* strains. The cells were resuspended in 1 mL of 50 mM sodium phosphate buffer, pH 7.4, containing 300 mM NaCl. The cells were lysed by sonication with three 15 s bursts on maximum power using the microtip. The samples were micro-centrifuged at 13,000 rpm for 10 min and the resulting supernatants were analyzed for activity. The end result indicated that KM71H, the cell strain demonstrating the slowest rate of expression, appeared to yield the highest level of soluble NCEH1 protein.

After it was determined that all further experimentation would be completed with KM71H, optimization of the expression was required. The original protocol, suggested by the EasySelect *Pichia* Expression Kit from Invitrogen, included growing

the cells overnight at 30 °C in a buffered minimal glycerol medium [100 mM potassium phosphate, pH 6.0, 1.34% yeast nitrogen base, 2.5 μM Biotin, 1% glycerol]. The overnights were then used to inoculate 1 L of buffered minimum glycerol medium in a 7 L flask and grown at 30 °C for 24 h. The cells were then pelleted by centrifugation at 8,200 x g using 10 min spins before being resuspended in 200 mL of buffered minimal methanol medium [100 mM potassium phosphate, pH 6.0, 1.34% yeast nitrogen base, 2.5 μM Biotin, 0.5% methanol] in a 2 L flask to induce expression. The cells were incubated while shaking at 30 °C for 24 h. Samples were taken a few hours into the expression and activity was detected by PNPA activity assays. After 24 h, the cells were harvested by centrifugation at 8,200 x g for 10 min. In the case of pPICZ, the pellet was retained for NCEH1 purification, while the supernatant was retained for secreted NCEH1 from the pPICZ $\alpha$  expression. Interestingly, neither system provided noticeable activity levels after 24 h in direct contrast to the samples tested previously after only a few hours. A time-course expression was carried out to determine the optimal expression period of NCEH1 in *P. pastoris* as shown in Figure 3.2. Three separate KM71H colonies containing residues 111-440 of NCEH1 were selected for expression verification. Small aliquots of each expression were removed for analysis by PNPA and the assay was carried out in a 1 mL volume with 50 mM Tris buffer, pH 7.8, containing 10 mM PNPA. Approximately 50 μL of crude sample was added to start the assay. The reaction was monitored at 405 nm for 30 s.

### KM71H Time Course Expression Trial



**Figure 3.2:** Time course expression of *Pichia pastoris*. Experiment conducted with 3 separate colonies of 111-440 NCEH1 in KM71H *P. pastoris*. A small aliquot of the expression of each colony was removed at the time point indicated and evaluated for activity with a PNPA assay. Samples were normalized to their starting time point at 0 h. It was concluded that the optimal time period for expression was approximately 8 h at 30 °C.

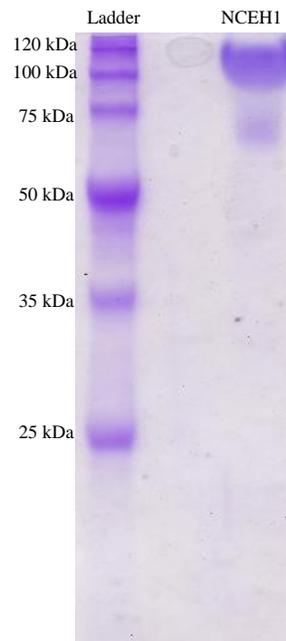
The results of the time course experiment indicated a loss of functional NCEH1 after approximately 8 h. This time period was confirmed in both pPICZ and pPICZ $\alpha$  expression systems. Therefore, further expressions were halted after 8 h of incubation with methanol.

#### 3.3.2 Purification Optimization of NCEH1 from pPICZ $\alpha$

Methodologies for the purification of NCEH1 were first evaluated from the expression constructs of pPICZ $\alpha$  because of the ability of the NCEH1 protein to be secreted from the cells. *Pichia pastoris* does not secrete a large number of native

proteins, and therefore the secretion of NCEH1 offered the opportunity of obtaining secreted protein with minimal efforts of purification. Expression trials were carried out on both 56-440 and 111-440 truncations of NCEH1 and no discernable difference was noted between the two. Truncation 111-440 was chosen for further experimentation because of the lack of the second hydrophobic region predicted to be associated to the membrane. It was predicted that removing this region may increase the ability of NCEH1 to crystallize, should a usable expression and purification system be achieved. Colonies of 111-440 NCEH1\_pPICZ $\alpha$  were expressed as previously described, and after 8 h of expression the cellular suspensions were centrifuged at 8,200 x g for 10 min and the supernatant was retained for experimentation. Ammonium sulfate precipitation of the protein from the resulting 1.2 L of expression was attempted, but due to the large volume and small protein concentration in solution, it was not successful. Therefore, a method to alleviate the problematic volume was required. Incremental incubation of the supernatant on a gravity flow column containing Ni-sepharose beads was considered, but these columns only hold 100 mL at a time and it would take a minimum of 12 h for sequential incubations. In the interim, significant protein could be lost. Andrew *et al.* evaluated the use of Aquacide II to concentrate protein solutions through dialysis tubing.<sup>4</sup> However, this method proved to be time consuming and inefficient for the volume of solution utilized. Concentration through centrifugation was also considered, but remained time consuming and not cost effective. In the end it was determined that passing the protein supernatant across an FPLC column was the best path forward. With a loading volume of approximately 5 mL/min all of the supernatant would be exposed to the FPLC column of choice within 4 h. Purification of NCEH1 by FPLC was achieved by the dialysis of the expression

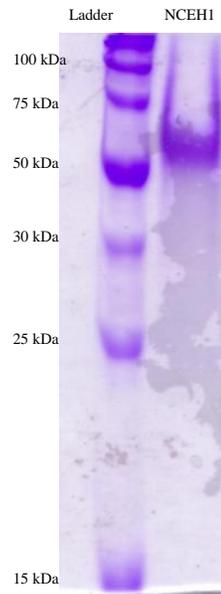
supernatant against 50 mM sodium phosphate buffer pH 7.4 containing 300 mM NaCl. The protein was eluted off of the column with an elution buffer, 50 mM sodium phosphate buffer pH 7.4 containing 300 mM NaCl and 300 mM imidazole resulting in purified protein shown in Figure 3.3.



**Figure 3.3:** SDS PAGE analysis of a purification of NCEH1 from pPICZ $\alpha$ . The expected molecular weight of 111-440 NCEH1 without glycosylation is approximately 35 kDa. The presence of 2 upper molecular weight bands at 100 and 75 kDa suggested the presence of hyper-glycosylation of NCEH1.

The purified protein sample, verified by activity analysis, contained two upper molecular weight bands at approximately 100 kDa and 75 kDa. The expected molecular weight of un-glycosylated 111-440 NCEH1 is approximately 35 kDa, thus suggesting the protein is hyper-glycosylating. In an effort to reduce the hyper-glycosylation, the temperature of the expression was reduced from 30 °C to 22 °C, and

the resulting bands reduced to 55 kDa as demonstrated in Figure 3.4. This observation added support to our hypothesis that secreted NCEH1 is hyper-glycosylated.

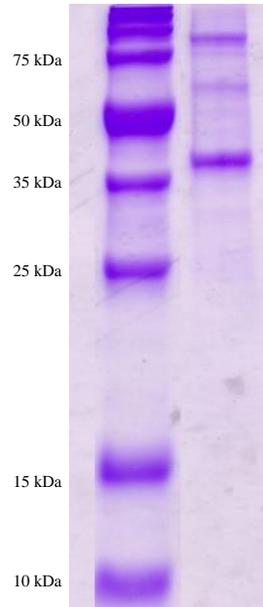


**Figure 3.4:** SDS analysis of a purification of 111-440 NCEH1 from pPICZ $\alpha$ . Expression was incubated at 22 °C. A single, although not homogeneous, band corresponding to the NCEH1 protein appears at approximately 55 kDa.

Although there was a reduction in hyper-glycosylation with expression at 22 °C, the reduction resulted in a protein band at approximately 55 kDa that was still considerably higher than the expected 35 kDa molecular weight of the non-glycosylated NCEH1 protein. The inability to decrease the amount of hyper-glycosylation further, without mutation, led to the exploration of a purification protocol for the NCEH1 pPICZ expression construct.

The pPICZ constructs expressed NCEH1 intracellularly. No difference was observed in the expression level of the truncations 56-440 versus 111-440 NCEH1 and

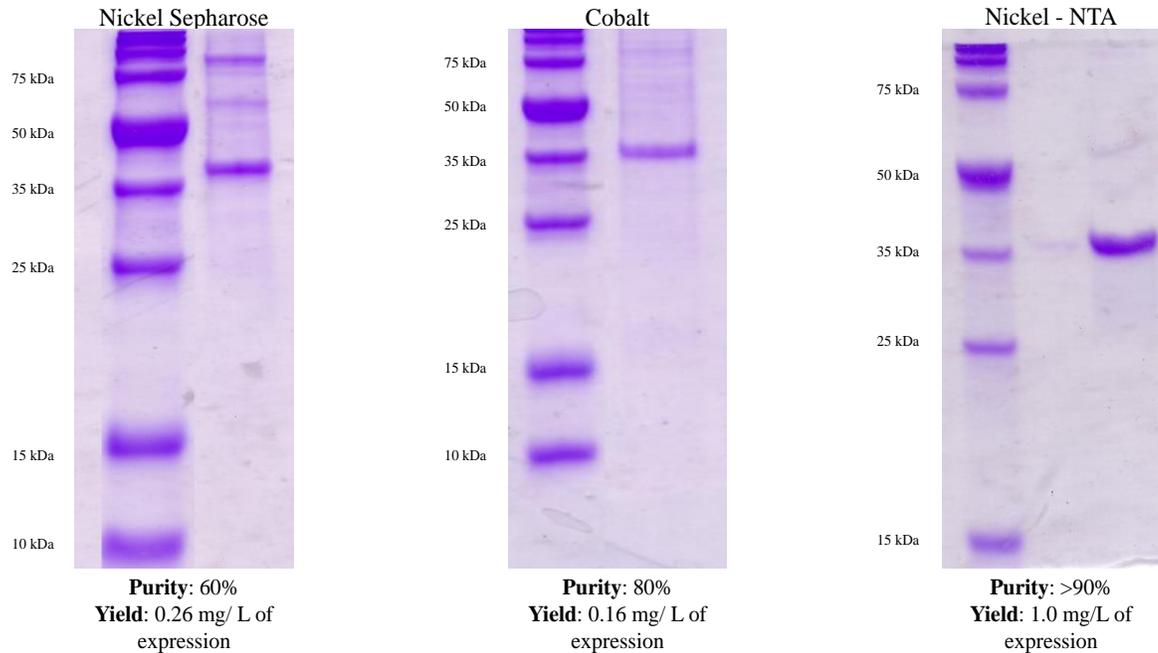
therefore, all further experimentation was carried out with truncation 111-440 with the anticipation that this construct would be preferable for structural characterization by X-ray crystallography. Cellular expression was halted after 8 h at 30 °C and the cells were pelleted by centrifugation at 8,200 x g for 10 min. The pellets were resuspended in 50 mM sodium phosphate buffer, pH 7.4, containing 300 mM NaCl and lysed by sonication. The solution was sonicated for three 1 min intervals with a rest on ice between each cycle. The sample had 1% tween 20 added and was rocked for 20 min at 4 °C before the membrane component was removed by 2 rounds of centrifugation at 23000 x g for 20 min each. The supernatant was originally incubated on Ni-sepharose beads for 1 h before being sequentially washed with a buffer containing 50 mM sodium phosphate, pH 7.4, containing 300 mM NaCl and a second wash buffer with 50 mM sodium phosphate pH 7.4 containing 300 mM NaCl and 100 mM imidazole. The protein was eluted with 3 consecutive washes with 10 mL of 50 mM sodium phosphate, pH 7.2, containing 300 mM NaCl, and 300 mM imidazole. The elutions were concentrated to 3 mL and dialyzed overnight into 10 mM HEPES buffer, pH 7.8, containing 10 mM NaCl. The resulting purified protein can be seen in Figure 3.5. Although a band was observed at the expected molecular weight of 35 kDa, the purified sample was estimated to be 60% pure by ImageJ analysis. The yield from the purified NCEH1 sample was 0.26 mg/L of expression.



**Figure 3.5:** SDS PAGE analysis of purified NCEH1 from the pPICZ expression vector. Purification completed with nickel sepharose resin (GE Healthcare) and resulted in a sample lacking in purity, approximately 60%, but a band present near the expected 35 kDa molecular weight.

The goal of this portion of the project was to obtain pure NCEH1 in a high yield for protein crystallization trials. Therefore, an increase in yield and purity was required to justify the use of the *Pichia* expression system. In an effort to increase the purity of the sample, a cobalt talon resin (Clontech) was utilized to replace the Ni-sepharose. The new cobalt resin increased the purity of NCEH1 to 80%; however, it reduced the yield to an average of 0.16 mg/L of expression. Ni-NTA provides a nitrotriacetic acid linker between the Ni<sup>2+</sup> and the resin, decreasing the leaching of Ni<sup>2+</sup> from the resin. Ni-NTA has been experimentally shown to increase yield and purity in comparison to the nickel sepharose.<sup>5</sup> The Ni-NTA resin was capable of

increasing the NCEH1 purity above 90% and still achieving a yield of approximately 1 mg/L of expression. The difference in purification resin can be visualized in Figure 3.6.



**Figure 3.6:** SDS PAGE analysis of the purification of NCEH1 using Ni-sepharose, cobalt, and Ni-NTA resins. The gel furthest to the left shows the purified protein lacking in purity with Ni-sepharose. Cobalt, shown in the middle, increased the purity of NCEH1, however decreases the yield in relationship to the Ni-sepharose column. The Ni-NTA resin, shown to the right, provided the highest purity and yields for the purification of NCEH1.

The purification of NCEH1 with Ni-NTA was the first successful purification to yield pure protein, and at a concentration viable for protein crystallization trials. Before screening for optimal crystal conditions could occur, the functionality of the

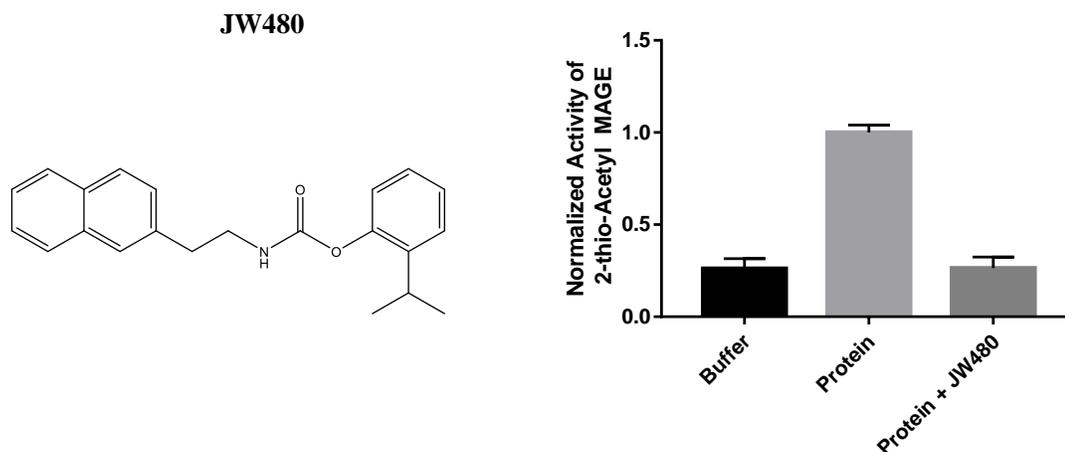
NCEH1 expressed from pPICZ was characterized. The presence of functional protein was verified using the 2-thio-acetyl MAGE assay before the sample was utilized for crystal screening.

### 3.3.3 Functional Characterization of NCEH1 from pPICZ

Igarashi *et al.* found that upon truncation of the N-terminus of NCEH1, the inability to glycosylate the protein led to a loss of activity. This information led to the prediction that glycosylation was required for NCEH1 activity.<sup>6</sup> In chapter 2 of this work, data was presented that also suggested that NCEH1 became unstable upon truncation of the N-terminus. Specifically, the protein was unable to be expressed when a histidine tag was added, but was well expressed when a larger chaperone protein fusion partner was added. *Pichia pastoris* is a eukaryote, and therefore offers intracellular machinery that is more appropriate for assisting in the folding of mammalian proteins, such as NCEH1. At this point, expression of NCEH1 in *Pichia* has been demonstrated with a simple histidine tag, showing a vast improvement in the folding environment for this protein. However, functionality of NCEH1 needed to be verified. The modified substrate 2-thio-acetyl MAGE, shown in Figure 2.3 in chapter 2, was utilized in an effort to confirm the functionality of NCEH1.

Due to the poor water solubility of the modified substrate, high enough concentrations could not be achieved to enable saturation Michaelis-Menten characterization. Therefore, an alternative strategy was explored, and activity was verified in concert with the use of the potent and selective inhibitor JW480 (shown in Figure 3.7) designed by Dr. Benjamin Cravatt.<sup>7</sup> Activity assays were conducted in 1 mL volumes of 75  $\mu$ M 2-thio-acetyl MAGE, 1 mM DTNB, and 50 mM sodium

phosphate pH 7.8. The assay was started by mixing approximately 0.1  $\mu$ M of protein into solution and monitoring the absorbance change at 412 nm for 60 s. The results demonstrated in Figure 3.7 show an increase in activity in the purified protein sample when compared to the assays run in the presence of only buffer or NCEH1 previously incubated with JW480.



**Figure 3.7:** The structure of the inhibitor JW480. The inhibitor was designed by Dr. Benjamin Cravatt for specificity against NCEH1<sup>7</sup>. The graph to right displays the normalized activity of the NCEH1 cleavage of 2-thio-acetyl MAGE in comparison to buffer and protein incubated with JW480.

The noted increase in activity of the purified protein sample indicated the presence of NCEH1. However, the decrease in activity to background levels of the protein sample incubated with JW480 confirmed that the hydrolysis of 2-thio-acetyl MAGE was in fact by NCEH1.

### 3.3.4 Crystallization Screening of NCEH1 from pPICZ

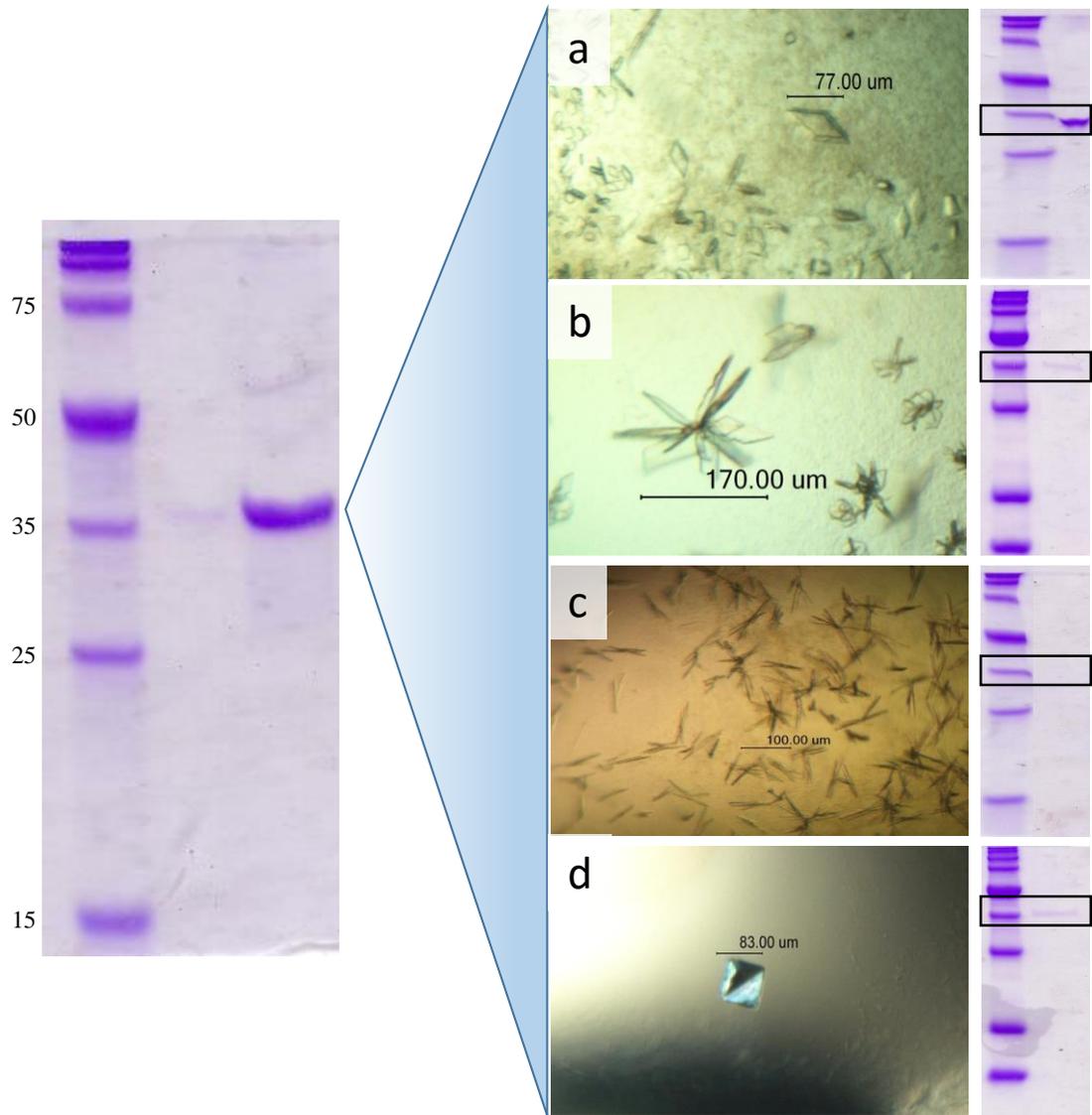
NCEH1 crystal trays were laid using the sample depicted in Figure 3.8. This sample had a concentration of 1.2 mg/mL, as measured using a BSA assay and a purity of greater than 95%, as determined by ImageJ. This sample was used to lay two crystal trays, screens 1 and 2. During these trials, multiple wells showed the presence of microcrystals after a 5-10 day incubation. After 4 weeks of monitoring, four conditions were selected that showed the most promise. These conditions are shown below in Table 3.4 and crystals are shown in Figure 3.8.

Hampton Research Screen	Well Number	Condition
1	45	.2M Zinc acetate dihydrate, .1M sodium cacodylate trihydrate pH 6.5, 18% PEG 8,000
1	46	.2M Calcium acetate trihydrate, .1M sodium cacodylate pH 6.5, 15% PEG 8,000
2	22	.2M Sodium acetate trihydrate, .1M Tris hydrochloride pH 8.5, 30% PEG 4,000
2	30	.2M Ammonium Sulfate, 30% PEG 8,000

**Table 3.4:** Crystallography solutions yielding microcrystals. Crystallography conditions arose from Hampton Research screens 1 and 2. The corresponding screen, 1 or 2, is listed in the first column of the table followed by the corresponding well solution number and chemical contents.

Figure 3.8 shows the original microcrystals obtained from the conditions listed above. Each of the conditions were replicated and after the formation of crystals, hanging drops were selected and carefully washed with the corresponding well solution. The wash removed un-crystallized protein from the sample, while leaving the precipitated crystals. The remaining crystals were selected and dissolved in water and 1x SDS PAGE loading dye. The samples were boiled for 10 min and loaded onto an SDS PAGE gel for analysis. Figure 3.8 shows the resulting SDS gels from each well

solution to the right of the corresponding crystal image. Each SDS gel shows the corresponding protein next to a broad range ladder and each of the conditions tested yielded the NCEH1 band with the exception of the calcium acetate hydrate solution, Figure 3.8 C. Density of the resulting SDS PAGE band was dependent on the number of crystals selected and dissolved. The decrease in the apparent concentration represents fewer dissolved microcrystals.



**Figure 3.8:** SDS PAGE gel representation of the NCEH1 sample used to lay hanging drop crystal screens 1 and 2. BSA analysis found the concentration of the NCEH1 to be approximately 1.2 mg/mL and purity to be greater than 95% as estimated by ImageJ. Several wells resulted in the formation of microcrystals as demonstrated in the figure. The crystals are ordered from top to bottom starting with the sodium acetate condition (a), the ammonium sulfate condition (b), the calcium acetate hydrate condition (c) and the zinc acetate dihydrate condition (d). Drops were carefully washed with well solution to remove excess un-crystallized protein before crystals were selected and resuspended for SDS PAGE analysis. With the exception of calcium acetate hydrate condition, each condition verified the presence of a protein at the expected 35 kDa band.

To optimize crystal formation, the sodium acetate (Figure 3.8a), ammonium sulfate (Figure 3.8b) and zinc acetate dihydrate (Figure 3.8d) solutions were prepared, filtered, and used as the well solution for an additive screen (Hampton Research). Crystal trays for each of the three conditions were laid, and the first forty-eight additives from the Hampton Research screen were sequentially pipetted into the wells, resulting in 48 unique solutions. In most cases the additive did not affect the formation of crystals; however, the additive screen for the sodium acetate solution (Figure 3.8a), with the addition of 2.5 mM adenosine 5 triphosphate disodium salt hydrate, yielded the crystal shown in Figure 3.9.



**Figure 3.9:** Crystal resulting from 0.2 M sodium acetate trihydrate, 0.1 Tris hydrochloride pH 8.5, 30% w/v PEG 4,000, 2.5 mM ATP.

Crystallography screens were halted when a similar molecular weight protein was isolated during a shortened purification of empty pPICZ vector, and work was started to determine its identity. The impurity was identified as mitochondrial alcohol dehydrogenase, mADH, with a molecular weight of 37 kDa. As demonstrated in Chen *et al.*, mADH can be removed from the Ni-NTA resin with imidazole concentrations of 50-100 mM.<sup>8</sup> The full purification protocol used to generate the crystals above utilized a 500 mL 100 mM imidazole wash of the Ni-NTA resin. This should be more than sufficient to remove mADH, and its sudden appearance is attributed to utilizing the shortened wash protocol without a 100 mM imidazole wash in the empty vector purification.

Future crystallography work will continue the screening for optimal crystallization conditions around the three verified crystallography conditions that yielded crystals. Variations in the pH, precipitant concentration, incubation temperature, and protein concentration can all be utilized in screening to further optimize crystallization conditions. The purified protein sample may also be subjected to second purification step utilizing an anion exchange FPLC column to guarantee the removal of mADH.

### **3.3.5 Glycosylation Mutagenesis of NCEH1 in pPICZ $\alpha$**

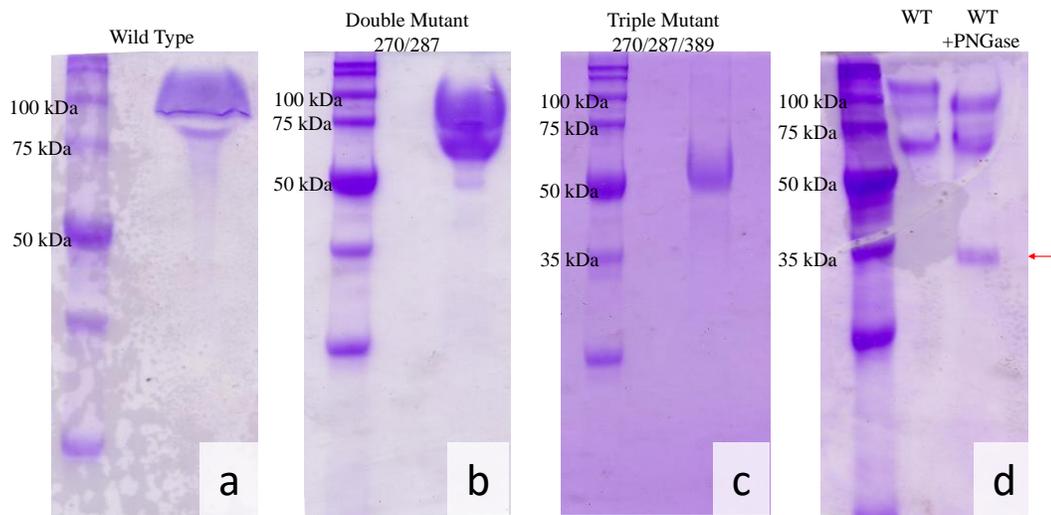
Previous experimentation with NCEH1 in the pPICZ $\alpha$  vector demonstrated the secretion of NCEH1; however, the protein appeared to be hyper-glycosylated. Experimentation with NCEH1 within the pPICZ vector showed the protein to be un-

glycosylated and demonstrated activity on its modified natural substrate 2-thio-acetyl MAGE. This observation suggested that the glycosylation sites (amino acids 270, 287 and 389) on NCEH1 purified from pPICZ $\alpha$  could be mutated and the protein would remain active NCEH1.

The three asparagine to glutamine mutations were introduced by PCR mutagenesis to the already verified NCEH1 construct ligated into the pPICZ $\alpha$  vector. The glutamine mutation was chosen to maintain relative size and charge at each of the amino acid locations, while preventing the actual glycosylation. The mutants were each transformed into DH5 $\alpha$  cells for DNA sequence confirmation before being inserted into the KM71H strain of *Pichia pastoris*. Previous experimentation, Figure 3.2, showed evidence of that the yeast cells became unstable after incubation in methanol for 8 h. Previously, it has been shown that *Pichia* cells cannot maintain themselves on methanol alone.<sup>1</sup> The expression medium containing 100 mM potassium phosphate, pH 6.0, 1.34% yeast nitrogen base, 2.5 $\mu$ M Biotin and 0.5% methanol was supplemented with 100 mM sorbitol allowing for sustained growth of *Pichia pastoris* without downregulating the AOX promoter, as would be seen with glycerol. This change in expression media allowed for the elongation of the expression protocol from 8 h to 48 h. The increase in expression translated to an increase in protein yield from 0.17 mg/L of expression to 1.0 mg/L of expression.

Initially the purification of NCEH1 from pPICZ $\alpha$  was completed by dialysis of the pPICZ $\alpha$  supernatant into 50 mM sodium phosphate buffer containing 300 mM NaCl at a pH of 7.4. The supernatant was passed through a Ni sepharose column for binding, washed with the sodium phosphate buffer and eluted with the same wash buffer but with the addition of 300 mM imidazole. The resulting purifications are

depicted in Figure 3.10. This figure shows a purification of the hyper-glycosylated wild type (a) as well as the corresponding decrease in glycosylation states that resulted from the elimination of glycosylation sites in the N270Q/N287Q double mutant (b) and N270Q/N287Q/N389Q triple mutant (c). Also shown in Figure 3.10 is the treatment of wild type protein with a de-glycosylase (PNGase), leading to the formation of 35 kDa band (d). The formation of the 35 kDa band indicates that the wild type NCEH1 is undergoing glycosylation as a posttranslational modification. Due to the high molecular weight observed for the triple mutant, the possibility of an unknown glycosylation site on the NCEH1 protein was explored. However, when the purified triple mutant N270Q/N287Q/N389Q was subjected to PNGase, no detectable change was observed indicating that either the additional glycosylation site was not accessible or that there was an uncontrolled and unintended modification to the protein.

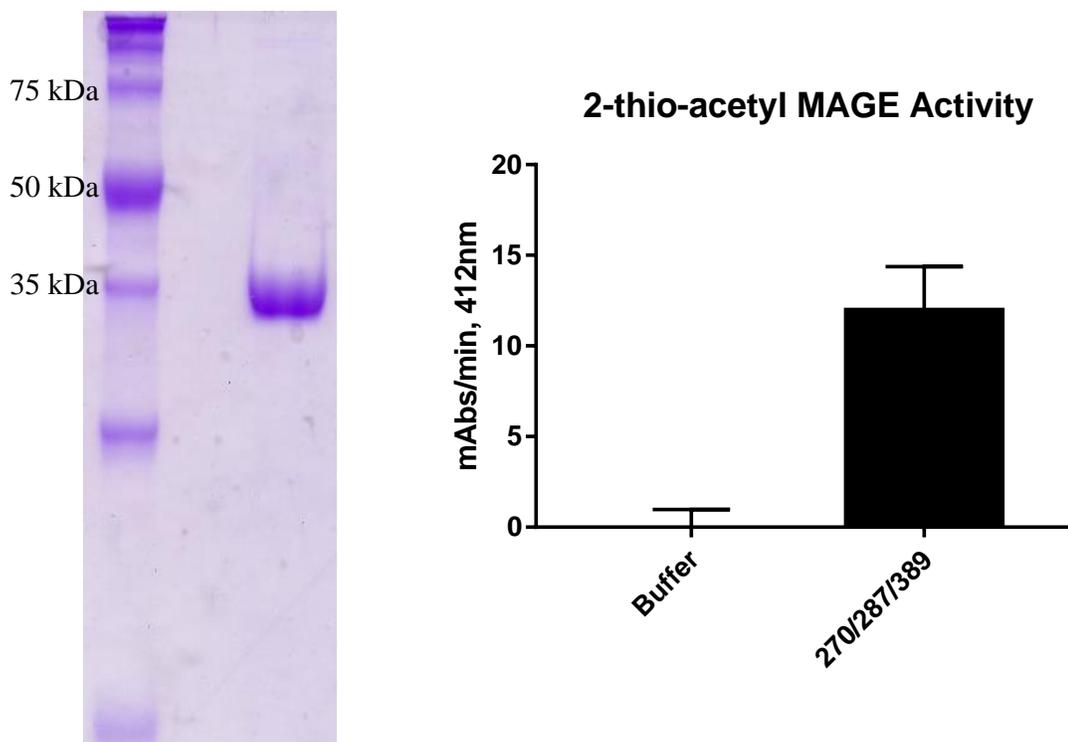


**Figure 3.10:** SDS PAGE analysis of the purifications of NCEH1 glycosylation mutagenesis from pPICZ $\alpha$ . This figure shows a purification of the hyper-glycosylated wild type (a) as well as the corresponding decrease in glycosylation states that resulted from the elimination of glycosylation sites in the N270Q/N287Q double mutant (b) and N270Q/N287Q/N389Q triple mutant (c). Also shown in Figure 3.10 is the treatment of wild type protein with a de-glycosylase (PNGase), leading to the formation of 35 kDa band (d). The formation of the 35 kDa band indicates that the wild type NCEH1 is undergoing glycosylation as a posttranslational modification. Although a clear decrease is apparent in the gels a, b, and c the final product of the triple mutant purification shown in panel c still yielded a protein at approximately 55 kDa while the predicted molecular weight remained at 35 kDa. Panel d demonstrates an incomplete deglycosylase reaction of the wild type NCEH1 sample. The formation of the appropriately sized 35 kDa NCEH1, marked by the red arrow, was observed proving the protein to be 35 kDa with the removal of glycosylation.

The existence of a band at 35 kDa for the wild type enzyme treated with PNGase strongly suggests all glycosylation sites are surface accessible, and the enzyme can be stably purified in the absence of glycosylation. The presence of a 55 kDa band for the triple mutant shown in Figure 3.10 does not have a satisfactory explanation, if the three glycosylation sites are at positions 270, 287 and 389, as one would expect to see a band at 35 kDa. To eliminate any anomalies of the triple mutant

that may have arisen due to a molecular biology issue, the triple mutant plasmid was reconstructed from the *E. coli* vector. The verified plasmid was grown, purified, and reinserted into the *Pichia pastoris* KM71H strain. The strain was grown in a small scale analysis, lysed, used as a template for PCR verification of the original DNA. The resulting DNA from the PCR was verified by sequencing to contain the mutations of interest.

The reconstructed triple mutant was expressed and because pPICZ $\alpha$  does not secrete many proteins an anion exchange FPLC column, HiTrap Q Sepharose (GE Healthcare) was chosen for purification over the previously used nickel chromatography. This modification of the original nickel purification was made to eliminate the possibility of nonspecific contamination. The supernatant was dialyzed into 50 mM Tris buffer pH 7.8 containing 150 mM NaCl and 1 mM EDTA. The sample was applied across the HiTrap column and washed with 50 mM Tris buffer, 1 mM EDTA, pH 7.8 and eluted with an NaCl gradient increasing to 1 M NaCl. Fractions were collected, and those containing protein were evaluated for activity by 2-thio-acetyl MAGE analysis. The fractions containing NCEH1 were pooled and concentrated. The purified NCEH1 was dialyzed overnight into 10 mM hepes, 10 mM NaCl, pH 7.4. The protein concentration for this purification was close to 1 mg/mL and can be visualized in Figure 3.11 along with its activity toward 2-thio-acetyl MAGE, in comparison to buffer.



**Figure 3.11:** SDS PAGE analysis of the pPICZ $\alpha$  expressed NCEH1 triple mutant. The triple mutant N270Q/N287Q/N389Q, at its expected 35 kDa molecular weight and its corresponding ability to cleave 2-thio-acetyl MAGE proving the presence of NCEH1 in the purified sample.

### 3.4 Conclusions

In the previous chapter, NCEH1 was shown to demonstrate significant instability. Due to this instability, the protein required a large chaperone protein, such as GST or MBP, to be fused to the N-terminus for optimal folding and expression. NCEH1 quickly destabilized if the fusion was removed resulting in the need for a more appropriate expression system. The yeast *P. pastoris* is a eukaryotic system that has been previously shown to increase the efficiency of mammalian protein expression relative to *E. coli*. The *Pichia* system also has the ability to either express the protein

of interest intracellularly (pPICZ), or to secrete the protein into the expression media (pPICZ $\alpha$ ). Expression of NCEH1 was carried out in both expression vectors, and after careful analysis, the KM71H strain was predicted to be the most appropriate expression strain for NCEH1. Initially, an expression time of 8 h was shown to be optimal for both intracellular and secreted expression methods. However, further expression experimentation showed that *Pichia pastoris* was unable to survive on methanol alone, and the expression media was regularly supplemented with sorbitol to provide an additional carbon source for growth without inhibiting the AOX operon. This increased the available expression time from 8 h to 48 h for the pPICZ $\alpha$  secreted expression, and importantly this significantly increased the yield of NCEH1 produced. Although this method has not been tested yet with the pPICZ expression, it is expected to have the same increased expression time and yield of NCEH1 produced.

The purification from the intracellular expression followed a very similar protocol to the approach previously discussed for *E. coli* expression. The pellet was resuspended, lysed, and centrifuged to remove the membrane component. The supernatant was applied to a Ni-sepharose column washed and eluted with buffers containing various amounts of imidazole. The purification yield of NCEH1 was improved from 0.26 mg/L of expression to 1.0 mg/L of expression by changing to the Ni-NTA affinity beads. Cobalt was also analyzed, and while it had moderately increased the purity of the sample it was at a significant cost in yield. The Ni-NTA purification was the clear winner as it increased both yield and purity of NCEH1 from pPICZ. This increase in yield allowed for purification of soluble NCEH1 to concentrations of greater than 1 mg/mL. The concentration of soluble NCEH1 in combination with purity of greater than 90% allowed for the screening of

crystallography conditions. Three main crystal conditions were verified to contain microcrystals of NCEH1 and further screening produced crystals slightly larger in size with addition of 25 mM ATP to the sodium acetate solution. An additional purification step with anion exchange chromatography should be completed to guarantee the removal of mADH before further screening of each of the three conditions around optimal pH, precipitant levels, temperature, and additives is still needed to continue the examination for optimal crystallization conditions of NCEH1.

NCEH1 was also expressed and purified from pPICZ $\alpha$ . In this expression protocol, the NCEH1 protein was secreted into the medium as an initial step to purification. However, the concentration of NCEH1 in the medium remained low enough to eliminate ammonium sulfate precipitation as an optimal method of purification from supernatant. Various methods to reduce the 1.2 L of supernatant were tested including concentration, a forced dialysis method with Aquacide II, and incubating the supernatant on a gravity flow affinity column. Each of these protocols failed to provide a viable method to retain a significant amount of the protein in a clean, timely and cost efficient manner. The best method consisted of dialyzing the supernatant against a loading buffer for an FPLC column, loading the supernatant across a nickel or anion exchange column, and testing eluted protein fractions for activity. After analysis of the eluted protein, it was discovered that the pPICZ $\alpha$  expression hyper-glycosylated the NCEH1 protein. In an effort to reduce the hyper-glycosylation three asparagine to glutamine mutations were made at amino acids 270, 287, and 389. Functional protein at the expected molecular weight of 35 kDa was

isolated with the use of an anion exchange column with a similar concentration and purity to the intracellular expression. These results represent the first time NCEH1 has been expressed and purified outside of a mammalian expression system. Not only was an increase in purity achieved from *P. pastoris* expression, but it was achieved without a large protein fusion partner on the N-terminus of NCEH1. The stability of NCEH1 from *P. pastoris* indicates that NCEH1 can maintain stability without the first 111 amino acids of the N-terminus of the protein. In addition, the three asparagine to glutamine mutations removed the posttranslational modifications of NCEH1 without effecting the activity of NCEH1, as seen with the turnover of the modified natural substrate 2-thio-acetyl MAGE. This confirms that glycosylation is not required for catalytic ability, and suggests more of a structural role for glycosylation in mammalian cells.

Although this expression and purification protocol has resulted in a large improvement in the ability to isolate purified NCEH1, it can still be improved. The intracellular expression has not been completed in the presence of sorbitol, and neither expression has been completed in a complex media. Both of these expression modifications have the potential to significantly increase the soluble yield of NCEH1 from *Pichia pastoris*. The increase in yield and purity of NCEH1 from the *Pichia pastoris* system will allow an in depth look at the functional and structural components of NCEH1 for the first time.

## REFERENCES

- [1] Daly, R., and Hearn, M. T. W. (2005) Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production, *Journal of Molecular Recognition* 18, 119-138.
- [2] Lueking, A., Holz, C., Gotthold, C., Lehrach, H., and Cahill, D. (2000) A system for dual protein expression in *Pichia pastoris* and *Escherichia coli*, *Protein Expression and Purification* 20, 372-378.
- [3] Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis, *Nature Methods* 9, 671-675.
- [4] Andrew, S. M., Titus, J. A., and Zumstein, L. (2002) Dialysis and concentration of protein solutions, *Current protocols in toxicology / editorial board, Mahin D. Maines (editor-in-chief) ... [et al.] Appendix 3, A.3H.1-5*.
- [5] Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J., and Schaefer, F. (2009) Immobilized-metal affinity chromatography (IMAC): A review, *Guide to Protein Purification, Second Edition* 463, 439-473.
- [6] Igarashi, M., Osuga, J.-i., Isshiki, M., Sekiya, M., Okazaki, H., Takase, S., Takanashi, M., Ohta, K., Kumagai, M., Nishi, M., Fujita, T., Nagai, R., Kadowaki, T., and Ishibashi, S. (2010) Targeting of neutral cholesterol ester hydrolase to the endoplasmic reticulum via its N-terminal sequence, *Journal of Lipid Research* 51, 274-285.
- [7] Chang, J. W., Nomura, D. K., and Cravatt, B. F. (2011) A Potent and Selective Inhibitor of KIAA1363/AADACL1 that Impairs Prostate Cancer Pathogenesis, *Chemistry & Biology* 18, 476-484.
- [8] Chen, Y., Li, Y., Liu, P., Sun, Q., and Liu, Z. (2014) Optimized expression in *Pichia pastoris* eliminates common protein contaminants from subsequent His-tag purification, *Biotechnology Letters* 36, 711-718.

## Chapter 4

### CONCLUSIONS AND FUTURE DIRECTIONS

#### 4.1 NCEH1 Expression and Purification from *E. Coli*

The truncation of the membrane protein's transmembrane helix has been shown to decrease the stability of NCEH1.<sup>1</sup> The work reported in Chapter 2, utilizing the recombinant *E. coli* system, supported the need for a chaperone protein fusion partner. The inability to express the N-terminally truncated enzyme alone, or with only a small histidine fusion tag lead to the discovery that the enzyme required a larger fusion protein such as GST or MBP to assist in the folding of NCEH1. Successful expression of NCEH1 was achieved with both GST and MBP fusion tags. However, these larger fusion partner constructs were still expressed almost exclusively into the insoluble fraction. In an effort to increase the solubility of NCEH1, N-terminal truncations were made to remove predicted hydrophobic regions. The NCEH1 expression construct obtained from Dr. Benjamin Cravatt had already removed the first 55 amino acids corresponding to the N-terminal transmembrane helix. After expression experiments showed that the protein remained in the insoluble fraction a second hydrophobic domain was removed creating the NCEH1 truncation 111-440. No significant improvement in solubility was observed. In a last attempt to improve solubility through truncation a construct of amino acids 165-440 was created to mimic the third human isoform of NCEH1. This truncation was not successfully expressed with either the GST or MBP fusion proteins, and has not been attempted in *Pichia pastoris*. The lack of solubility led to the requirement of purifying large volumes of expressed NCEH1 in an effort to recover a large enough quantity of enzyme for functional and structural studies.

Purification of NCEH1 from the *E. coli* recombinant system also provided evidence of the protein's instability as the protein was unable to be cleaved successfully from its fusion tag. Thrombin and Factor Xa, were both incapable of complete cleavage of the MBP-NCEH1 fusion proteins. The NCEH1 that was cleaved by either protease was rapidly destabilized and degraded. Due to the limitations in protein cleavage, functional and structural characterization of the MBP-NCEH1 fusion protein was pursued without the cleavage of the MBP fusion tag. Modifications to the expression and purification protocol increased the yield of recovered protein from 0.08 mg/L of expression to 0.15 mg/L of expression.

Functional characterization was completed by monitoring the hydrolysis of para-nitrophenyl acetate (PNPA) and the modified natural substrate 2-thio-acetyl MAGE. The poor solubility of the natural substrate precluded Michaelis-Menten kinetics; however, the use of this substrate confirmed the activity was attributed to NCEH1. As a substitute for 2-thio-acetyl MAGE, PNPA was used for kinetic studies and a  $K_M$  and  $k_{cat}$  were calculated to be 5.3 mM and 4 s<sup>-1</sup> respectively.

Samples demonstrating soluble fusion protein concentrations greater than 0.5 mg/mL, and with a purity of greater than 90%, were utilized for crystallization trials. Hanging drop preparations utilizing Hampton Research screens 1 and 2 were laid and incubated at room temperature. Microcrystals were achieved; however, the protein concentration remained low and the size and quality of crystals grown could not be improved.

Previous work by Nomura *et al.* suggested directed mutagenesis could improve the solubility of NCEH1.<sup>2</sup> Due to the lack of a verified structure for the enzyme, this random mutagenesis was used instead in an effort to increase the soluble fraction of

the enzyme. Random mutagenesis was accomplished with the addition of modified dNTPS, 8-oxo-2'-deoxyguanosine-5'-triphosphate and 2'-deoxy-P-nucleoside-5'-triphosphate, to a PCR protocol containing the DNA template 56-440 NCEH1 with a predicted mutation rate of 6-19%. Mutagenic DNA was transformed directly into BL21-CodonPlus(DE3)-RIPL cells, colonies were screened for activity, and three colonies were selected for large scale expression and purification. The average yield of recovered NCEH1 was 0.47 mg/L of expression, an improvement over the previous recovery of 0.15 mg/L of expression.

#### **4.2 *Pichia Pastoris***

The instability of NCEH1 from *E. coli* led to the decision to switch expression systems and attempt to obtain purified enzyme from *Pichia pastoris*, a eukaryotic system better equipped for the folding mammalian proteins. Both the 56-440 and 111-440 NCEH1 truncations were sub-cloned into two expression vectors, pPICZ and pPICZ $\alpha$ , selected for the ability to intracellularly express or secrete NCEH1, respectively. Both vectors contained a short C-terminal histidine tag and lacked a large chaperone on the N-terminus of the protein. NCEH1 was successfully expressed in both vectors. In contrast to NCEH1 expressed from *E. coli* and regardless of the lack of an N-terminal chaperone, the successful expression of NCEH1 from *Pichia* strongly suggested that the N-terminus was not required for stability of the protein and that the protein benefited from the *Pichia* folding mechanisms.

The expression of NCEH1 demonstrated a time dependent curve that maximized protein production at 8 h following methanol addition. This was caused by the inability of the KM71H strain to completely utilize methanol as a carbon source. The addition of sorbitol to the expression medium elongated the expression time of

pPICZ $\alpha$ , increasing the available NCEH1. This method has not been tested for pPICZ, but is predicted to have the same elongating effect.

Purification of NCEH1 from *P. pastoris* has resulted in a protein concentration of 1 mg/L of expression for both expression vectors. The protein was estimated to be 90% pure by ImageJ, resulting in the first time NCEH1 has been successfully purified with the lack of a chaperone protein in concentrations and purity high enough for crystallization trials. Hanging drop preparations of crystal screens 1 and 2 were laid using protein samples with concentrations greater than 1 mg/mL resulting in several conditions yielding microcrystals. Microcrystals from four separate conditions were re-solubilized and analyzed via SDS PAGE. Three of the four showed an apparent band at the predicted 35 kDa providing evidence to pursue crystallization screening around each of the three conditions.

A disadvantage to the pPICZ expression vector was the expression of the endogenous *Pichia* protein mitochondrial alcohol dehydrogenase or mADH. This protein was found to have a very similar molecular weight to NCEH1 at 37 kDa. Research has shown that this enzyme elutes from a Nickel column at imidazole concentrations of 50-100 mM while NCEH1 has been shown to elute above 100 mM.<sup>3</sup> The difference in elution characteristics make it unlikely that mADH is co-purifying with NCEH1. Although not demonstrated in this work, mADH could also be removed with anion exchange chromatography as the proteins do not have the same pI. Fortunately, the co-expression of mADH can be circumvented using pPICZ $\alpha$ , which will excrete NCEH1 into the media. Unfortunately, the secretion of NCEH1 led to a hyper-glycosylated state. The three predicted glycosylated asparagines located at amino acid 270, 287, and 389 were successfully mutated to glutamines, eliminating

the hyper-glycosylated protein. Although abnormalities arose with this protocol, modifications to the purification scheme resulted in the recovery of a functional 35 kDa band of NCEH1 as previously predicted without glycosylation.

The demonstration of activity of NCEH1 expressed from both *E. coli*, which is incapable of glycosylation, and *P. pastoris* with all three glycosylation sites mutated demonstrated functional protein can be prepared without glycosylation. These results contradict previous notions that glycosylation is required for catalytic activity<sup>1</sup>. Current evidence appears to signify glycosylation plays a role in stabilizing NCEH1 during folding. This was demonstrated by the inability of NCEH1 to maintain stability without a chaperone protein in *E. coli*, but maintain stability in *P. pastoris* without a chaperone.

### **4.3 Future Directions**

#### **4.3.1 *Escherichia coli* Expression**

Due to the instability of NCEH1 from the *E. coli* expression system, a large chaperone protein was required to remain fused to the enzyme. The ability to isolate higher concentrations of NCEH1, without a chaperone, from *Pichia pastoris* eliminates the need for future work to improve the recovery of NCEH1 from *E. coli*. However, further experimentation involving the random mutagenesis of NCEH1 is still best completed with the *E. coli* system to enhance our understanding of therapeutic approaches to NCEH1.

Random mutagenesis within this work has already shown colonies with an apparent increase in solubility, however the sequences from the random mutagenic strains have not been isolated. Thus, it is unclear if the apparent increase in solubility

is due to mutation or another cellular variation. The sequencing of these variants is the next step to answer the question of whether mutation or unknown variation of expression is the reason for the increase in recovered NCEH1. The difficulty of sequencing from BL21 cells requires these plasmids to be retransformed into DH5 $\alpha$  cells for propagation and subsequent sequencing analysis. Understanding the mutations present could improve the solubility of NCEH1, and if needed, could then be incorporated into the sequences of the *P. pastoris* expression system. This improvement could enhance the yield of NCEH1, creating better opportunities for crystallization, as well as our ability to perform the directed mutagenesis required to increase in solubility.

An additional direction for mutagenic NCEH1 is to increase the functional ability to cleave organophosphorus compounds. As discussed in appendix A of this work, random mutagenesis of platelet-activating factor acetylhydrolase (PAF-AH), a similar process has already been initiated with PAF-AH. A similarly designed protocol could also be carried out with NCEH1 in an effort toward improved organophosphate hydrolase activity. This protein engineering would be beneficial for the catalytic bioscavenging of nerve agents and pesticides used within the environment.

#### **4.3.2 *Pichia Pastoris***

The ability to purify NCEH1 in concentrations and purity high enough for crystallization trials is a critical first step for the elucidation of the structure of the enzyme. Although this work reports concentrations of soluble protein above 1 mg/mL, this is still on the low end for optimal crystallization screening. Future work to increase the concentration of soluble NCEH1 retrieved for the *P. pastoris* expression system includes the following paths. As previously mentioned, the incorporation of

mutations from the sequencing of the *E. coli* random mutants may increase the soluble fraction of NCEH1. Solubility may also be increased by switching the growth and expression media from minimal to complex by including yeast extract and peptone for the growth and stability. In addition, the NCEH1 truncation 165-440, created to replicate human isoform c, was never constructed for *Pichia* expression. The lack of expression in *E. coli* does not necessarily preclude stability within the *Pichia* expression system, as noted by the ability of *P. pastoris* to express stable NCEH1 lacking an N-terminal fusion protein. The 165-440 construct may prove critical in the ability to achieve high concentrations of soluble enzyme with an increase in homogeneity critical for the crystallization of NCEH1.

Initial crystallization trials indicated that there were three possible crystallography conditions to optimize for further screening. Future work around the crystallization of NCEH1 will include the screening of conditions previously mentioned and will be performed on highly pure samples, such as shown in Figure 3.11. Variables such as pH, temperature, enzyme concentration and precipitant concentrations will be varied, and oil overlays, and new additives will be evaluated. Crystallization screening with the inhibitors bound, such as the Cravatt ABPP inhibitor JW480 should also be explored.

Functional characterization of the intrinsic kinetic parameters of  $K_M$ ,  $V_{max}$ , and  $k_{cat}$ , have also only been explored with the fusion MBP\_NCEH1 construct from the *E. coli* expression system. However, the instability of the enzyme upon cleavage of the fusion MBP would suggest an imperfection in the folding of NCEH1. The NCEH1 enzyme expressed from *Pichia pastoris* does not require a large fusion chaperone for stability, and therefore this method of expression most likely produces a more

accurately folded enzyme. Functional characterization of the intrinsic kinetic parameters should be further studied on the enzyme resulting from the *Pichia* system.

At this time purified NCEH1 has also never been demonstrated to catalytically cleave cholesterol esters. A coupled catalytic cholesterol esterase assay should be performed with the purified enzyme to further study its role as a neutral cholesterol esterase.

### **4.3.3 Other Expression Systems**

As previously mentioned in Chapter 1 of this work, NCEH1 is predicted to have three human isoforms; however, extensive research has not been carried out to identify the importance of these isoforms or the role they play in different cell types. Further experimentation is required to identify the isoforms of NCEH1 associated with various cellular types, such as macrophages and platelets, as well as disease states involving atherosclerosis and cancer. Experimentation should also focus on the effect of the different isoforms on natural substrates.

The isolation of purified and functional NCEH1 from both *E. coli* and *P. pastoris* strongly suggests that glycosylation plays a role in the stability of NCEH1 while folding; although, the *Pichia* system lacks the ability to glycosylate the protein identically as the mammalian system would. The actual physiological role of glycosylation cannot be adequately studied in either *E. coli* or *Pichia*, and therefore requires an in-depth analysis from a mammalian system. Igarashi *et al.* demonstrated an importance of glycosylation pattern of the mouse form of NCEH1<sup>1</sup>. However, glycosylation is only predicted to share two similar sites between the mouse and human form. In addition to this, the human form is only predicted to contain the three glycosylation sites. However, only two molecular weights were visualized at 45 kDa

and 50 kDa in mammalian cells<sup>1,2,4</sup>. Upon treatment with a deglycosylase the molecular weight was visualized at 40 kDa<sup>4</sup>. The only plausible experiment to study the role of glycosylation would require the expression and functional characterization of human NCEH1 within a mammalian expression system. It could be beneficial to the therapeutic approaches to NCEH1 to confirm the three glycosylation sites, as well as any pattern these glycosylation sites may play within the various cell types containing NCEH1. Investigation into the glycosylation pattern involving both healthy and cancerous cell lines could improve our knowledge of the functional activity of NCEH1 pertaining to aggressive cancers and atherosclerosis. This information could be coupled to testing constructs of various glycosylation patterns, in their physiological conditions, to isolate any correlation between their differences and their functional ability pertaining to the endogenous substrates 2-acetyl MAGE, cholesterol esters, and organophosphates.

## REFERENCES

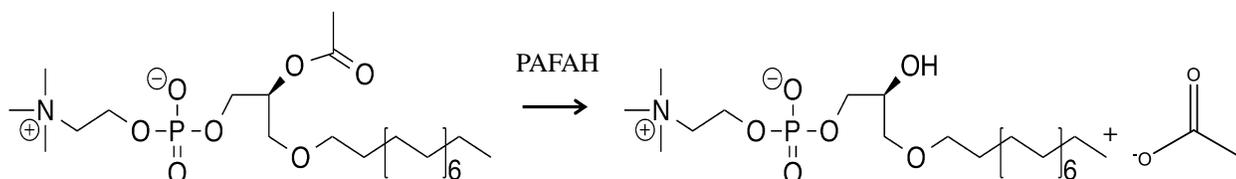
- [1] Igarashi, M., Osuga, J.-i., Isshiki, M., Sekiya, M., Okazaki, H., Takase, S., Takanashi, M., Ohta, K., Kumagai, M., Nishi, M., Fujita, T., Nagai, R., Kadowaki, T., and Ishibashi, S. (2010) Targeting of neutral cholesterol ester hydrolase to the endoplasmic reticulum via its N-terminal sequence, *Journal of Lipid Research* 51, 274-285.
- [2] Nomura, D. K., Fujioka, K., Issa, R. S., Ward, A. M., Cravatt, B. F., and Casida, J. E. (2008) Dual roles of brain serine hydrolase KIAAL1363 in ether lipid metabolism and organophosphate detoxification, *Toxicology and Applied Pharmacology* 228, 42-48.
- [3] Chen, Y., Li, Y., Liu, P., Sun, Q., and Liu, Z. (2014) Optimized expression in *Pichia pastoris* eliminates common protein contaminants from subsequent His-tag purification, *Biotechnology Letters* 36, 711-718.
- [4] Jessani, N., Liu, Y. S., Humphrey, M., and Cravatt, B. F. (2002) Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness, *Proceedings of the National Academy of Sciences of the United States of America* 99, 10335-10340.

## Appendix A

### RANDOM MUTAGENESIS OF PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE

#### A.1 Introduction

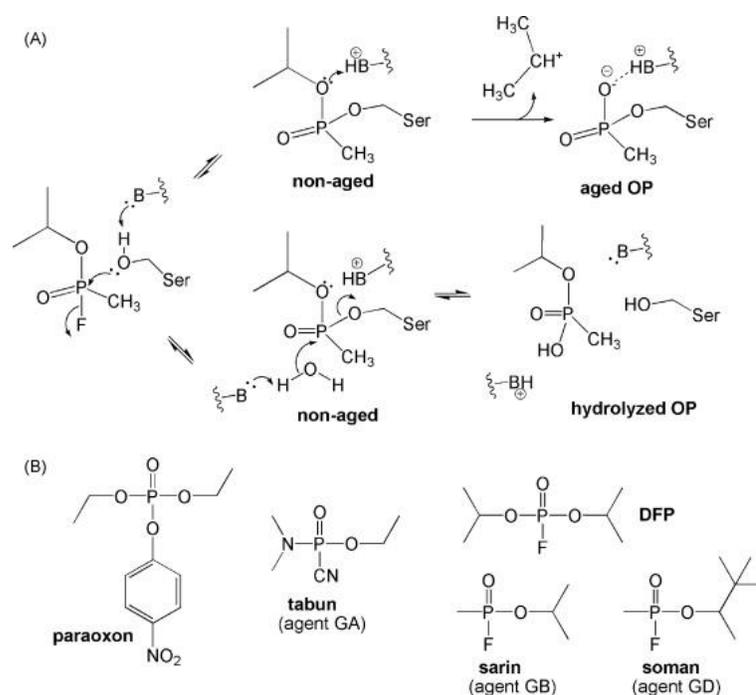
As discussed in the introduction of this work platelet-activating factor (PAF) is a potent lipid mediator that causes inflammation and atherosclerosis.<sup>1-4</sup> The regulation of PAF is a part of numerous biological pathways including the inflammatory response, anaphylaxis, hypotension, and smooth muscle contraction.<sup>3</sup> In addition to the hydrolysis 2-acetyl MAGE by NCEH1, PAF is predominantly regulated by platelet-activating factor acetylhydrolase, (PAFAH) also via hydrolysis. This reaction is shown in Figure A.1.



**Figure A.1:** The hydrolysis of PAF by PAFAH. The reaction forms a biologically inactive lyso-PAF and acetate ion.

PAFAH is a serine hydrolase exhibiting the classic  $\alpha/\beta$ -hydrolase fold whose structure was previously reported by Samanta *et al.* in 2008.<sup>5</sup> Although significant research on the PAFAH enzyme has involved elucidating its role in atherosclerosis, it also has implications in the hydrolysis of organophosphate (OP) compounds such as paraoxon, sarin, tabun, and soman shown in Figure A.2. This figure, originally printed

in Samanta *et al.* in 2009, demonstrates the process of OP hydrolysis and its complications with PFAFH.<sup>6</sup> As can be seen in Figure A.2, the organophosphate compound binds to the active site of the serine hydrolase, progressing the enzyme first to a “non-aged” complex. This complex is an inhibited form PFAFH, but theoretically could be hydrolyzed with the addition of a catalytic base to activate a water molecule.



**Figure A.2:** Serine hydrolase reaction with organophosphates. A) The interactions of organophosphate compounds with serine hydrolases. In enzymes such as acetylcholinesterase, the OP compound binds the active site of the serine hydrolase inactivating the enzyme causing the formation of an “aged” complex. The addition of an optimally placed basic compound can participate in the turnover of the OP preventing the “aged” complex from forming. B) Example of toxic OP compound structures paraoxon, tabun, sarin, and soman.

Reprinted from “Crystal structures of human group-VIIA phospholipase A2 inhibited by organophosphorus nerve agents exhibit non-aged complexes”, (2009) *Biochemical Pharmacology* 78, 420-429., with permission from Elsevier.

Wild type PAFAH lacks the basic residue required for the dephosphorylation of the active site serine, deactivating it in the presence of OP compounds. However, the crystallization of PAFAH in the presence of sarin, soman, and tabun led to directed mutagenesis of the tryptophan residue at position 298 to a histidine residue.<sup>7</sup> Kirby *et al.* showed that this mutation was able to successfully dephosphorylate the catalytic serine residue in the active site of PAFAH. This was a critical first step to the analysis of PAFAH as a relevant catalytic bioscavenger against potentially life threatening OP compounds. However, the novel somanase activity  $k_{cat}$  was only reported as  $5 \text{ min}^{-1}$ .<sup>7</sup> This novel somanase activity led to the belief that further directed, or random, mutagenesis was a possible means to increase the catalytic efficiency of PAFAH against OP compounds. However, the screening of mutagenic DNA could be time consuming and inefficient.

Cho *et al.* demonstrated the successful screening of randomly mutated DNA through cell surface display of an organophosphorus hydrolase (OPH) in *E. coli*.<sup>8</sup> This method provided the next step in the consideration of other mutations that could increase the catalytic activity of PAFAH against OP compounds. The following work presented in this appendix will discuss the current state of the mutagenesis of PAFAH, and its ability to cleave the OP compound methyl paraoxon.

## **A.2 Materials and Methods**

### **A.2.1 Random Mutagenesis**

The PAFAH DNA template spanning amino acids 49-423, and already containing the W298H mutation, was obtained from Emily Hartzell and utilized for random mutagenesis. Random mutagenesis was introduced into the PAFAH DNA sequence with the use of 10 mM mutagenic dNTPS 8-oxo-2'-deoxyguanosine-5'-triphosphate and 2'-deoxy-P-nucleoside-5'-triphosphate (trilink biochemical), and a low fidelity taq polymerase (Fisher scientific). The DNA was amplified with the forward primer 5'-  
GATGATGGATCCCGGCAGCTTCGGCCAAACTAAAATCCCCCG -3' and reverse primer 5'-GATGATAAGCTTTCAGTTAATGTTGGTCCCTGGAATAAG -3' for 20 cycles with fresh DNA template added after the 10<sup>th</sup> cycle. The PCR reaction was run in duplicate in an effort to maximize the mutagenic library and a mutation rate of 6-19% was expected for this reaction.<sup>9</sup> The amplified DNA was PCR purified with a Qiagen purification kit and eluted in ddH<sub>2</sub>O. This sample was stored as the library of mutagenic DNA for further PCR reactions.

### **A.2.2 Subcloning of PAFAH into the pINCOP Vector.**

The pINCOP plasmid originally described in Shimazu *et al.*<sup>8, 10</sup> utilized throughout this portion of the experiment was obtained from Dr. Wilfred Chen and contained a truncated ice nucleation protein (INP). INP created an anchor used to display PAFAH on the surface of the cell. Cell surface display of PAFAH was achieved previously by Emily Hartzell by fusing the enzyme onto INP within the pINCOP vector. PAFAH was confirmed to be displayed by Emily in activity assays showing the cells ability to hydrolyze a modified natural substrate of PAFAH 2-thio-

PAF and was utilized as the template for a second PCR reaction, lacking the mutagenic dNTPs and running for 30 cycles for maximum amplification. The pINCOP vector and amplified DNA was digested with BamHI and HindIII for 2 h at 37 °C. The digested samples were gel purified with a Qiagen extraction kit and ligated overnight at 14 °C with T4 Ligase (New England Biolabs).

Ligated DNA was subcloned into BL21(DE3) cells (New England Biolabs) by a heat shock method. DNA was incubated with BL21(DE3) cells on ice for 5 min, heat shocked at 42 °C for 45 s, and placed back on ice for 2 min. The samples were incubated, shaking, at 37 °C for 1 h in 1 mL of LB. The mixture was then added to 5 mL of LB containing 100 µg/mL of ampicillin for growth overnight. This sample contained a mixture of PAFAH mutants and was stored in aliquots at -80 °C.

### **A.2.3 Screening of Mutagenic PAFAH**

BL21 cells containing the mutated PAFAH DNA were plated on numerous LB agar plates supplemented with 100 µg/mL of ampicillin. The BL21 cell were also plated on M9 agar plates supplemented with 0.1% tryptone, 0.05% yeast extract, 0.1% casamino acids, 0.2% glucose, 100 µg/mL ampicillin, and 10 µM CoCl<sub>2</sub>. Plates were grown at 37 °C for 16 h before a top agar containing 0.7% agar, 50 mM phosphate-citrate buffer pH 8.0, and 1 µM methyl paraoxon (Sigma-Aldrich). Colonies developing a yellow color were selected and grown up overnight at 37 °C in 5 mL LB containing 100 µg/mL ampicillin. Overnight samples containing wild type PAFAH and W298H PAFAH constructs, which were ligated into the pINCOP vector as controls.

Small aliquots of each overnight were stored at -80 °C for future experimentation. Overnights were then centrifuged at 4000 rpm for 10 min and

resuspended in 2 mL of 50 mM phosphate-citrate buffer pH 8.0. After resuspension an additional volume of 50 mM phosphate-citrate buffer containing methyl paraoxon was immediately added to each tube to a final reaction volume of 2.5 mL containing 1  $\mu$ M methyl paraoxon. The tubes were mixed and an initial spectrophotometer absorbance at 405 nm was measured. The reaction proceeded for 10 min before a second spectrophotometric reading was taken. Colonies expressing activity higher than the OP hydrolase mutant, W298H, were saved for future experimentation.

#### **A.2.4 Subcloning of Mutagenic DNA into pGEX-6P1**

Colonies selected during the screening of mutagenic PAFAH for an increase in activity toward methyl paraoxon were grown up overnight 5 mL of LB containing 100  $\mu$ g/mL ampicillin. The pINCOP plasmid containing the mutagenic PAFAH was purified using a Qiagen extraction kit and used as template for PCR amplification for the subcloning of the mutagenic DNA into pGEX-6P-1 expression vector. The DNA was amplified with the forward primer 5' – AAAAAAGGATCCTCCTTTGGCCAAACT – 3' and the reverse primer 5' – AAAAAACAGCTGGTTAATGTTGGTCCCT – 3'. The underlined segments represent the corresponding BamHI and SalI restriction sites, respectively. The amplified DNA was PCR purified by a Qiagen kit, and eluted with ddH<sub>2</sub>O. The purified DNA and empty pGEX-6P-1 vector were digested with BamHI and SalI for 1 h at 37 °C before the addition of shrimp alkaline phosphatase (rSAP) obtained from New England Biolabs. Following the addition, the digestion was continued at 37 °C for 2 h before the entire sample was gel purified by a Qiagen extraction kit. The purified, digested PAFAH DNA was mixed with electrocompetent DH5 $\alpha$  cells and electroporated with 2.5 kV. The cells were mixed with 1 mL of LB and incubated

shaking at 37 °C for 1 h. The various quantities of the transformation were plated on LB agar containing 100 µg/mL ampicillin.

### **A.3 Results and Discussion**

The wild type truncation 49-423 PAFAH and the mutant W298H of the same truncation were previously ligated into the pINCOP vector by Emily Hartzell, creating the pINCPAF and pINCPAF298 expression vectors. This work had demonstrated that PAFAH was successfully displayed at a high yield on the surface of the bacterial cells by the functional hydrolysis of 2-thio-PAF, a modified natural substrate of PAFAH. In an attempt to improve the OP hydrolase activity of PAFAH random mutagenesis was pursued in conjunction with the cell surface display protocol described in Cho *et al.*<sup>8</sup> for a medium-throughput screening of the catalytic activity of PAFAH.

Template DNA for random mutagenesis was obtained from the previously expressed 49-423 truncated PAFAH. Random mutagenesis was introduced to the PAFAH template DNA via a reduced efficiency PCR. The reaction included a low fidelity taq polymerase coupled with the modified dNTPS 8-oxo-2'-deoxyguanosine-5'-triphosphate and 2'-deoxy-P-nucleoside-5'-triphosphate to induce random mutagenesis at a predicted rate of 6-19%.<sup>9</sup> The PCR was run for 20 cycles, with fresh template added after 10 cycles to increase the library of mutagenic DNA. The thought process behind the addition of fresh template was to correct any amplification bias toward mutations that may have occurred in the first few rounds of amplification. The reaction was also run in duplicate, with all of the PCR amplified DNA purified

together to create one clean mutagenic library. The DNA library was utilized as the template for a second PCR reaction in which the only goal was amplification of the mutagenic library. This step was completed across four separate preparations that were pooled together and PCR purified by a Qiagen extraction kit.

Plasmid preparation for the subcloning of the mutagenic DNA proceeded via the amplification and purification of the previously created pINCPAF vector. The original vector, obtained from Dr. Wilfred Chen, already contained an OP hydrolyzing enzyme which was removed and replaced by PAFAH. Since PAFAH is incapable of hydrolyzing the OP compounds without mutation, this vector was used to eliminate false positives created by incomplete digestion of the original vector before subcloning the mutagenic DNA. If the original vector made it through the digestion and ligation process, the resulting colony would not be able to hydrolyze methyl paraoxon during the screening phase. The pINCPAF vector, along with the purified and amplified DNA, were both digested with BamHI and HindIII, and ligated overnight with T4 ligase. The ligated sample was mixed directly with BL21 (DE3) cells and transformed via a heat shock method. In contrast to a typical procedure, where the grown transformation would be plated for single colonies, the mixture was grown up overnight at 37 °C. This was done so that a solution of BL21 cells that represented the whole library of mutagenic DNA could be stored for further experimentation.

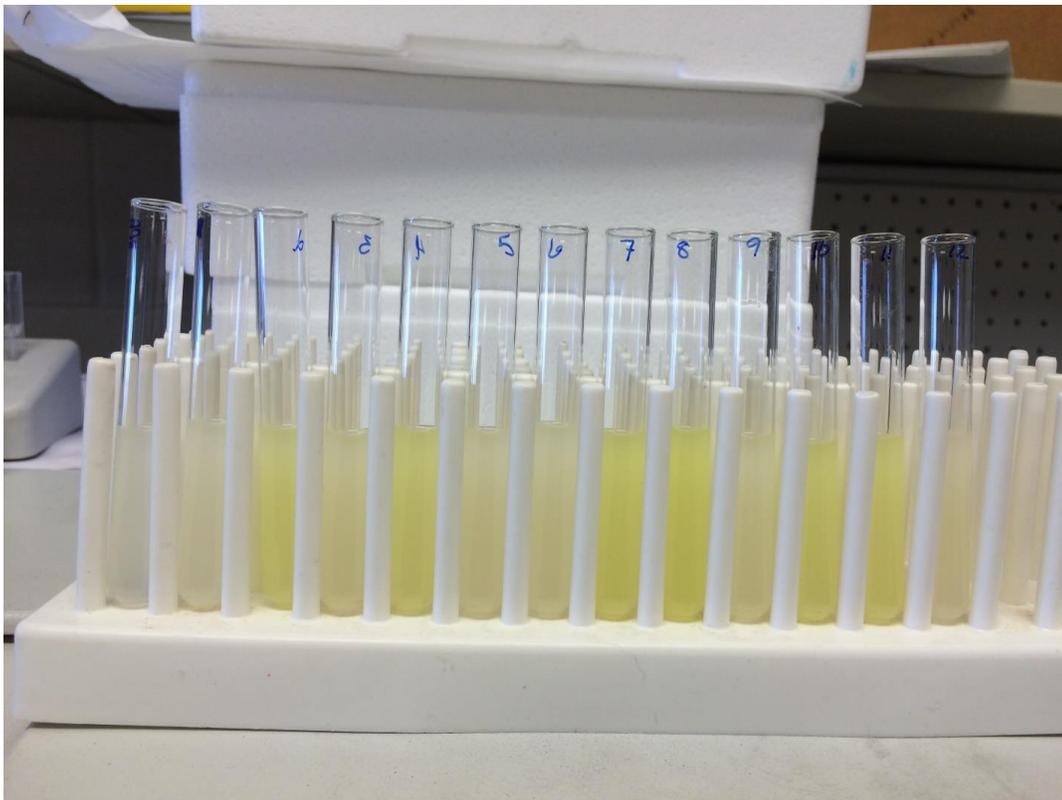
The overnight sample, containing a variety of mutagenic PAFAH vectors, was then used for a medium throughput screening for OP hydrolyzing ability. The subcloned BL21 cells were plated in various amounts on both LB and supplemented

M9 agar, which were previously described in the materials and methods. The M9 supplementation was used to increase growth, without an increase in background. LB has a yellow color and using it might introduce interferences into the analysis. The plates were incubated at 37 °C for 16-24 h before an agar overlay containing a 50 mM phosphate-citrate buffer and 1 μM paraoxon was added to each plate. The colonies were monitored for 30 min for the production of yellow color. Colonies exhibiting the greatest qualitative change in color were selected for overnight 5 mL growth in 100 μg/mL ampicillin. Even with the background yellow hue, LB agar produced a noticeable variation of yellow color. The slightly more labor intensive protocol of making M9 plates was deemed unnecessary, and all further experimentation was completed on LB containing 100 μg/mL of ampicillin.

After the selection of colonies for further screening overnights were prepared. In parallel we also included overnights of the wild type pINCPAF and W298H pINCPAF as positive and negative controls. Upon binding of the OP compound, wild type PAFAH becomes inactivated and is unable to hydrolyze the compound. Thus, no formation of color should be evident in this sample. The directed mutagenesis of the tryptophan to histidine mutation at amino acid 298, previously described by Kirby *et al.* provided the ability to hydrolyze the OP compound albeit with a relatively slow  $k_{cat}$  of 5 min<sup>-1</sup>.

Small aliquots of the reactive overnight samples were removed and stored at -80 °C as stocks. The overnight samples were then centrifuged at 3,200 x g and the supernatants were discarded. The pellets were resuspended with 2 mL of 50 mM

phosphate-citrate buffer, pH 8.0 and turbidity measured for the normalization of cellular content. Additional 50 mM phosphate-citrate buffer containing 5  $\mu$ M of methyl paraoxon was added to each of the resuspended tubes to a final reaction volume of 2.5 mL containing 1  $\mu$ M of methyl paraoxon to begin the reaction. The tubes were immediately mixed and absorbance at 405 nm measured. The reaction progressed for 10 min before a second absorbance at 405 nm was measured and colonies maintaining normalized catalytic ability above that of the W298H PAFAH mutant were retained. An example of this assay can be visualized in Figure A.3.



**Figure A.3:** The colorimetric reaction of PAFAH. Reaction follows the hydrolysis of methyl paraoxon to para-nitrophenol displayed in wild type, W298H, and the randomly mutagenic PAFAH. The reaction includes 2.5 mL of resuspended bacteria, containing a cellular surface display of the associated PAFAH protein, in 50 mM phosphate-citrate buffer pH 8.0 and 1  $\mu$ M methyl paraoxon. The first tube on the left contains the negative control of wild type PAFAH, which is incapable of hydrolyzing the OP compound. The second tube from the left contains the positive W298H PAFAH control, which displays a small level of hydrolysis. Each of the remaining tubes account for colonies selected from previous screening of LB agar plates treated with an overlay containing 1  $\mu$ M methyl paraoxon.

This assay demonstrates the effectiveness of using random mutagenesis in conjunction with the pINCOP cell surface display vector to screen for colonies with an increased ability for PAFAH to hydrolyze OP compounds. As can be visualized in Figure A.3, several colonies were produced with a greater activity than the positive

control, W298H. It is also noticeable that several of the colonies exhibited behavior similar to the positive control. During the initial screening phase, the change in color was monitored after a 30 min incubation. This change was observed qualitatively, meaning that colonies with a small to negligible change in hydrolytic capability were not selected, but could be with a more quantitative approach. This quantitative assay is one of the future goals of this project, outlined in the following section.

This qualitative assay does not suggest that each of the active colonies contain separate mutations. Due to the amplification process during mutagenesis, a mutation created early on in the PCR process would have been amplified exponentially throughout both PCR reactions. A single mutation may be represented across several colonies within this screening process. Colonies exhibiting the highest activity against methyl paraoxon were selected for further screening.

The pINCOP vector was utilized significantly in this work due to its ability to localize proteins to the cell surface; however, it is not useful for large scale expression and purification methods. For sequencing purposes, as well as continued screening, the mutated PAFAH will need to be subcloned into the pGEX-6P-1 vector. The DNA was amplified from the pINCOP vector via PCR primers for the 49-423 PAFAH truncation containing a BamHI and SalI restriction enzymes. Experimental difficulties have arisen while using the SalI restriction enzyme and the subcloning of the mutated PAFAH into pGEX-6P-1 has not been successfully obtained. A new reverse primer changing the restriction enzyme will need to be designed for insertion into pGEX-6P-1 to successfully insert the mutated sequences.

#### **A.4 Conclusions and Future Directions**

Cell surface display of the PAFAH enzyme proved to be critical to achieving medium throughput screening of PAFAH mutant activity against the OP compound methyl paraoxon. This work demonstrates that the combination of random mutagenesis and cell surface display of the mutated enzymes is a successful approach for the monitoring the catalytic activity of PAFAH. The results obtained in this work provided several colonies of interest. These colonies exhibited an increase in methyl paraoxon hydrolysis efficiency and will be utilized moving forward.

Going forward, a new reverse primer for the subcloning of each of the colonies of interest should be designed for the successful insertion of mutated DNA into the pGEX-6P-1 expression vector. This vector will not only allow for the sequencing of each of these mutagenic DNA samples, but also the expression and purification of the mutants for kinetic characterization of purified enzyme. Colonies containing mutations of increased catalytic efficiency could be crystallized, leading to a structural explanations for their activity. These mutants with increased OP hydrolase activity will next be evaluated for activity against OP nerve agents in collaboration with USAMRICD. Finally these colonies could be used as templates for further mutagenesis, random or directed, to further increase the catalytic efficiency and create viable catalytic bioscavengers for the clearance of dangerous acetylcholinesterase inhibitors.

## REFERENCES

- [1] Nomura, D. K., Fujioka, K., Issa, R. S., Ward, A. M., Cravatt, B. F., and Casida, J. E. (2008) Dual roles of brain serine hydrolase KIAAL1363 in ether lipid metabolism and organophosphate detoxification, *Toxicology and Applied Pharmacology* 228, 42-48.
- [2] Buchebner, M., Pfeifer, T., Rathke, N., Chandak, P. G., Lass, A., Schreiber, R., Kratzer, A., Zimmermann, R., Sattler, W., Koefeler, H., Froehlich, E., Kostner, G. M., Birner-Gruenberger, R., Chiang, K. P., Haemmerle, G., Zechner, R., Levak-Frank, S., Cravatt, B., and Kratky, D. (2010) Cholesteryl ester hydrolase activity is abolished in HSL<sup>-/-</sup> macrophages but unchanged in macrophages lacking KIAA1363, *Journal of Lipid Research* 51, 2896-2908.
- [3] Thevenin, A. F., Monillas, E. S., Winget, J. M., Czymbek, K., and Bahnson, B. J. (2011) Trafficking of Platelet-Activating Factor Acetylhydrolase Type II in Response to Oxidative Stress, *Biochemistry* 50, 8417-8426.
- [4] Prescott, S. M., Zimmerman, G. A., Stafforini, D. M., and McIntyre, T. M. (2000) Platelet-activating factor and related lipid mediators, *Annual Review of Biochemistry* 69, 419-445.
- [5] Samanta, U., and Bahnson, B. J. (2008) Crystal Structure of Human Plasma Platelet-activating Factor Acetylhydrolase STRUCTURAL IMPLICATION TO LIPOPROTEIN BINDING AND CATALYSIS, *Journal of Biological Chemistry* 283, 31617-31624.
- [6] Samanta, U., Kirby, S. D., Srinivasan, P., Cerasoli, D. M., and Bahnson, B. J. (2009) Crystal structures of human group-VIIA phospholipase A2 inhibited by organophosphorus nerve agents exhibit non-aged complexes, *Biochemical Pharmacology* 78, 420-429.
- [7] Kirby, S. D., Norris, J., Sweeney, R., Bahnson, B. J., and Cerasoli, D. M. (2015) A rationally designed mutant of plasma platelet-activating factor acetylhydrolase hydrolyzes the organophosphorus nerve agent soman, *Biochimica Et Biophysica Acta-Proteins and Proteomics* 1854, 1809-1815.
- [8] Cho, C. M. H., Mulchandani, A., and Chen, W. (2002) Bacterial cell surface display of organophosphorus hydrolase for selective screening of improved hydrolysis of organophosphate nerve agents, *Applied and Environmental Microbiology* 68, 2026-2030.
- [9] Zaccolo, M., Williams, D. M., Brown, D. M., and Gherardi, E. (1996) An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues, *Journal of Molecular Biology* 255, 589-603.

- [10] Shimazu, M., Mulchandani, A., and Chen, W. (2001) Cell surface display of organophosphorus hydrolase using ice nucleation protein, *Biotechnology Progress* 17, 76-80.

**Appendix B**  
**LICENSE AGREEMENTS**

**ELSEVIER LICENSE  
TERMS AND CONDITIONS**

Jul 17, 2016

This Agreement between Meghan M Klems ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	3911490969220
License date	Jul 17, 2016
Licensed Content Publisher	Elsevier
Licensed Content Publication	Biochemical Pharmacology
Licensed Content Title	Crystal structures of human group-VIIA phospholipase A2 inhibited by organophosphorus nerve agents exhibit non-aged complexes
Licensed Content Author	Uttamkumar Samanta, Stephen D. Kirby, Prabhavathi Srinivasan, Douglas M. Cerasoli, Brian J. Bahnson
Licensed Content Date	15 August 2009
Licensed Content Volume Number	78
Licensed Content Issue Number	4
Licensed Content Pages	10
Start Page	420
End Page	429
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	
Original figure numbers	1
Title of your thesis/dissertation	Functional Characterization of the Serine Hydrolase Neutral Cholesterol Ester Hydrolase 1.
Expected completion date	Aug 2016
Estimated size (number of pages)	100
Elsevier VAT number	GB 494 6272 12
Requestor Location	Meghan M Klems University of Delaware Department of Chem. and Biochem. Drake Hall 316 NEWARK, DE 19716 United States