Development of an open-source data analysis platform for high throughput screening in drug discovery and its application to E3 ligase IDOL inhibitor discovery project



Submitted in partial fulfilment of the requirements of Doctor of Philosophy

By Thomas Christott

Green Templeton College

University of Oxford

Hilary Term 2024

Abstract

Alzheimer's disease (AD) and cardiovascular diseases (CVDs) pose a major challenge in the coming decades considering an increasingly ageing population and demographic changes. MYLIP, also known as inducible degrader of the LDL receptor (IDOL) is an E3 ligase that is involved in the downregulation of lipoprotein receptors: the low-density lipoprotein receptor (LDLR) and the very low-density lipoprotein receptor (VLDLR). LDLR and VLDLR are involved in plasma cholesterol clearance in the liver and Aβ plaque removal in the brain, respectively. Therefore, MYLIP is of great interest as a potential therapeutic target for two of the biggest challenges in aged care. Here, I have developed methods to express MYLIP and other proteins in the ubiquitination cascade of the lipoprotein receptors, assays to screen small molecule libraries for inhibitors of the function of MYLIP and used the assay to conduct a pilot screen with a subset of a diversity library held by the Oxford Drug Discovery Institute. The resulting data set was then analysed in a workflow in the software that I have developed in this project.

Reliable, easy to use data analysis software is a corner stone of drug discovery efforts. Especially in academia, there exists a need for easy to use, affordable, reliable, and adaptable software that can be used in place of disparate templates for different pieces of software that have often been developed by long gone group members. Here, I have developed such a software that offers ease of use, is modular and allows for quick development of new workflows for novel assays. In addition, it is easily integrated into an Oracle database. Biophysical and Biochemical Quantifier and Qualifier (BBQ) is now used in our department for the analysis of routine peptide displacement assays, enzymatic assays, and differential scanning fluorimetry, to determine IC50 values, ΔT_m values, and screening compound libraries in single dose screens.

Declaration

This thesis and the work presented in it were completed at the University of Oxford (Nuffield Department of Medicine) between October 2019 and December 2023 under supervision of Professor Dr Paul E Brennan and Dr Oleg Fedorov.

All contributions made by others have been clearly stated in the text. This work has not been submitted in full or partial fulfilment of any other degree.

Oxford, 25 January 2024



Thomas Christott

Acknowledgements

First and foremost, I would like to express gratitude to Dr Suzanne English MBBS Hons (UWA) DPhil (Oxon) MRCP, my wonderful wife and mother our wonderful children, Thomas and Norah. Without her continued emotional support, insight, and wisdom, none of this would have been possible. I dedicate this work to her and the little ones, who had to do without me for so long during the COVID lockdowns while I was doing the lab work.

I would like to thank my supervisors, Dr Oleg Fedorov and Prof Paul Brennan, for giving me the opportunity to pursue this project and for their continued support and guidance. Special thanks to Dr Oleg Fedorov for encouraging me to pursue a DPhil in the first place.

Furthermore, I would like to thank my fellow team members, James Bennett, Charline Giroud, and Karly Buchan for their support, advice or simply letting me bounce ideas off them. Thank you! Charline Giroud deserves a special thanks for her work on the PTB domains and phosphormimetic mutants of MYLIP while I was on paternity leave. I would also like to thank Dr Emma Murphy, who has been working alongside us in the Alzheimer Research UK funded Oxford Drug Discovery Institute for many years and has also provided invaluable scientific and moral support.

I would also like to thank the members of my thesis committee Dr Emma Mead, Dr Opher Gileadi, and Prof Alex Bullock. They have helped me stay on track and have provided invaluable constructive criticism.

My gratitude also to everyone in the Centre for Medicine Discovery's Biotechnology group who have supported me with the high throughput cloning pipeline, especially Dr Alejandra Fernandez-Cid, Dr Eleanor Williams, Leela Shrestha, as well as Korvus Wang, who has supplied me with several expression plasmids from the departmental library.

Further constructs and proteins have been supplied by Dr Vittorio Katis, helping me get to the bottom of a false positive result, as well as Dr Michael Fairhead and Dr Zhuoyao Chen, allowing me to establish the MYLIP auto-ubiquitination assay.

The computational side of the project would not have been possible without Lucas Martins Ferreira, who introduced me to Python while I was briefly seconded to efforts to establish a

data analysis platform to determine the Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers during the first lockdown in 2020. Thank you, Lucas!

I would also like to thank Prof Daniel Ebner, Dr Stephanie Bundy Hatch and Dr Timothy Johanssen for letting me support their work on phenotypic screens.

Finally, I would like to thank my proofreaders, Dr Emma Murphy, Dr Eleanor Williams, Dr Francesca Nicholls, and Dr Oleg Fedorov who have helped me get this work over the line.

Table of Contents

Abstract 2

Declaration 3

Acknowledgements 4

Table of Contents 6

Abbreviations 14

Chapter 1 – Introduction 16

1.1. Ageing and life-style related diseases 16

1.1.1. What are ageing-related diseases? 17

1.1.2. What is a 'lifestyle' related disease? 17

1.2. Alzheimer's Disease 18

1.2.1. Epidemiology and economic impact 18

1.2.2. Diagnosis and Neuropathophysiology 18

1.2.3. Genetic risk factors 21

1.2.4. Current treatment options 22

1.2.5. Future treatment options 23

1.3. Cardiovascular disease 24

1.3.1. Cholesterol homeostasis in health and disease 26

1.3.2. Cholesterol transport in the blood - Low density lipoprotein 28

1.3.3. Dysregulation of cholesterol homeostasis 30

1.3.4. Prevention and current treatments for CVD 34

1.4. Molecular targets of interest 35

1.4.1. Choice of molecular targets for study 35

1.4.2. Inducible degrader of the low-density lipoprotein receptor (IDOL) 35

1.4.3. Assays 44

1.5. An overview of data analysis in high throughput screening 47
1.5.1. The SGC/CMD Biophysical and Biochemical Screening Group – a case study 47
1.5.2. Identifying requirements and choosing an environment 48
1.6. Summary and aims 48
1.6.1. Alzheimer's and cardiovascular disease 48
1.6.2. Software development 50
Chapter 2 – Materials and Methods 52
2.1 General methods 52
2.1.1 Plasmid Isolation (Mini Preps) 52
2.1.2 Plasmid Concentration determination 53
2.1.3 Bacterial Transformation 53
2.1.4 Expression Vector Selection 54
2.1.5 Agarose gel electrophoresis 55
2.1.6 T4 DNA polymerase treatment 56
2.1.7 Polymerase Chain Reaction 56
2.1.8 Ligation independent cloning 58
2.1.9 Glycerol stock plate 60
2.1.10 Small scale test expression and purification 60
2.2 MYLIP construct design and cloning 61
2.2.1 Site directed mutagenesis 62
2.2.1.1 Reaction set-up 62
2.2.1.2 PCR Protocol 62
2.3 Enzymes of the ubiquitination cascade 63
2.3.1 Ubiquitin-activating enzyme (E1) 63
2.3.2 Ubiquitin-conjugating enzymes (E2s) 63
2.3.3 Ubiquitin 64

2.4 LDLR adapter proteins 65
2.5 Protein Expression 66
2.5.1 Media and antibiotics 66
2.5.2 Expression strains 67
2.5.3 Transformation 67
2.5.4 Expression 67
2.6 MYLIP Protein purification 68
2.6.1 Buffers 68
2.6.2 Lysis 68
2.6.3 Immobilised Metal Affinity Chromatography (IMAC) 68
2.6.4 Protein concentration 69
2.6.5 Size Exclusion Chromatography (SEC) 69
2.6.6 TEV protease digest 69
2.6.7 Protein concentration determination via UV_{280} absorbance 69
2.7 Murine E1 purification 69
2.7.1 Lysis 70
2.7.2 IMAC 70
2.7.3 Ion Exchange Chromatography (IEX) 70
2.7.4 Size exclusion chromatography 71
2.8 Ubiquitin purification 71
2.8.1 Lysis 71
2.8.2 IMAC 71
2.8.3 BCA assay 71
2.9 UBE2D construct purification 72
2.10 Quality control and protein identification 72
2.10.1 Gel electrophoresis 72

2.10.2 Intact mass liquid chromatography and mass spectrometry (LC/MS) 72
2.10.3 Thermal stability 73
2.11 Peptides 73
2.12 Biochemical/biophysical assays 74
2.12.1 Peptide displacement assays 74
2.12.2 Microscale Thermophoresis (MST) 80
2.12.3 Biolayer Interferometry (BLI) 80
2.12.4 Isothermal Titration Calorimetry (ITC) 82
2.12.5 Differential Scanning Fluorimetry (DSF) 82
2.13 Enzymatic assays – substrate ubiquitination 83
2.13.1 Detection of LDLR ubiquitination via Western blot 83
2.13.2 Detection of LDLR ubiquitination via gel fluorescence 84
2.13.3 HTRF based ubiquitination assay – discrete timepoints 84
2.13.4 Continuous HTRF based ubiquitination assay 86
2.13.5 MUBE1-p002 batch validation 87
2.13.6 Testing additional members of LDLR superfamily 88
2.14 MYLIP autoubiquitination assay 89
2.14.1 MUBE1A-p003.2 batch validation 89
2.14.2 Compound dispensing 90
2.14.3 Pilot screen 91
Chapter 3 - Results of Protein Production and Assay Development 92
3.1 Introduction 92
3.2 Protein construct design and production 92
3.2.1 MYLIP construct design, cloning and test expressions 93
3.2.2 Full scale MYLIP protein expression and purification 96
3.2.3 Ubiquitin cascade enzyme production 105

3.2.4 LDLR adapter protein construct design and production 117
3.2.5 Summary and Discussion 117
3.3 Binding of MYLIP FERM domain to LDLR family receptor tail peptides 119
3.3.1 Titrations of MYLIP constructs against LDLR family receptor tails in HTRF, AlphaScreen and Fluorescence polarisation 121
3.3.2 Cooperativity of MYLIP and LDLR adapter proteins 125
3.3.3 Other biophysical assays 128
3.3.4 Summary and Discussion 135
3.4 Ubiquitination of receptor tail peptides by full length recombinant MYLIP 136
3.4.1 Detection of LDLR ubiquitination via Western Blot 136
3.4.2 HTRF based ubiquitination assay with discrete timepoints 136
3.4.3 Continuous HTRF based ubiquitination assay 138
3.4.4 Summary and Discussion 146
3.5 Using the MYLIP autoubiquitination assay to screen for MYLIP inhibitors 147
3.5.1 Time window for auto-ubiquitination assay 147
3.5.2 MYLIP auto-ubiquitination assay pilot screen 147
3.5.3 Identified “Hits” 151
3.5.4 Summary and Discussion 156
Chapter 4 – Results of Software Development 157
4.1 Introduction 157
4.2 Identification of requirements 157
4.2.1 All in one solution 157
4.2.2 Accessibility 158
4.2.3 Automated and manual data processing 159
4.2.4 Database integration 160
4.2.5 Future proofing 162

4.3 Developed analysis workflows 162
4.3.1 Common UI elements 162
4.3.2 Commonalities in behind-the-scenes data processing 165
4.3.3 End Point Single Dose assays 168
4.3.4 End Point Dose Response assays 168
4.3.5 Differential Scanning Fluorimetry 171
4.3.6 Cell Based Compound Screen 177
4.3.7 Enzymatic Reaction Rate Dose Response assays 181
4.4 Saving completed analyses for later use 183
4.5 Workflow editor 186
4.5.1 Assay definition file format 186
4.5.2 Workflow editor steps 187
4.5.3 Workflow interpreter 193
4.6 Dissemination 193
4.7 Summary and Discussion 193
Chapter 5 – Discussion 196
5.1 MYLIP 196
5.1.1 Introduction 196
5.1.2 Isolated MYLIP FERM domain expresses well in *E. coli* 196
5.1.3 Isolated MYLIP FERM domain does not bind to receptor tail peptides in AlphaScreen or HTRF assays 197
5.1.4 PTB domains of DAB1 and DAB2 bind to receptor tails, but cannot facilitate binding to MYLIP FERM domain 199
5.1.5 A serendipitously developed MYLIP auto-ubiquitination assay 200
5.1.6 MYLIP auto-ubiquitination pilot screen. 202
5.1.7 Limitations 203

5.1.8 Validation of hits 204
5.1.9 Future perspectives 205
5.2 Software development discussion 206
5.2.1 Introduction 206
5.2.2 End point single dose (EPSD) workflow 207
5.2.3 End Point Dose Response (EPDR) workflow 207
5.2.4 Differential Scanning Fluorimetry (DSF) 208
5.2.5 Cell Based Compound Screen (CBCS) 208
5.2.6 Enzymatic Reaction Rate Dose Response (ERDR) 208
5.2.7 Extensibility 208
5.2.8 Summary 209
5.2.9 Future directions 209
References 211
Appendix 231
DNA templates 231
Vector Properties 234
MYLIP protein constructs 235
MYLIP Primers 237
MYLIP test expressions 241
Full scale MYLIP expressions and purifications 244
LDLR adapter protein constructs 255
LDLR adapter proteins purification 258
Phospho-mimetic MYLIP mutants 270
Site directed mutagenesis primers: 270
Sequencing results: 270
Murine E1 construct 274

Protein-peptide titrations in AlphaScreen, HTRF and fluorescence polarisation assays ...275

3.5.5 HTRF – MYLIP$_{(1-276;\ C\text{-}His6)}$ vs LDLR$_{(811-834)}$-K-biotin 275

3.5.6 HTRF – MYLIPA$_{(1-273;\ N\text{-}His6\ SUMO)}$ vs LDLR$_{(811-834)}$-K-biotin.....275

3.5.7 HTRF assay –different buffers 276

3.5.8 HTRF – Reagent titration MYLIPA-p011, biotin-ApoER2$_{(849-871)}$.....278

4.7.1 LDLR adapter protein PTB domains and LDLR family peptides.....280

Abbreviations

APC	allophycocyanin
AD	Alzheimer's disease
APP	amyloid precursor protein
APOE	apolipoprotein E
ApoER2	apolipoprotein E2 receptor
ARIA	amyloid related imaging abnormalities
ABC	ATP binding cassette
ABCA1	ATP binding cassette protein A1
BLI	biolayer Interferometry
CVD	cardiovascular disease
CSF	cerebrospinal fluid
CT	computer tomography
DDM	n-Dodecyl β-D-maltoside
DSF	differential scanning fluorimetry
EEG	electroencephalography
FRET	Förster resonance energy transfer
LISA	enzyme-linked immunosorbent assay
ELN	electronic lab notebook
ER	endoplasmic reticulum
HA	hemagglutinin tag
HTRF	homogenous time resolved Förster resonance energy transfer
IMAC	immobilised metal affinity chromatography
IDOL	inducible degrader of the LDL receptor
IEX	ion exchange chromatography
ITC	isothermal titration calorimetry
KNIME	Konstanz Information Miner
LDLR	LDL receptor
LXR	liver X receptor
LDL	low-density lipoprotein
LB	Lysogenic broth
MCI	mild cognitive impairment
mAb	monoclonal antibody
MYLIP	myosin regulatory light chain interacting protein
NCD	non-communicable disease
PAGE	poly-acrylamide gel electrophoresis
PCR	polymerase Chain Reaction
PET	positron emission tomography
RCF	relative centrifugal force

RXR	retinoid X receptor
RCT	reverse cholesterol transport
RPM	revolutions per minute
SDM	site directed mutagenesis
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SRE	Sterol regulatory element
SREBP	sterol regulatory element binding receptor
TB	terrific broth
TCEP	tris(2-carboxyethyl)phosphine
VLDL	Very low-density lipoprotein
VLDLR	VLDL receptor

Chapter 1 – Introduction

1.1. Ageing and life-style related diseases

Ageing populations are a major concern for many developed and developing nations. In the EU alone, the proportion of the population 80 years and older grew from 3.4% to 6.1% over a 20 year period, while there has been a marked decline in the age groups 19 years and under (23.2% in 2001 to 20.2 in 2022) and 20-39 years (29.3% in 2001 to 23.8% in 2022)[1](**Figure 1**). With this comes an increase in non-communicable diseases (NCDs), many of which are age- and ageing-related(as reviewed in [2]). These diseases are exerting pressures on healthcare systems and the aged care sector. Assuming governments and the healthcare sector follow "business as usual", one study predicts an increase in per-capita cost of 30% for all NCDs combined by 2050[3]. Even under best case predictions, the costs are projected to rise significantly. Given this predicted burden, new approaches to treatment and prevention must be found. Targeting modifiable risk factors comes at a relatively low cost and has the greatest potential to reduce costs.

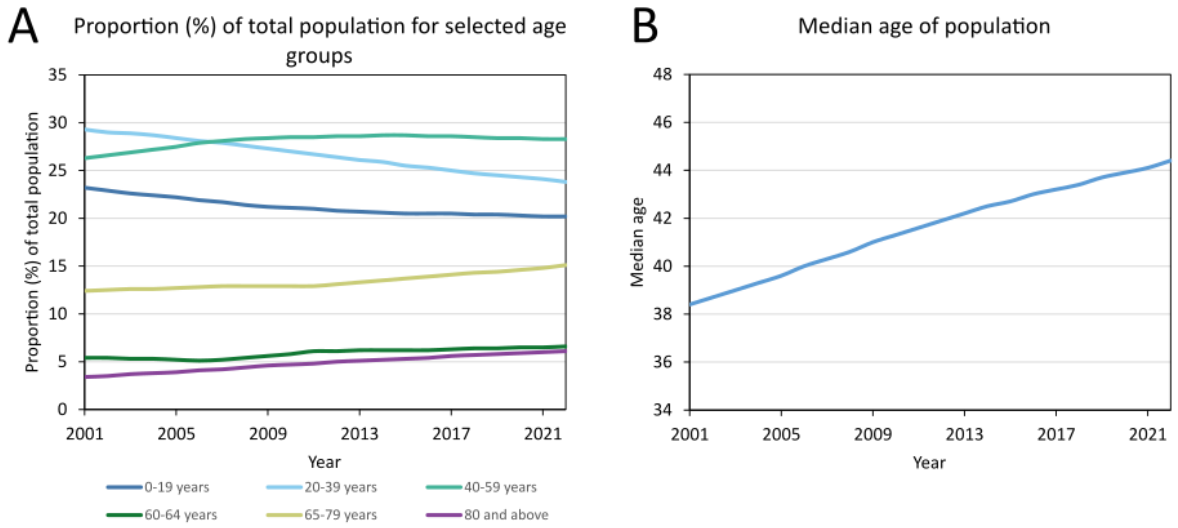


Figure 1: Development of population age in EU27 from 2001 to 2022. **A:** Proportion (%) of total population for selected age groups. **B:** Development of median age. Data taken from [1]

1.1.1. What are ageing-related diseases?

Ageing is the greatest risk factor for many human diseases, including cancer, diabetes, cardiovascular and neurodegenerative diseases. In 2013, López-Otin et al. defined 9 hallmarks of ageing[4], namely:

- Altered intercellular communication,
- Genomic instability,
- Telomere attrition,
- Epigenetic alterations,
- Loss of proteostasis,
- Deregulated nutrient sensing,
- Mitochondrial dysfunction,
- Cellular senescence, and
- Stem cell exhaustion.

Not only do some of these hallmarks lead to an increased risk of tipping the body into pathological dysfunctions by themselves, but they also signify a decline in the human body's ability to deal with damage incurred from environmental (e.g., UV, pollution) and lifestyle dependent factors (e.g., diet, lack of exercise, sports or work-related injuries, sleep quality).

It is worth making a distinction between "age-related" and "ageing-related" diseases. Since these terms are often used interchangeably (as discussed by Kuan et al. [5]), it can be difficult to differentiate between diseases that are non-communicable and develop due to the hallmarks of ageing and diseases that are simply occurring at different life stages. Following the definition of Kuan et al., Le Couteur and Thillainadesan developed an epidemiological perspective of ageing related diseases[6]. Besides cancer, the two most represented disease areas, were cardiovascular diseases and dementias. Not only do the incidence of Alzheimer's disease, cardiomyopathy and myocarditis follow a near exponential trajectory with age, they are also very similar for high and low sociodemographic indices ([6], Figure 2 in reference).

1.1.2. What is a 'lifestyle' related disease?

In contrast, while many lifestyle-related diseases do have non-modifiable genetic risk factors, they are chiefly defined by modifiable risk factors and thus they are especially amenable to prevention, i.e., lifestyle changes[7].

However, lifestyle is not just defined by active choices, but also opportunities (or a lack thereof), such as access to education, access to healthy food and the resources to prepare it, access to exercise, security and safety (including workplace safety), as well as social contacts[8]. Like ageing related diseases, they are often non-communicable.

1.2. Alzheimer's Disease

Alzheimer's Disease (AD) is the most common form of dementia, accounting for 60-70% of cases in adults aged 65 and over[9]. First described in 1907 by Alois Alzheimer, it is a progressive and irreversible neurological disorder affecting cognitive function, memory, language, behaviour and personality[10], [11]. This leads to progressive disability in all domains of daily life, often resulting in a complete inability to function independently and increasing susceptibility to other serious medical complications and death.

1.2.1. Epidemiology and economic impact

Around 100,000 people die of AD every year in the EU[12]. Prevalence of AD is generally higher in women than in men, with rates ranging from 0.6 for the 60-64 age range up to 21.9 for the 85-89 age range for both sexes. Recent projections indicate that the number of people living with any type of dementia will double in Europe by 2050, from 1.57% to 3.00% of the total population. Worldwide, the prevalence is estimated to triple [13].

Dementia patients require almost constant care and thus the costs associated are high. They are also region dependent. A meta-analysis of studies from 2008 to 2021 showed estimated annual costs per patient ranging from around €74,000 (British Isles) to €8000 (Eastern Europe and Baltics)[14], outpacing the median equivalised disposable income for those regions[15].

1.2.2. Diagnosis and Neuropathophysiology

Criteria for the diagnosis of AD were originally defined in 1984 by a working group convened by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA) [16] (from here on referred to as NINCDS-ADRDA criteria), while distinguishing between *probable*, *possible* and *definite* AD.

The diagnostic criteria for *probable* AD as defined by the NINCDS-ADRDA working group are:

> “
> i. dementia established by clinical examination and documented by the “Mini-Mental Test”, “Blessed Dementia Scale”, or some similar examination, and confirmed by neuropsychological tests
> ii. deficits in two or more areas of cognition
> iii. progressive worsening of memory and other cognitive functions
> iv. no disturbance of consciousness
> v. onset between ages 40 and 90, most often after age 65; and
> vi. absence of systemic disorders or other brain diseases that in and of themselves could account for the progressive deficits in memory and cognition
> ”

For a *definite* diagnosis, histopathologic evidence is required, either obtained via biopsy or autopsy *postmortem*. To further differentiate between AD and other types of dementia, Electroencephalography (EEG), computer tomography (CT), measurement of regional cerebral blood flow (rCBF), positron emission tomography (PET) and magnetic resonance imaging (MRI) are discussed.

Amyloid beta (Aβ) plaques are generally regarded as the cause of AD (Amyloid hypothesis, **Figure 2**), although there is still no consensus in the field about whether they are causing the disease or are merely a symptom. The amyloid precursor protein (APP) is sequentially cleaved by the enzymes β-secretase (BACE1) and γ-secretase (made up out of four subunits) into Aβ peptides of varying length, most commonly $A\beta_{38}$, $A\beta_{40}$ and $A\beta_{42}$. These peptides then aggregate outside neurons in the brain parenchyma, leading to neuronal dysfunction and death, and therefore to a decline in cognition and eventually the patient’s death. Aβ plaques were first described by Alzheimer, although the identity of these “minute miliary foci”[11] was not determined until the 1980s[17]. Interestingly, Aβ peptides are not formed if APP is first cleaved by α-secretase instead of β-secretase[18].

Since the definition of the original criteria, the understanding of the underlying pathology of AD has advanced. Similarly, biomarkers for diagnosis and new imaging and other diagnostic techniques have become available, prompting some to review old diagnoses. An analysis of

the American National Alzheimer's Coordinating Center Uniform Data Set (2005 – 2010) showed that 39% of dementia cases clinically diagnosed as not being AD did show AD neuropathophysiology [19], underlining the need to bring these new developments into the clinic.

In 2011, the NINCDS-ADRDA criteria were updated, and the National Institute on Aging (NIA) and the Alzheimer's Association work group published a series of articles in Alzheimer's and Dementia (reviewed in [20]) and proposed a distinction between three stages of AD:

1. Preclinical, where significant changes in biomarkers are observed before the onset of symptoms.
2. Mild cognitive impairment (MCI), where first changes in memory and other cognitive abilities manifest, such as difficulty paying bills and making plans. While these changes are said to not interfere with daily life, they are significant enough to be detected.
3. Dementia caused by AD pathophysiology.

Today, AD is diagnosed by a combination of clinical/functional tests, PET imaging of $A\beta_{42}$ plaques and tau tangles (in separate scans), and cerebrospinal fluid (CSF) tests for the presence of $A\beta_{42}$, phosphorylated tau (p-tau) and total tau (t-tau). A meta-analysis by Olsson et al.[21] found that, generally, the ratio of AD to healthy control for these biomarkers are about 0.56 for $A\beta_{42}$, 1.99 for p-tau and 2.54 for t-tau. Meaning they are lowered, elevated, and elevated, respectively, compared to healthy individuals. The authors note that there was great variability between laboratories and commercial assays, none of which were clinically certified at the time of the studies. Since then, four commercial tests from Roche have received "FDA breakthrough device designation to enable a timely diagnosis" of AD[22].

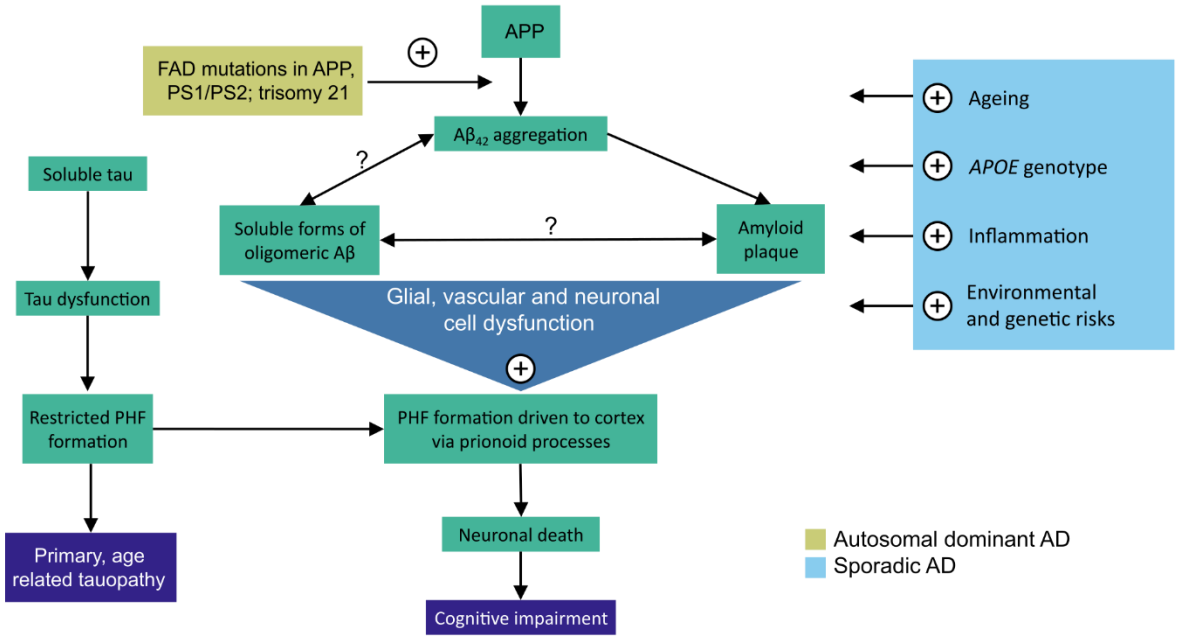


Figure 2: Schematic of amyloid cascade hypothesis. PHF: paired helical tau filaments, FAD: familial Alzheimer's disease, APP: Amyloid precursor protein, PS: presenilin. Reproduced from [23] under licence from Springer Nature, licence number 5673221299410, 20 November 2023.

1.2.3. Genetic risk factors

While other types of dementia have mostly modifiable risk factors, the risk of developing AD is defined 60-80% by heritable factors [24]. There are three genes which confer the largest genetic risk of developing AD, all of which are involved in Aβ processing, leading to increased Aβ. Namely the genes encoding the two catalytic subunits of γ-secretase (presenilin 1 (PSEN1) and presenilin 2 (PSEN2)), and amyloid precursor protein (APP).

Another major risk factor is carrying the *APOE* (the gene encoding for apolipoprotein E (ApoE)) allele *APOE-ε4*, which is also implicated in atherosclerosis. ApoE is part of several classes of lipoprotein particles, e.g., VLDL, and interacts with the lipoprotein receptors low density lipoprotein receptor (LDLR), very-low-density lipoprotein receptor (VLDLR) and apolipoprotein E receptor 2 (ApoER2). ApoE is shown to bind to Aβ plaques and facilitates both their proteolytic degradation in microglia and by extracellular proteases[25].

1.2.4. Current treatment options

Currently, four drugs (donepezil, galantamine, rivastigmine and memantine) are approved for symptomatic treatment of AD in the US[26]–[29], UK[30] and EU[31], [32]. All four are only targeting neurotransmitter imbalance without modifying the underlying pathology and are thus only able to slow down disease progression.

Recently, two disease modifying treatments have become available in the USA: Two monoclonal antibodies (mAb) targeting Aβ plaques and marking them for removal by macrophages. The first drug, aducanumab, was approved in 2021[33], although it remains controversial as there was conflicting data on its efficacy[34] and the drug was approved despite 10 of 11 members of an advisory committee finding that the drug was not effective[35]. Biogen has subsequently withdrawn the marketing authorisation application with the European Medicines Agency[36]. The second drug, lecanemab (trade name Leqembi, Eisai, Inc.), was approved in January 2023 in the US. However, it carries a "black box warning", warning of possible intracerebral haemorrhages for homozygotic carriers of the *APOE-ε4* allele[37]. While promising, lecanemab only leads to a "moderately less decline on measures of cognition and function". Two further Aβ targeting mAbs (gantenerumab[38] and donanemab[39], [40]) have also recently completed phase three trials. However, these trials have also only shown a modest reduction in the rate of cognitive decline, and only in early stages of the disease. With all four mAbs having recently been developed, long term data on their effect on disease progression is still outstanding.

Table 1: AD drugs currently in the clinic

Drug	Trade name(s)	Indication	Target/mechanism	First Approval
Donepezil	Aricept	Mild to moderate AD	Acetyl-choline esterase inhibitor	2006[26]
Galantamine	Reminyl	Mild to moderate AD	Acetyl-choline esterase inhibitor	2001[27]
Rivastigmine	Exelon	Mild to moderate AD	Acetyl-choline esterase inhibitor	2001[28]
Memantine	Ebixa	Moderate to severe AD	*N*-methyl-d-aspartate receptor antagonist	2002[31]
Aducanumab	Aduhelm	Mild cognitive impairment, mild dementia	Anti-Aβ monoclonal antibody	2021[33]
Lecanemab	Leqembi	Early-stage AD	Anti-Aβ monoclonal antibody	2023[41]

1.2.5. Future treatment options

Based on currently available data, all available treatment options are only treating the symptoms of AD or showing a moderate slowing of the progression of the disease. Additionally, all four mAbs discussed carry a high risk of amyloid related imaging abnormalities with oedema (ARIA-E) or with haemorrhages (ARIA-H) at the site of plaque clearance.

There is a clear unmet need to halt or ideally reverse the formation of Aβ plaques and NFTs. This could be accomplished by triggering the removal of Aβ plaques/NFTs or preventing their formation in the first place. Potential targets could include the secretases processing APP into Aβ peptides or the enzymes that regulate Aβ plaque clearance in the brain. Although γ-secretase inhibitors have been abandoned due to severe side effects resulting from the impact on other γ-secretase substrates. Therefore, modulation instead of inhibition of γ-secretase has become a focus of study (as reviewed in [42]).

1.3. Cardiovascular disease

Cardiovascular diseases (CVDs) are currently the most common cause of death (e.g. 38% of all deaths in the EU, **Figure 3**, [12]) and a major cause of global disease burden (20% for women, 24% for men, Global Burden of Disease (GBD) study, 2016, as reviewed in [43]). In addition to the healthcare costs, CVD also leads to a significant loss in economic productivity, with one study by the European Society of Cardiology (ESC) estimating a loss in 2018 of €62 billion across all ESC members, with €37 billion incurred by the 28 EU member states[44].

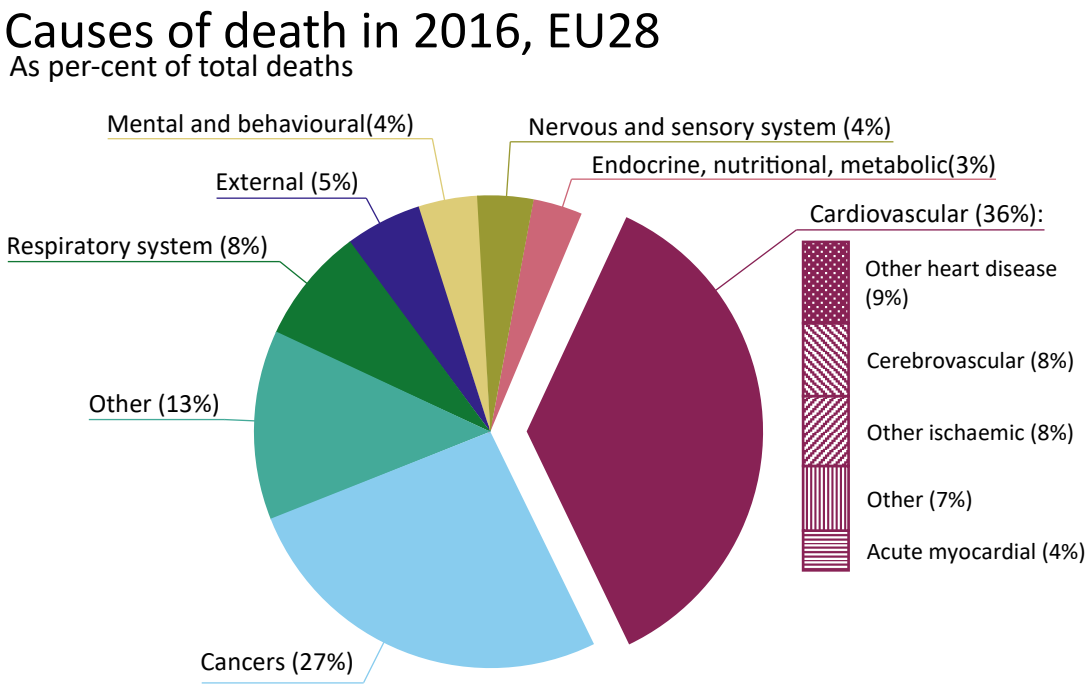


Figure 3: Causes of death in the EU28. Percentages calculated from absolute numbers for all age groups from 2016 [12]. Percentages in data labels are rounded to integers. Mental and behavioural category includes deaths caused by substance abuse.

There are two broad categories of CVDs according to the WHO: CVDs caused by atherosclerosis and others [45](**Table 2**). Atherosclerosis is a local hardening of elastic and muscular arteries, caused by lipid deposits and subsequent inflammatory processes[46], where immune cells attempt to remove the lipid deposits and infiltrate the arterial wall, leading to the development of complex lesions. The resulting atheromatous plaques are made

up of a lipid core, endothelial and smooth muscle cells, macrophages, mast cells and other lymphocytes[47], [48].

Table 2: Types of cardiovascular diseases[45]

Cause	Disease
Atherosclerosis	Ischaemic heart disease or coronary artery disease
	Cerebrovascular disease (e.g., stroke)
	Diseases of the aorta and arteries, (e.g., hypertension, peripheral vascular disease)
Other	Congenital heart disease
	Rheumatic heart disease
	Cardiomyopathies
	Cardiac arrhythmias

Table 3: Risk factors for atherosclerosis

Modifiable risk factors	Non-modifiable
Western style diet (high fat, salt, sugar)	Age
Abdominal obesity	Genetic background, (including ethnicity, somatic mutations, family history)
Insulin resistance	Coronary anatomy and branch pattern
Diabetes	
Dyslipidaemia	
Hypertension	
Smoking	

With the initial cause of atherosclerosis being the deposition of plasma lipids, mainly from low density lipoproteins (LDL), dyslipidaemia is a major risk factor. Dyslipidaemia is the dysregulation of the lipid metabolism brought on by genetic disposition or environmental and

lifestyle factors. The latter, modifiable risk factors, include smoking, diabetes (i.e., compliance with treatment regime), hypertension (e.g., brought on by stress and/or diet), and lack of exercise.

These modifiable risk factors are on the rise due to increasing prosperity in parts of the developing world, bringing about a proliferation of a fatty and salty western style diet and an increasingly sedentary lifestyle[7]. Pleasingly, increasing awareness and advances in medicine are leading to a decrease in disease burden in the western world[43].

1.3.1. Cholesterol homeostasis in health and disease

Cholesterol is an important building block for cell membranes, regulating their flexibility and fluidity. In particular, it is essential for the formation of lipid rafts, stiffer section of the membrane facilitating the formation of certain membrane protein complexes [49], [50]. Additionally, it is a vital precursor for a whole class of hormones, steroid hormones (progestogens, mineralocorticoids, glucocorticoids, androgens and oestrogens)[51]. Excess cholesterol in the form of plasma circulating LDL is a major risk factor of atherosclerosis[52].

1.3.1.1 Cholesterol sensing in the endoplasmic reticulum

Many genes involved in the import and metabolism of cholesterol are under the control of sterol regulatory elements (SREs), and a corresponding transcription factor, sterol response element binding protein 1 (SREPB-1, with two isoforms SREBP-1a and SREBP-1c), first described in 1993[53].

Low levels of cellular cholesterol are sensed by a three-part complex in the in the membrane of the endoplasmic reticulum (ER) [54](**Figure 6D**): SREBP, SREBP cleavage activating protein (SCAP)[55] and insulin induced gene proteins INSIG-1[56] and INSIG-2[57]. In low cholesterol conditions, INSIG is not bound to the complex and SREBP/SCAP are translocated to the Golgi apparatus, where SREBP undergoes two sequential cleavages and its bHLH-LZ domain is released and free to enter the nucleus[54]. Three core genes of cholesterol homeostasis are regulated by SREBP-1, *LDLR* and *PCSK9*[58], [59](**Figure 6D**), and HMG-CoA synthase [60], the first enzyme in the mevalonate pathway, in which cholesterol is synthesized.

1.3.1.2 Increasing intracellular cholesterol

Intracellular cholesterol can be increased either by import or *de novo* synthesis. Import of cholesterol is facilitated by the low-density lipoprotein receptor (LDLR), eponymous member

of a family of receptors, which in turn is part of a super-family of proteins (**Figure 5**). Members of the LDLR family share a common architecture, with an N-terminal substrate recognition domain, a single membrane spanning α-helix and a C-terminal intracellular domain which contains the endocytic NPxY motif (**Figure 7**).

LDLR abundance is regulated post-translationally both in its nascent and mature form. Proprotein convertase subtilin/kexin-9 is a protease that binds LDLR, VLDLR and ApoER2[66]–[68] and can act both systemically and locally. PCSK9 circulating in the bloodstream can bind to the EGF-like repeat A of the extracellular domain of LDLR, VLDL and ApoER2. A PCSK9 bound receptor is then internalised and translocated to the lysosome, where it is degraded. Locally, proPCSK9 in the ER can bind to the nascent LDLR (nLDLR) and route it to the to the lysosome for degradation[68](**Figure 6D,E**).

De-novo synthesis of cholesterol occurs as part of the mevalonate pathway (reviewed in [69]), which also produces many essential isoprenoids, some of which are required for the synthesis of tRNAs, HaemA, ubiquinone and post-translational modification of some proteins. The first enzyme in this pathway is 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) synthase, which synthesises HMG from acetyl-CoA and acetoacetyl-CoA. Like LDLR and PCSK9, it is under regulatory control by SREBPs[60]. The second enzyme is HMG-CoA reductase (HMGR), which reduces HMG to mevalonate. It is regulated by multiple products of the mevalonate pathway, both sterol-derived and non-sterol derived[70]. Multiple avenues of regulation allow for fine-tuned regulation of the pathway under cholesterol abundance. In presence of excess cholesterol or LDL, HMGR activity is reduced to between 10-25% of normal levels, not completely suppressed, so that products of the pathway are still available for other syntheses [71].

1.3.1.3 Cholesterol efflux

While low levels of cholesterol can be sensed directly, high levels of cholesterol are measured by the abundance of its metabolites. Other genes are under the control of the oxysterol sensing transcription factor liver X receptor (LXR, isoforms α and β), which is active as a heterodimer with retinoid X receptor (RXR)[72]. Two main genes under its regulation are *ABCA1*[73] and *MYLIP*[74](**Figure 6G**).

ATP binding cassette protein A1 (ABCA1) is a transmembrane transporter that facilitates reverse cholesterol transport (RCT). It binds circulating lipid-poor high-density lipoproteins (HDL) on its extracellular domain and shuttles excess cellular cholesterol from the plasma membrane and incorporates it into HDL (**Figure 6A**).

Excess cholesterol in cellular membranes can also be processed to cholesteryl esters by the ER bound enzyme sterol *O*-acyltransferase (SOAT, also known as acyl-Coenzyme A: cholesterol acyltransferase, ACAT)[61], [62], with two different genes showing different tissue localisation. In contrast to the other enzymes discussed here, SOAT1 is not under transcriptional regulation bound to cholesterol. Instead, it is allosterically activated by cholesterol[63]. SOAT1 protein is the dominant form in most of the body except for the small intestine. The cholesterol esterified by SOAT1 forms lipid droplets in the cytoplasm[64]. Microsomal triglyceride transfer protein (MTTP) in turn then binds nascent lipoproteins on the cell surface and loads them with the cholesteryl esters from the lipid droplets[65] (see also **Figure *4*A** for the localisation in the lipoproteins). Foam cells, which form part of atherosclerotic plaques, are showing high levels of SOAT1 expression and get their appearance from the abundance of lipid droplets in their interior[64]. It is postulated that foam cells with high lipid droplet content predispose atherosclerotic plaques to rupture. Thus, SOAT1/2 are also interesting targets for CVD.

1.3.1.4 Downregulation of cholesterol import

MYLIP, now commonly referred to as inducible degrader of the LDL receptor (IDOL) is an E3 ligase whose substrates include the nascent forms of LDLR, VLDLR and ApoER2 at the surface of the ER[74]–[78]. After poly-ubiquitination, its targets are routed to the lysosome for degradation. MYLIP is self-regulated via auto-ubiquitination, ensuring that it is depleted quickly once excess oxysterols are no longer sensed by LXR/RXR.

1.3.2. Cholesterol transport in the blood - Low density lipoprotein

There are five different classes of circulating lipoproteins in humans, each with unique compositions and densities.

- Chylomicrons (ultra-low-density lipoproteins, ULDL)
- Very low-density lipoprotein (VLDL)
- Low density lipoprotein (LDL)
- Intermediate density lipoprotein (IDL)

- High density lipoprotein (HDL)

VLDL is the precursor for LDL and IDL. It is produced by the liver and intestine, where apoB-100 is synthesised at the rough endoplasmic reticulum (ER). A single apoB-100[79] associates with lipids, phospholipids (phosphatidylethanolamine (PE), phosphatidylinositol (PI), sphingomyelin (SM), phosphatidylcholine (PC), Lyso-PC[80]), and triacylglycerides in the smooth ER. After post-translational modifications of apoB-100 in the ER it is then secreted into the bloodstream[81]. Accumulation of cholesterol esters, ancillary apoliporpoteins (ApoA-I, ApoA-II, ApoA-IV, ApoC-I, ApoC-II, ApoC-III, and ApoE[82]) from HDLs transforms them into mature VLDLs. Interaction with hepatic triglyceride lipase turns them into LDLs. Mature LDL is approximately 200 Å in diameter[83]. The shell is made up of a scaffold of ancillary apolipoproteins). The core is comprised of mainly esterified cholesterol and some triacylglycerol (**Figure 4**).

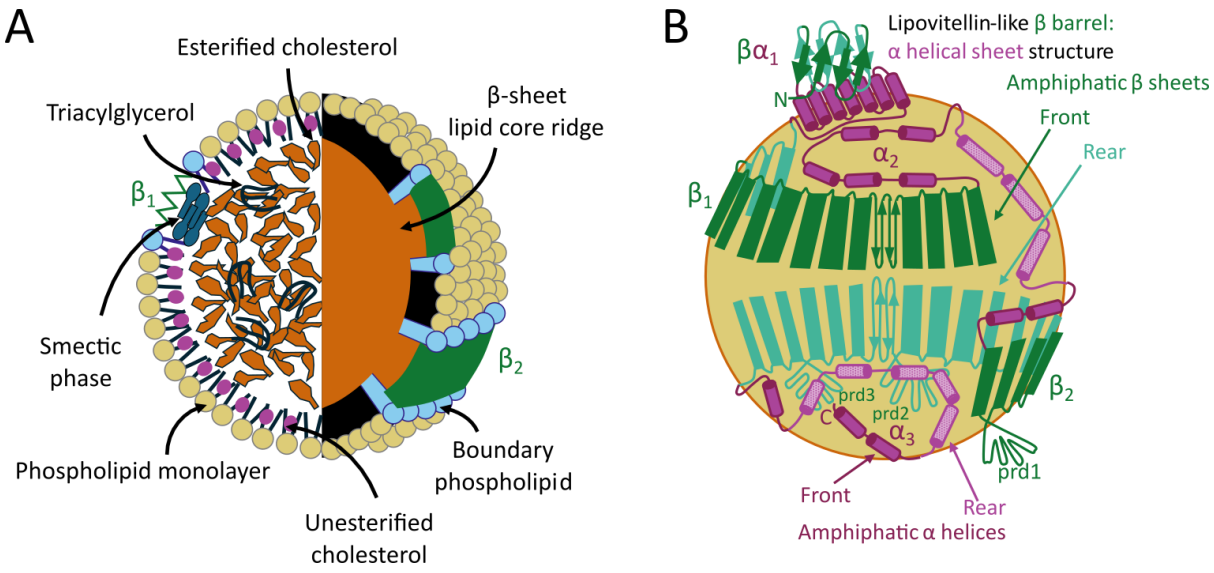


Figure 4: Three-dimensional organisation of an LDL particle. **A:** Proposed organization of lipids within the particle: ■ surface phospholipid; ■ core lipid including amphipathic β sheet-induced lipid-core ridges; ■ boundary phospholipid; ■ amphipathic β sheets. **B:** Proposed organization of apoB-100 on the LDL particle surface. β structure is shown in green, α-helical structure is shown in purple. Structures at front of sphere are shown in darker colours, structures at the back of the sphere are shown in lighter colours. Adapted from [78], Figure 11 under Creative Commons CC-BY licence[79], parts of the figure legend are direct quotes.

1.3.3. Dysregulation of cholesterol homeostasis

Dysregulation of or mutations affecting the function of any of the discussed proteins are likely to lead to dyslipidaemia, an imbalance in the body's lipid homeostasis. Of particular concern for CVDs is an excess of lipids, hyperlipidaemia, a major risk factor for atherosclerosis.

1.3.3.1 Tangier disease - ABCA1

Patients suffering from Tangier disease (TD) exhibit a massive reduction or total depletion of circulating HDL and show accumulation of cholesteryl esters in various tissues. Although patients with TD also show a reduction in LDL, this did not confer a protective effect against CVD. Loss of function mutations in ABCA1 were identified as the driver of TD[84].

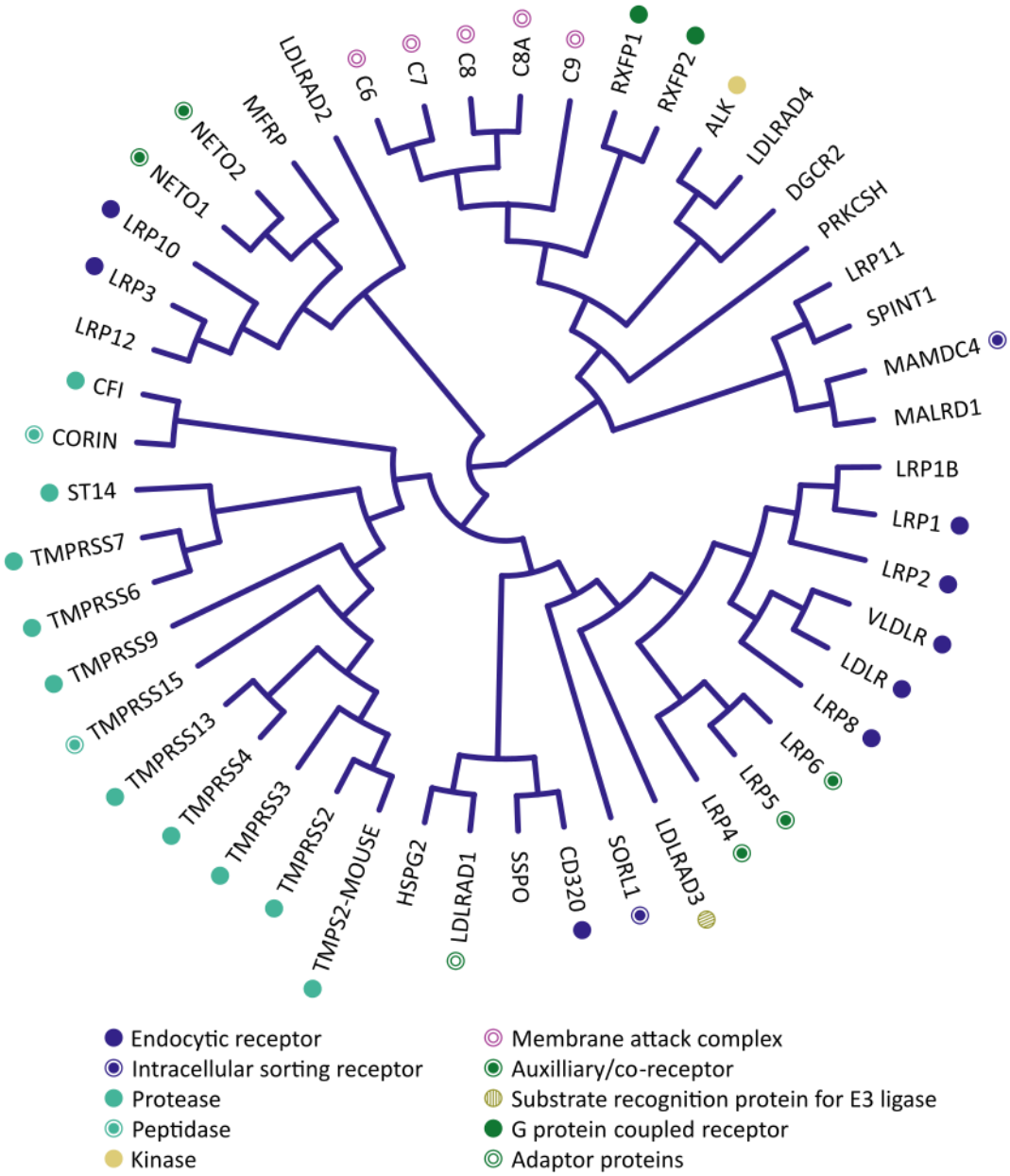


Figure 5: Phylogenetic tree of the LDL receptor superfamily. Generated with Polymorph (**https://polymorph.sgc.utoronto.ca/index.php**), accessed 2021-02-19, and annotated with information pulled from UniProt[85].

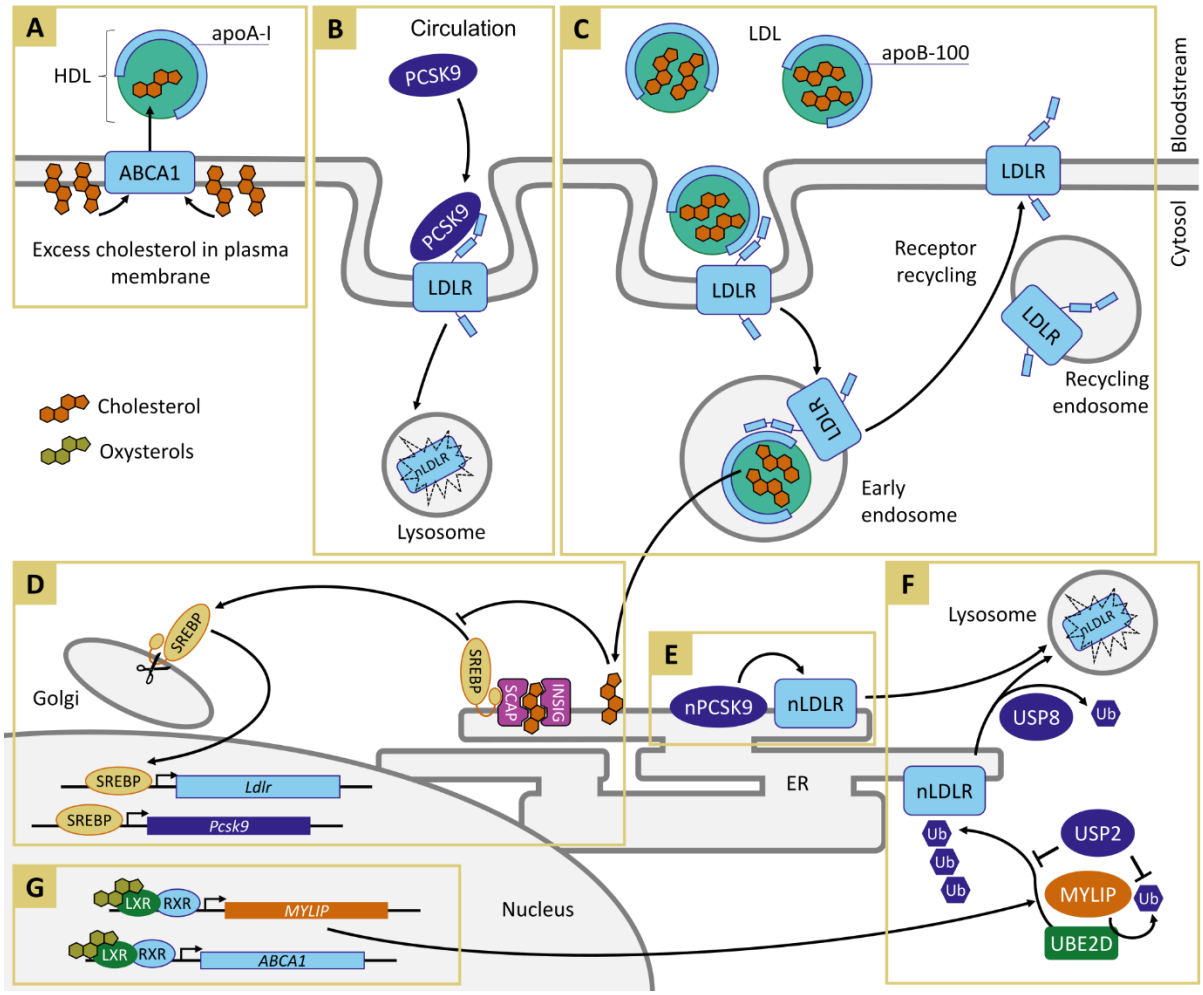


Figure 6: Schematic of LDLR and IDOL regulation. **A:** Excess cholesterol gets translocated into lipoproteins via the ABCA1 transporter[84]. **B:** LDLR is bound by circulating PCSK9, internalised and destined for lysosomal degradation **C**: LDL is imported into the cell by LDL receptor. It is then trafficked into the early endosome and eventually the endoplasmic reticulum (ER)[86] where **D:** Cholesterol levels are sensed in the ER by SREBP-SCAP-INSIG. In cholesterol depleted cells, SREBPs get translocated to the Golgi apparatus, where they undergo cleavage and the bHLH-Zip domain is released, free to enter the nucleus and bind to SREs[53], [87]. **E:** LDLR is regulated post-translation by PCSK9. Nascent PCSK9 (nPCSK9) acts on nLDLR, which gets degraded in the lysosome. **F:** nascent LDLR (nLDLR) is poly-ubiquitinated by MYLIP[76], [88]. Activated ubiquitin is supplied by UBE2D[77]. Ubiquitinated nLDLR is then trafficked to the lysosome for degradation and ubiquitin is recycled by the ubiquitin specific protease USP8. USP2 removes poly-ubiquitin chains from both nLDLR and MYLIP[89]. **G:** *MYLIP* and *ABCA1* genes under regulation by LXR/RXR.

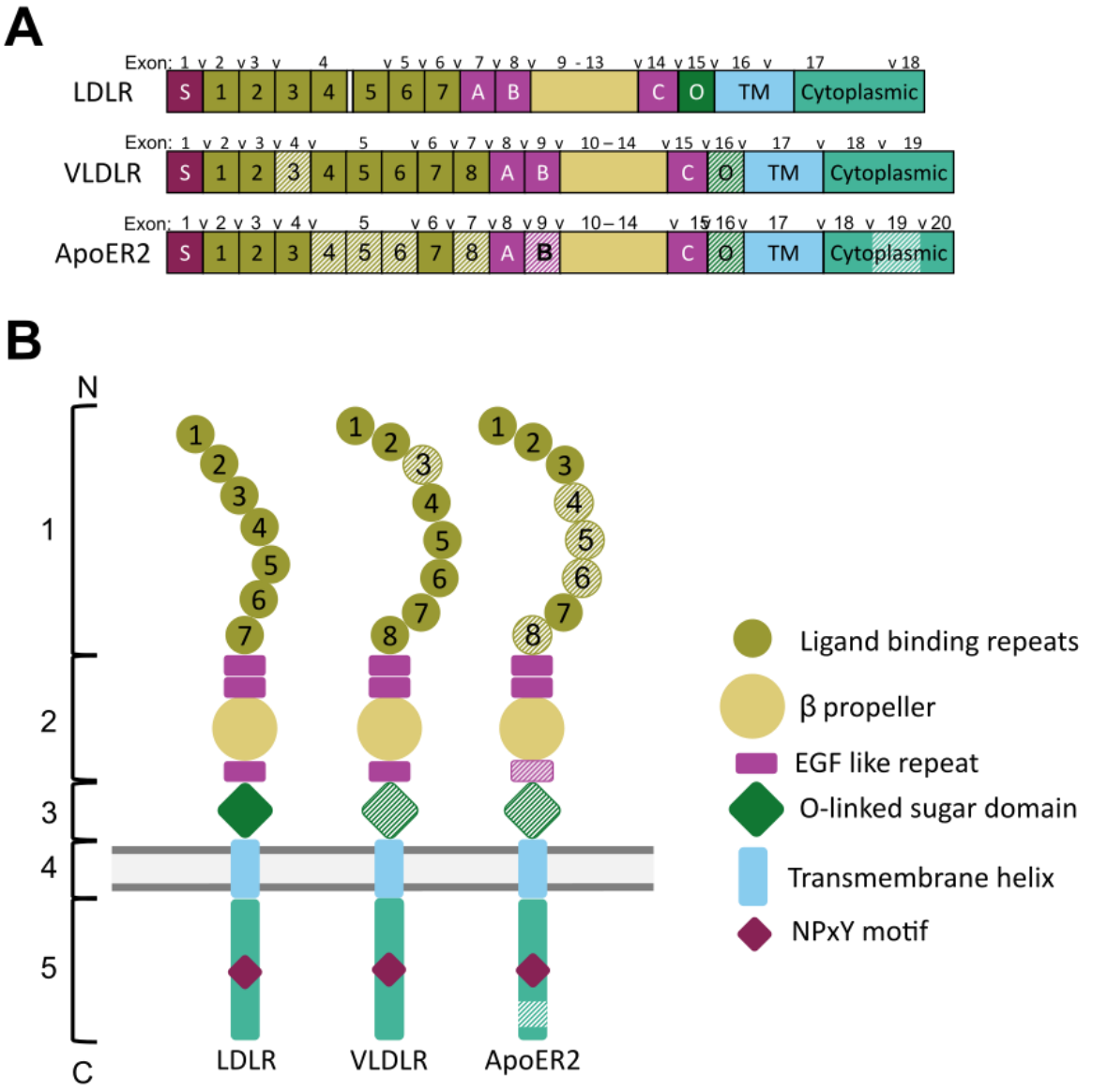


Figure 7: Structure of the LDL receptor and its two most closely related family members (see also **Figure 5**). **A:** LDLR[90], VLDLR and ApoER2[91] genes. Exon boundaries are indicated by "v"s and alternative splicing variants are indicated in patterned areas. **B:** All three are characterised by a five-part structure. **(1)** A substrate binding domain made up of leucine rich repeats. LDLR contains seven repeats, VLDLR eight and ApoER2 five. **(2)** EGF region with EGF-like repeats and a β-propeller. **(3)** O-linked glycosylation domain. **(4)** A single-span a-helical transmembrane domain. **(5)** A cytoplasmic, c-terminal tail that contains the NPxY endocytic motif (see also **Figure 9**). Domains lost in alternative splicing variants are shown as patterned areas.

1.3.3.2 Familial hypercholesterolemia – LDLR, ApoB, PCSK9 and LDLRAP1

While four genes have been identified as drivers of familial hypercholesterolemia (FH) [92], loss of function mutations in LDLR are the most common cause of FH. In 2015, 1946 mutations had been deposited in the Human Genome Mutation Database (as reviewed by Gabcova-Balaziova[93]). Even so, all four proteins are involved in the same mechanism – LDL import. Gain of function mutations of PCSK9, loss of function mutations for LDLR and LDLRAP1 (an adapter protein required for the localisation of LDLR on clathrin coated pits, without which the receptor cannot be imported[94], [95]) all lead to increased plasma cholesterol levels.

1.3.4. Prevention and current treatments for CVD

Modification of diet and increase in exercise are widely regarded to improve health outcomes in almost all diseases. American, European and British guidelines reviewed by Stewart, Manmathan and Wilkinson[96] unanimously recommend aerobic exercise and strength training. While there are some differences in dietary recommendations, reducing intake of saturated fats and sugars, as well as increasing the intake of fruit and vegetables is again highly recommended. Abstaining from smoking and alcohol consumption are also widely recommended.

Several of the pathways discussed above are currently targets for treatment of CVD:

i. Statins are a class of small molecule drugs targeting HMG-CoA reductase, thus leading to a reduction on cholesterol synthesis. Concerns around their safety have led to finely tuned recommendations for their use(recently reviewed in[97]).

ii. Hypertension can be managed by a range of drugs, thiazide-based diuretics[98], calcium channel blockers, and beta-blockers[99] being the most common categories.

iii. PCSK9 has been described relatively recently as new target. Currently, two mAbs interfering with its binding to LDLR are in the clinic, alirocumab[100] and evolocumab[101]. Additionally, an siRNA based drug, inclisiran, has recently gained FDA approval[102], [103].

1.4. Molecular targets of interest

1.4.1. Choice of molecular targets for study

With the background of ageing populations and the impact that CVD and dementia will have on health systems globally, identifying targets that can modify these diseases at an early stage is critically important. On a practical level, an ideal target is one that is well characterised and can be effectively studied *in vitro* in biophysical, biochemical, and cell-based assays. Once the mechanism of action has been elucidated, it must be amenable to high throughput screening to identify agents (small molecules, antibodies) that can effectively modify the function of the target to pivot away from the disease phenotype.

1.4.2. Inducible degrader of the low-density lipoprotein receptor (IDOL)

MYLIP was first discovered as an E3 ligase targeting the myosin regulatory light chain B (MRLC-B), affecting cell motility and cytoskeletal reactions, especially neurite outgrowth[104], [105], hence it was dubbed myosin regulatory light chain interacting protein (MIR) at the time. Northern blot analysis of various tissues showed a wide distribution, both in adult and foetal tissue[104]. It was later shown to also target members of the LDL receptor family and was given the name inducible degrader of the low-density lipoprotein receptor (IDOL). It is worth noting here that the use of two different names, MYLIP and IDOL, has a potentially confounding effect on research. A quick PubMed search for IDOL and MYLIP revealed that IDOL is preferentially used in CVD and AD research, while MYLIP is preferentially used in cancer research. For this work, it will be referred to as MYLIP, as this is still the recognised gene name, and it was registered in the departmental database following this convention.

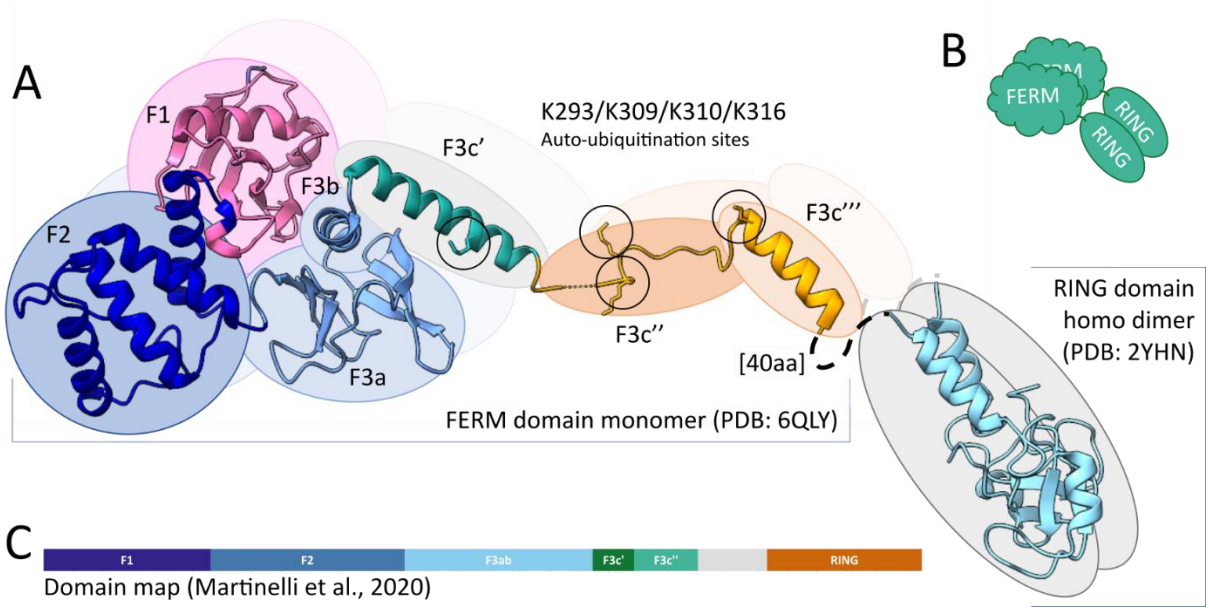


Figure 8: Structure of MYLIP. **A:** Composite image of published crystal structures of MYLIP FERM domain (PDB: 6QLY) and RING domain homo dimer (PDBL 2YHN). Domains are highlighted as ellipses. Based on the structure, F3c'' according to Martinelli et al. [106] has been further subdivided into the non-structured F3c'' and helical F3c'''. Auto-ubiquitination sites K293, K309, K310 and K316 are highlighted. The second FERM domain monomer is indicated as pale ellipses in the background. **B:** Cartoon representation of MYLIP homo dimer. **C:** Domain map according to Martinelli et al.[106].

Structurally, it is now known to consist of an N-terminal substrate recognising FERM (4.1, ezrin, radixin, moesin) domain[107] and a C-terminal RING domain, which facilitates dimerisation and E3 ligase activity[108], [109], including auto-ubiquitination of MYLIP (auto-ubiquitination sites K293, KK309, K316, K320[78]), which is degraded in the proteasome.

1.4.2.1 Cell motility

Further studies into the connection between MYLIP and the cytoskeleton and cell motility identified a MYLIP antagonist, MSAP (MIR-interacting saposin-like protein, also known as protein canopy homolog 2, CNPY2). Stimulation of neurite outgrowth by overexpression of MSAP could be counteracted by expression of MYLIP[108]. Importantly, the full length wild type was required. Nagano et al. then further probed MYLIP's role and looked at whether there was a connection to platelet-derived growth factor (PDGF), a known potent mitogen and inductor of fibroblast motility[110]. They treated Swiss 3T3 fibroblasts with PDGF in the culture medium and demonstrated upregulation of *MYLIP*[110] expression and increased ubiquitination and proteasomal degradation of MRLC. This induces a reorganisation of actin structures and a motile phenotype. Suppression of *MYLIP* with siRNA showed a signification

reduction of cell migration and wound healing in scratch assays. The mRNA binding protein hnRNP-K was identified as the mediator between PDGF treatment and increased MYLIP expression: phosphorylated hnRNP-K was shown to bind to MYLIP mRNA. Nagano and colleagues hypothesised that this stabilises the MYLIP mRNA, resulting in an increase in protein levels. Additional treatment of PDGF stimulated cells with the Src kinase family inhibitor SU6656[111] prevented the phosphorylation of hnRNP-K and subsequently the upregulation of *MYLIP*.

1.4.2.2 Evidence for involvement of MYLIP in cholesterol regulation

In 2009, Zelcer and colleagues[74] investigated the link between the heterodimeric liver X receptor transcription factors (LXRαβ) and the low-density lipoprotein receptor (LDLR). Treating several cell lines with synthetic LXR agonists (GW3965[112] or T0901317/T1317[113]), they were able to show that activation of LXR resulted in a decrease of LDLR on the cell surface and its redistribution to intracellular compartments independent of intracellular sterol levels. However, they observed only in a moderate reduction in LDLR mRNA levels. Furthermore, the agonists lead to an increased level of the transporter ABCA1. Conversely, treatment of *Lxrαβ*$^{-/-}$ mouse embryonic fibroblasts (MEFs) with the LXR agonists resulted in no effect on LDLR levels. Analysing transcriptional profiles, they then identified MYLIP as a potential mediator between LXR activation and LDLR degradation and showed that LXR agonist treatment induced expression in primary and established cell lines (primary hepatocytes, primary macrophages, MEFs and HepG2[114]). Similarly, *MYLIP* and *ABCA1* expression was induced by LDL in cell culture medium. Ectopic expression of full-length wild type MYLIP and catalytically inactive MYLIP C387A[109] showed that redistribution of LDLR from the plasma membrane to intracellular compartments was dependent on the E3 ligase activity of MYLIP. Zelcer and colleagues then investigated the effect of MYLIP on other members of the LDLR superfamily and found that ApoER2 was “marginally” affected, and no effect could be shown for LRP1-GFP (aka APOER), β-amyloid precursor protein-GFP (aka APP), Lrp4, or SorLA. Interestingly, Zelcer and colleagues reported no effect on MRLC (their supplemental figure 4), in contradiction to previous findings, where ubiquitination of MRLC in presence of MYLIP was shown[108], [110].

Whereas Zelcer et al. immunoblotted against their proteins of interest with whole cell lysates of HEK293 ectopically expressing MYLIP wt and MYLIP C387A and compared band intensity in

presence of endogenic, wt or mutant MYLIP, Bornhauser et al. and Nagano et al. pulled down their proteins of interest directly using monoclonal antibodies on protein G Sepharose beads, followed by immunoblotting against ubiquitin.

Other results from previous studies were, however, confirmed: MYLIP possesses auto-ubiquitination activity[109] and is degraded via the proteasome[110]; knockdown of MYLIP expression with shRNA reduces its effect on its target proteins (as measured by an increase of LDL uptake by Zelcer at al.) and the LXR agonist mediated degradation of its target was attenuated by MYLIP shRNA. Furthermore, the levels of LDLR mRNA were not affected, as with MRLC, and the C387A mutation abolished the ubiquitination of MYLIPs target. Since LDLR is a membrane bound protein, the route of degradation is likely to be different to the cytoplasmic MRLC. Zelcer et al describe that MYLIP targets immature LDLR in the endoplasmic reticulum and poly-ubiquitinated LDLR is trafficked to the lysosome for degradation. In vivo effects were also investigated, and it was found that ectopic expression of *MYLIP* increased plasma levels of total and un-esterified cholesterol in $LDLR^{(+/+)}$ mice, but not in not $LDLR^{(-/-)}$ mice.

Having demonstrated that LDLR is not the only target of MYLIP in the LDLR superfamily, Hong and colleagues then turned their attention to VLDLR and ApoER2[76] and showed that the LXR agonists (GW3695 and T0901317/T1317) affect VLDLR in the same fashion as LDLR, leading to a decrease in protein levels while mRNA levels remained unaffected. Therefore, the substrate specificity of MYLIP overlaps with that of the proprotein convertase subtilisin/kexin type 9 (PCSK9) [67], [68]. Ectopic expression of wild type and catalytically inactive MYLIPC387A showed the same reduction and lack thereof, respectively, as for LDLR.

Conversely, the depletion of endogenous oxysterol LXR antagonists by overexpression of the oxysterol sulfotransferase Sult2b1 leads to increase in VLDLR protein levels without affecting gene expression[74]. This effect could be reversed by treatment of Sult2b1 expressing cells with GW3695. Changes in expression of *MYLIP* and *ABCA1* in GW3695 treated mice follows *in vitro* pattern, with a decrease in VLDLR protein levels but unchanged gene expression.

Having observed two bands in immunoblots against VLDLR and ApoER2, Hong and colleagues hypothesised that these represent immature and fully glycosylated versions of VLDLR and ApoER2 (Zelcer and colleagues had also previously observed two bands for LDLR and also attributed these to the immature and glycosylated form[74]). To determine where MYLIP

effects its targets and how they are degraded, they conducted pulse-chase experiments with 35S cysteine and 35S methionine in HEK293 cells to interrogate the change of VLDLR protein levels over time. In cells with ectopically expressed MYLIP, the mature receptor could not be detected in significant quantities even after 6 hours. In the control cells the mature form started to appear at or before 2 hours. To determine the route of degradation, cells were treated with the proteasome inhibitor carbobenzoxy-Leu-Leu-leucinal (MG132, reviewed by Lee and Goldberg[115]) and found no effect. Treatment with ammonium chloride and chloroquine, which impair the activity of the lysosome, resulted in maintenance of VLDLR levels after concomitant treatment with GW3695 in agreement with previously published result by same group[74].

After identifying the conserved lysine in the cytoplasmic domain at which VLDLR is ubiquitinated, they investigated the related receptors LRP1B and epidermal growth factor receptor (EGFR) which also carries the same sequence motif as LDLR, VLDLR and ApoER2 (NPxY, **Figure 9**). However, they found no decrease in LRP1B and EGFR levels. Similarly, no decrease of the related APP was observed. Grafting the intracellular domain of LDLR on APP did, however, lead to a decrease, showing that the entire motif is essential for MYLIPs activity on these receptors.

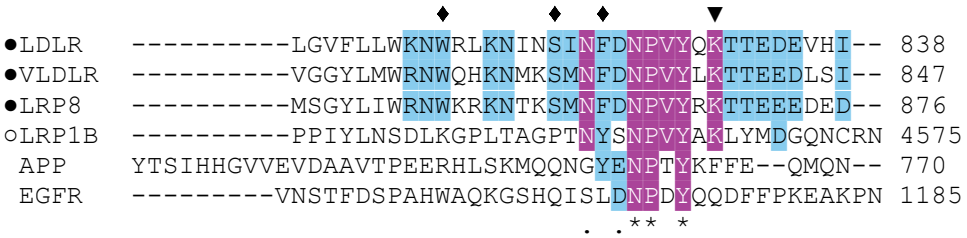


Figure 9: Sequence alignment of LDLR superfamily members and EGFR investigated as substrate for MYLIP. ▼ denotes conserved ubiquitination site. ♦ Denotes residues identified by Calkin et al.[78] as required for binding. Shaded residues denote conservation according to Hong et al., 2010[76]. ● circles are members of the LDLR superfamily, full circles are substrates of MYLIP. Sequence alignment performed on UniProt.org with ClustalO.

1.4.2.3 Mechanism of action

Ubiquitin is a small protein that is found in most tissues and is involved in post-translational regulation of protein abundance (**Figure 10**). Ubiquitin is first activated by ubiquitin activating enzymes (E1) in an ATP-Mg^{2+}-dependent fashion, where it is bound to the E1 via a thioester. In two subsequent transthioesterification reactions, the activated ubiquitin is transferred from the E1 first to ubiquitin conjugating enzymes (E2), from where it is either directly transferred to a ubiquitin protein ligase (E3, or E3 ligase), in the case of HECT/RBR type E3 ligases, or the E2 forms a complex with RING type E3 ligases. The ubiquitin is then conjugated to its target protein in an isopeptide bond on a lysine. After the initial ubiquitination, poly-ubiquitin chains are formed, the shape of which depends on the specific E3 ligases (as reviewed in [116]). These ubiquitin marks can also be removed by deubiquitinating enzymes (DUBs) in an additional layer of regulation. Soluble ubiquitinated proteins are then degraded via the proteasome while membrane proteins are recycled either in the proteasome or shuttled to lysosomes via sorting/recycling endosomes (as reviewed in [117]).

The really interesting new gene (RING) domain, which facilitates binding to E2 enzymes and also contains the catalytically active cysteine, was first described by Freemont and colleagues in 1991[118] and its structure solved via NMR three years later[119], while the exact molecular mechanism by which RING type E3 ligases transfer ubiquitin from E2s was solved by Branigan et al in 2020[120]. Zhang et al. then identified the specific E2 partner enzymes of MYLIP[77]. They employed an auto-ubiquitination assay and screened a panel of 19 E2s that are known to work with RING type E3 ligases. In an *in vitro* system generated from *E. coli* lysates (E2 expression system, without affinity chromatography steps), MYLIP from HEK293 lysates (N-terminally tagged with 3x FLAG and Strep-tag) was incubated with recombinant UBE1, hemagglutinin (HA) tagged ubiquitin, and ATP. After the reaction, Streptactin beads were used to pull out MYLIP from the reaction mix and ubiquitinated product was detected via SDS-PAGE followed by Western blot for HA tag. The UBE2D family (UBED2D1-4) were the only E2s that showed activity. MYLIP bearing the catalytically inactive C387A mutant was not detected.

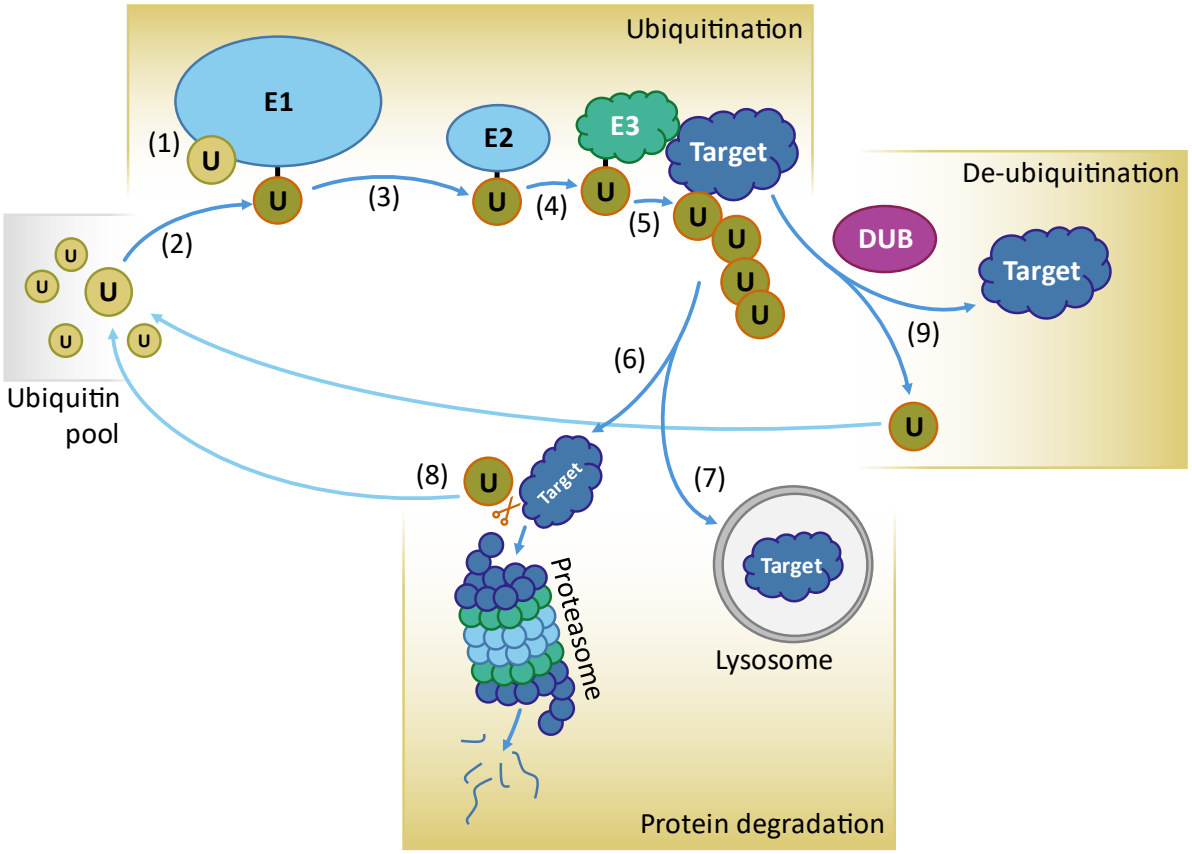


Figure 10: Ubiquitination mediated protein degradation, specifically for RING type E3 ligases. **(1)** One unit of ubiquitin binds to the ubiquitin activating enzyme (E1) at an allosteric site. **(2)** A second unit of ubiquitin binds to and gets activated. **(3)** The ubiquitin gets transferred to the ubiquitin conjugating enzyme (E2) in a first transthioesterification reaction. **(4)** The E2 then binds to the ubiquitn ligase (E3) and transfers the ubiquitin in a second transthioesterification reaction. The E3 enzyme then covalently binds the ubiquitin to the target protein. **(6)** Soluble proteins are degraded via the proteasome, **(7)** membrane bound proteins are shuttled to the lysosome via sorting/recycling endosomes. **(8)** Ubiquitin is cleaved from the target protein at the entrance to the proteasome and re-enters the ubiquitin pool of the cell[121]. Non-degradative ubiquitinations (often in the form of mono-ubiquitination or poly-ubiquitination with different branching points) are not shown.

1.4.2.4 Evidence for involvement if MYLIP in Alzheimer's disease

Given strong the strong association of ApoE and in Alzheimer's disease, Choi and colleagues investigated the role of lipoprotein receptors in the brain[122]. Working with transgenic mice, they recapitulated key findings *in vivo*: LXR agonists increased *MYLIP* mRNA levels and led to a reduction of LDLR levels; LDLR levels are increased in the brains of *MYLIP*$^{-/-}$ mice, with the

levels of LDLR being intermediate in *MYLIP*$^{+/-}$ mice compared to *MYLIP*$^{-/-}$ and *MYLIP*$^{+/+}$ mice. Specifically, *MYLIP* mRNA was detected in microglia. They then bred *MYLIP*$^{+/+}$, *MYLIP*$^{+/-}$ and *MYLIP*$^{-/-}$ mice carrying the *APPswe* and *PSEN1dE9* transgenes (*APP/PS1*), which cause increased Aβ depositions in the mouse brain[123] and found a significant correlation between the *MYLIP* genotype and Aβ40 and Aβ42 deposition, as measured by thioflavin S staining and immunostaining of brain tissue sections. Intriguingly, the abundance of microglia directly correlated with the *MYLIP* genotype, but the levels of microglial markers (*Cd45*, *Cd68* and *Ym1*) and the ability of microglia to be stimulated with lipopolysaccharides remained unchanged. This indicates that the activity of microglia in APP/PS1 mice is not influenced by MYLIP. In line with previous studies[25], [124], the authors were able to show that increasing LDLR levels via siRNA knockdown of MYLIP increased extracellular Aβ clearance. In a follow up study to [125], Gao and colleagues used a MYLIP anti-sense oligonucleotide (ASO) to knock down MYLIP expression in *APP/PS1* mice[126]. They were able to demonstrate significant Aβ plaque reduction, reduction of total Aβ species and ApoE levels. They attributed this to increased phagocytic activity of brain microglia and theorised that increase of ApoER2 levels on neurons might also have a protective effect, in line with previous findings[127].

1.4.2.5 Target rationale

Given its role in two areas of ageing related disease, MYLIP presents itself as a tempting target for drug discovery efforts. Indeed, some mutations of MYLIP have been identified as playing a driving role in health and disease. Several mutations in humans and knockout studies in mice have shown a beneficial effect on LDL cholesterol (LDL-C) levels:

- p.Arg266X results in a complete loss of function and associated lowered levels of LDL-C[128]
- G51S impairs the homo-dimerisation of MYLIP and thus its auto-ubiquitination, increasing its half-life from circa 2 hours to about 6 hours. This results in increased LDLR degradation and increased LDL-C levels[129]
- Conflicting findings exist for N342S and the associated SNP rs3757354, with some studies showing an association with lowered LDL-C levels[130], others showing none and one Chinese study showing variability between two different ethnic groups.
- Total inactivation of MYLIP attenuates diet-induced obesity in mice[131]

Furthermore, Leitch et al. demonstrated inhibition of MYLIP function with a cyclic peptide. They were able to show a dose dependent increase in LDLR abundance after treatment of HepG2 cells with the cyclic peptide[132]. In a similar vein, the mouse knockout studies discussed above, and the MYLIP-ApoER2-ApoE-Aβ-plaque axis present a strong case for drug development efforts.

1.4.2.6 Risk profile and other substrates

When targeting a protein with an inhibitor, it is desirable to have as few off-target effects (effects not directly related to the function of the protein or the pathway we intend to inhibit) as possible to limit the potential for unacceptable side effects.

The three most well characterized targets of MYLIP are the lipoprotein receptors LDLR, VLDLR and ApoER2. These are all involved in the disease mechanisms of interest in this study and indeed, the studies investigating MYLIP loss-of-function mutations have not given any indication as to a deleterious effect on other aspects of health. While ApoER2 is playing an important role in the developing brain, it is unlikely that this would be of consequence in the prospective patient cohort.

Another recently discovered substrate of MYLIP is the transcription factor androgen receptor(AR)[133]. Ito and colleagues investigated the interplay between AR and the known MYLIP antagonist CNPY2/MSAP in pancreatic cancer cell lines and found that CNPY2/MSAP knockdown lead to a decrease in AR abundance. They theorised that MYLIP is responsible and were able to show with co-immunoprecipitation and *in vivo* ubiquitination assays that MYLIP does indeed ubiquitinate AR. Consistent with previous findings that CNPY2/MSAP inhibits MYLIP mediated ubiquitination of MRLC[108], CNPY/MSAP overexpression rescued AR abundance.

In the same vein, Zhao and colleagues identified MYLIP as a target for miR-19b in breast cancer after bioinformatical studies revealed an association between the two[134]. They were able to show a direct link between miR-19b expression and *MYLIP* repression. While other genes are affected by miR-19b, MYLIP's role in regulating cell motility should not be underestimated and warrants further investigation.

Furthermore, SNP rs3757354 of MYLIP has been associated with early onset preeclampsia[135]. While this is a severe condition, it is again not relevant to the patient

cohort dealing with CVD and AD as age- and ageing-related diseases. Should MYLIP be considered as a target for familial hypercholesterolaemia in younger patients, the pre-eclampsia risk and effects on neuronal development need to be considered.

1.4.2.7 Summary

Several loss-of-function mutations and knockout studies show a protective effect against atherosclerosis as well as a neuroprotective effect. Although MYLIP plays a role in neuronal development, learning and neuroplasticity, these factors are most likely not relevant for the main patient cohort (65 years of age and above). Given its role in both CVD and AD, MYLIP/IDOL appears to be an attractive target worth pursuing.

To identify tool compounds and novel starting points for pharmaceutical intervention, we have decided to take a target-based approach. Since the target is well known and there is ample evidence for the mechanism of target recognition, this work is setting out to develop protein-peptide-interaction based assays to screen small molecule libraries for inhibitors of the binding of the LDLR tail to the FERM domain of MYLIP. The CMD has considerable expertise in this area[136]–[140] and infrastructure dedicated to it, which will aid in the effective prosecution of this target.

1.4.3. Assays

The following section gives a brief overview of assay techniques that are suitable for the discovery of small molecule inhibitors of MYLIP.

1.4.3.1 Protein-protein interaction (PPI) assays

In these assays, the protein of interest (POI) (or part thereof, e.g., a substrate recognition domain) and the disease relevant interaction partner (or part thereof, e.g., a peptide) are usually tagged with biotin or an affinity tag that allow for the use of detection reagents that generate a signal when the interaction partners are bound to each other (**Figure 11**). These assays are amenable to high-throughput screening using microtiter plates (384 or 1536 wells) and plate readers. To achieve high throughput, automated liquid handling is essential, especially for compound dispensing. Overall, they require considerable initial capital investment. Reagents are often universal and therefore comparatively cheap, while proteins can often be easily produced in *E. coli* and purified. The most common assay technologies are AlphaScreen and homogenous time resolved FRET (Förster resonance energy transfer; HTRF).

The biggest limitation of these systems is their reductive nature – cellular context is entirely missing and cell permeability, viability, and selectivity studies to ensure no off-target effects are necessary. Enzymes can be studied as well with fluorescence polarisation or HTRF if the substrate is amenable to conjugation with an appropriate fluorophore[141].

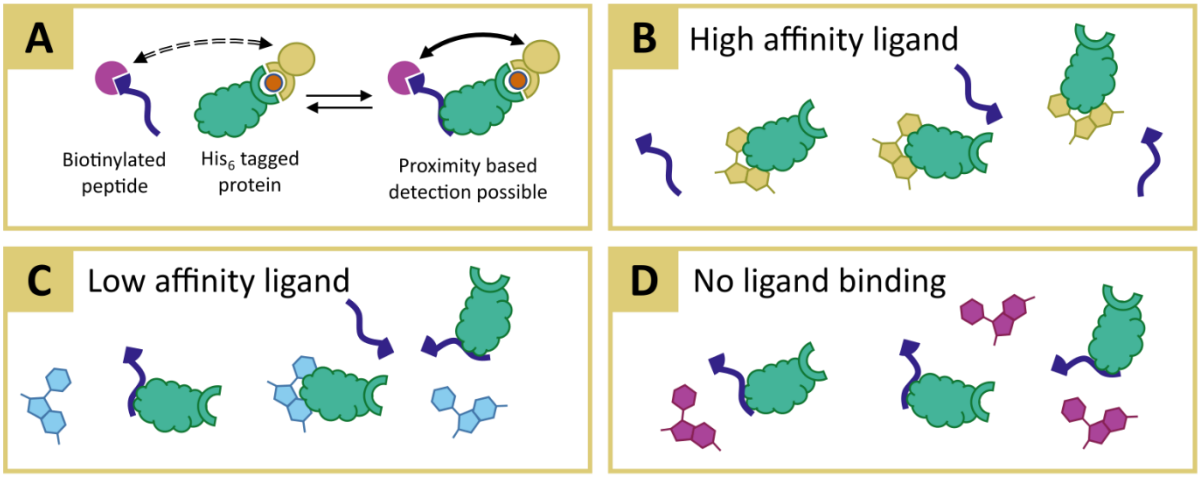


Figure 11: Screening for inhibitors of protein-peptide-interaction with proximity-based assays. **A:** Equilibrium of bound and unbound states of the protein-peptide pair. In bound state, detection is possible with methods such as HTRF and AlphaScreen. Detection moieties are represented as full circles interacting with biotin on peptide and affinity tag on protein. **B:** Binding is completely abolished in presence of a high affinity ligand. **C, D:** Reduction in binding and therefore signal is proportional to the affinity of the ligand.

1.4.3.2 Enzymatic assays

These assays rely on the direct detectability of reaction products, e.g., via fluorescence, luminescence, or mass spectrometry. Assays can also be set up where the product is converted to something that can be detected. In our lab, Promega's proprietary ADP-Glo™, AMP-Glo™ and UDP-GLo™ assays are used to detect ADP, AMP and UDP, respectively, as end products of enzymatic reactions.

The chief limitation of these assays is that the product or substrate must be amenable to be converted. If assay kits are readily available from a supplier, however, they are relatively easy to establish and amenable to high throughput screening, provided the necessary infrastructure is in place, as discussed above.

1.4.3.3 Biophysical assays

Several methods are available for investigating the binding of ligands to proteins, be they small molecules, peptides, or other proteins. Isothermal titration calorimetry and biolayer interferometry are the most used methods and discussed below.

Isothermal titration calorimetry (ITC)[142] is a label-free method to determine the dissociation constant of an interaction. Here, binding partner one is titrated into binding partner two in a calorimeter and the heat change upon binding is measured. Because neither of the interaction partners is immobilised on a surface, possible binding of the free partner to the sensor surface or the orientation of the immobilised partner on the sensor surface are not of concern and don't need to be controlled for. Due to the absence of these interfering factors, ITC is often considered a gold standard method. Despite this, it is not without limitations. Heat of dilution of the titrated component and solvent/buffer mismatch between the two partners need to be corrected for with control experiments, i.e., titrating of one partner in absence of the other. Which component is titrated into the sample cell depends on the solubility and solvent use. To keep the mismatch minimal, small molecules are often used in the larger volume (i.e., the sample cell) and the protein in the titration syringe supplemented with an equal amount of solvent. Otherwise, a high concentration of solvent would generate heat of dilution that could obscure the binding events.

Biolayer Interferometry (BLI) is another label free technology available in the department. This technology uses disposable fibre-optic sensors to measure biomolecular interactions. It measures white light reflected from the internal reference surface and the ligand-binding surface, creating an interference pattern with peaks and troughs along the visible light spectrum. Binding to the surface of the sensor increases its thickness and distance from the reference surface and therefore a shift in the peaks in the interference pattern. This shift is measured in nanometres (nm) and displayed as a sensogram. At saturation, the sensogram reaches a plateau, the height of which is dependent on the affinity of the analyte to the immobilised ligand, the concentration of the analyte, and the density of the immobilised ligand on the sensor surface.

Experiments are carried out by dipping the sensor(s) into different concentrations of analyte on 96 or 384 well microtiter plates. The dissociation constant K_D can be measured by fitting

an appropriate model directly to the sensogram (e.g., a one-phase association) or by fitting the model to the plateaus reached at different analyte concentrations.

1.5. An overview of data analysis in high throughput screening

Analysing data and maintaining research integrity is just as vital as designing and carrying out sound experiments. Unfortunately, conversations with colleagues over the years at conferences, meetings and tea breaks in the department have shown that many academic labs are using a mixture of home-grown templates and pipelines of varying quality. These are often handed down from researcher to researcher and the original author has often left the lab some time ago, leading to knowledge gaps where researchers are essentially using a black-box approach. The staff turnover rate in academia leads to a regular loss of knowledge about these workflows. Between the academic years 2020/21 and 2021/2022, the number of academic staff in the UK increased from 224,530 to 233,930, with 45,800 starters and 36,475 leavers[143]. On the other hand, researchers could be individually creating their own templates and pipelines, leading to a duplication of efforts and lost time. I will be discussing the historical situation in our department as a case study and then lay out in what ways the issues can be addressed.

1.5.1. The SGC/CMD Biophysical and Biochemical Screening Group – a case study

The Biophysical and Biochemical Screening Group (BBG) was originally established as part of the Structural Genomics Consortium (SGC). The SGC is a private-public partnership that is committed to accelerating drug discovery via open science and sharing all its data freely with the scientific community[144]. To this end, databases were established at each site to store results from a range of different assays (e.g. protein-protein interaction assays such as AlphaScreen and HTRF, differential scanning fluorimetry[145], isothermal titration calorimetry) and make the results accessible between sites. When developing new assays, members of the department also developed analysis workflows, often using a variety of programs and templates and while our group had a relatively low level of turnover, the developers of the templates and workflows all eventually left the department.

The workflow used for analysing raw data from AlphaScreen based protein-peptide interaction assays to determine IC_{50} values for small molecule inhibitors of protein-protein interaction will be discussed as an example: First, a KNIME[146] pipeline was built to process

liquid handler transfer reports to extract concentrations and sample IDs. A MS Excel template was then used to accept the output from the KNIME pipeline and to re-arrange and normalise raw data from microtiter plate readers in preparation for curve fitting in a GraphPad PRISM template. This template would perform a fit with constraints for the top and bottom of the curve as well as a free fit. After obtaining the results of the curve fits in PRISM and pasting them back into the MS Excel template, the users would then have to sight the data and decide which fit to use, depending on the completeness of the dose response curve. The MS Excel template then transposes the results from the curve fitting and the normalised data in a results table to be manually uploaded to our in-house database. This approach was limited to processing one plate at a time, leading to a lot of time spent copying and pasting data, which can be complicated by user errors and memory errors.

1.5.2. Identifying requirements and choosing an environment

The Python distribution Anaconda[147] is part of the standard image used by our IT department to set up new computers. This distribution comes with many useful libraries pre-installed and is therefore an ideal platform to develop assay data analysis workflows for our department. Since these libraries are open-source software, licensing costs and IP rights are not a factor that needs consideration.

Python offers a wide selection of libraries and modules that allow the readout of raw data files and liquid handler transfer reports, normalisation of raw data, processing of data to associate data points with sample IDs and concentrations, fitting of data points to models, and the display of results in the form of tables and interactive graphs. Additionally, user interface libraries make it possible to create a user experience that will feel familiar to users of MS Windows PCs, lowering the barrier for entry and increasing the adoption rate. This will also cut down on training time, which can be crucial to keep drug discovery programs running smoothly despite staff turnover. This also facilitates cross training of users, again protecting projects against disruption due to staff turnover or illness.

1.6. Summary and aims

1.6.1. Alzheimer's and cardiovascular disease

While MYLIP knockout cell lines and animal models exist and siRNA knockdown strategies have been developed, potent and selective small molecule inhibitors of its function are still

missing from the scientific toolbox. In addition to being a precursor to therapeutics, these molecules serve as a valuable tool to study the function of a protein *in vitro* and *in vivo* without the side effect of disturbing the context in which the protein exists. If an enzyme of interest also fulfils a role as a scaffolding protein in a larger complex, it is difficult to ascertain whether the effect of a knockout/knockdown is due to the abrogation of its enzymatic function alone or whether the function of the entire protein complex has been compromised. This study aims to develop methods to screen small molecule libraries for inhibitors of MYLIP, and to apply these assays to screen libraries in the portfolio of the department with the goal to discover a first in class inhibitor of MYLIP. The specific milestones required to achieve the overall goal are laid out below:

1.6.1.1 Protein construct design and production

In the first instance, protein constructs exploring the FERM domain of MYLIP need to be designed for the development of a protein-protein interaction assay and biophysical assays based on the replication of the findings of Calkin et al. [78]. For enzymatic assays, catalytically active, full-length protein needs to be produced. For these constructs, efficient protocols for expression and purification must be developed to facilitate the production of amounts of protein of sufficient purity and yield to support the development and conducting of assays.

1.6.1.2 Protein-peptide binding assay

Once proteins have been produced, the results of the fluorescence polarisation experiments of Calkin et al.[78] must be reproduced with either the originally described technology (fluorescence polarisation) or other available technologies (1.4.3.1). Based on this, the conditions of the assay need to be optimised with regards to data quality (signal to background and signal to noise ratios), reagent consumption, and ease of use (e.g. a minimal number of reagent addition and incubation steps) to make it suitable for a high-throughput assay. This can include procurement of further LDL family receptor peptides to explore the impact of the boundaries on the affinity to the MYLIP FERM domain and to identify other members of the LDLR family that might be potential substrates for IDOL.

1.6.1.3 Enzymatic activity assay

To complement the protein-peptide binding assay, an enzymatic assay utilising the full ubiquitination cascade around MYLIP/LDLR can be established. This can either be optimised

for high throughput screening or used to further validate and characterise any hits identified via the protein-peptide binding assay.

1.6.1.4 Pilot screen

Once an assay has been developed, a small pilot screen should be conducted to assess the viability of it for high throughput screening. The main criteria for this would be whether the Z' and Z'_{robust} values [148], [149] for individual plates are satisfactory (>0.5) and whether signal levels are consistent over the batch of assay plates being screened. If the signal level deteriorates too quickly while a large batch of plates is being handled, the system needs to be stabilised or the screen broken up into smaller batches.

Furthermore, the hit-rate of the assay, i.e. the number of true or apparent inhibitors of the protein-peptide interaction or enzymatic reaction, cannot be too high. If the system is too easily disrupted, the number of false positive hits will make deconvolution of true inhibitors of the MYLIP-LDLR interaction and compounds that disrupt the underlying assay technology prohibitive in terms of time and reagent consumption. Generally, a hit rate of 0.5% – 1% is considered good.

1.6.2. Software development

The goal of this study is to develop a freely available, adaptable data analysis platform for high throughput screening. This software will enable academic labs to develop, establish and maintain analysis workflows for routine assays. While an understanding of the underlying mathematical and statistical theory is critical (i.e., choice of model to fit experimental data to, and which measures to use to assess data quality and significance of results), a lack of knowledge of how to implement these should not stand in the way of high throughput data analysis. Use of such a software will also protect groups against the inevitable loss of knowledge that occurs with the high turnover of staff in academia. Database integration will ensure that results can be accurately and reliably recorded and easily shared with collaborators.

Complete solutions do exist, but they are often cost prohibitive, running into thousands of pounds per workstation per year. An in-house developed solution based on an open-source software ecosystem can lead to significant savings, which can be reinvested in new assay technologies and/or additional staff to increase the capacity of the department.

Since Python is an interpreted language as opposed to a compiled language, there is greater transparency, assuming that the code is sufficiently annotated and documented. Nevertheless, the flipside of standardisation of data analysis across the department is that errors in the code can affect the results of more people. Rigorous testing of workflows and benchmarking it against the existing solutions is essential.

Throughput is another key point that needs to be addressed. While the traditional approach of having separate templates for data preparation and fitting to models requires individual processing of assay plates, an automated workflow allows for faster processing of larger sets of assay plates, either sequentially or in parallel. This means that throughput will not be primarily limited by the speed of the operator, but by the efficiency of the code. Here, Python offers another advantage as Python handles the memory management for the programmer.

Chapter 2 – Materials and Methods

Note: Construct and purification IDs follow the pattern established in the department for data capture in the database. Therefore, IDs might not appear be sequential when others have worked on the same protein(s).

2.1 General methods

Procedures for cloning were based on a protocol previously published by the SGC [150] which has been updated in the intervening years with regards to reagent and equipment availability.

2.1.1 Plasmid Isolation (Mini Preps)

Plasmid Isolation was carried out using either QIAprep Spin Miniprep kits or NucleoSpin Plasmid EasyPure kits, depending on availability, used according to manufacturer's instructions (see below). All centrifugation steps were carried out in a standard benchtop micro centrifuge.

2.1.1.1 Overnight cultures for mini-prep

E. coli cells (MACH-1) carrying the desired plasmid (e.g., empty expression vector; template DNA; or ligation product ready to be transformed into expression strain) were grown over night at 37°C in a standard shaker incubator at 180 RPM. In a 50 ml conical tube, 5 ml LB medium (See 2.5.1.1) with the appropriate selection antibiotic were inoculated from glycerol stocks. Tubes were sealed with a portion of a breathable microplate seal (StarLabs, Cat. No. E2796-3015). Cells were harvested via centrifugation at 3900 RPM in a standard benchtop centrifuge with swinging bucket rotor and supernatant was carefully removed.

2.1.1.2 QIAprep Spin Miniprep Kit (QIAgen, Cat. No. 27104)

Resuspend pelleted cells in 250 µl resuspension buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA 100 µg/ml RNase A) and transfer into a 2 ml microcentrifuge tube. Cells were then lysed by adding 250 µl alkaline lysis buffer (200 mM NaOH, 1 % (w/v) SDS). After 5 minutes, the reaction was stopped by addition of 350 µl neutralisation buffer (exact composition proprietary, 25-50% guanidinium hydrochloride, 10-25% acetic acid, pH 4.3). Cell debris was then removed via centrifugation (10 min at ≥ 17,000 RCF) and supernatant transferred onto a column in a fresh microcentrifuge tube. Plasmid DNA was then bound to the membrane by passing all liquid through the membrane via centrifugation for 1 minute. Flow through was

discarded. The column was then washed with 500 µl buffer PB (exact composition proprietary, high guanidine hydrochloride and isopropanol concentration to remove proteins and lipids) followed by a wash with wash buffer PE (composition proprietary, contains at least 80% ethanol). Any residual buffer was then removed with an additional centrifugation step. The column was then transferred into a fresh tube and 50 µl pre-warmed (42C) elution buffer (10 mM Tris-HCl, pH 8.5) was added to the centre of the membrane. After 15 minutes of incubation, elution was completed by centrifugation for 15 minutes and the column discarded.

2.1.1.3 NucleoSpin Plasmid EasyPure (Macherey-Nagel, Cat. No. 740727.50)

Pelleted cells were resuspended in 150 µl resuspension buffer (Buffer A) and transferred to 1.5 ml tube. Cells were then lysed by alkaline lysis with 250 µl Buffer B (NaOH lysis) and incubated for 2 min at room temperature. 350 µl neutralisation buffer (Buffer A3, acetic acid) and invert until lysate has turned colourless. Cell debris was then removed by centrifugation for 15 min at ≥ 17,000 RCF. Kit protocol states 3 minutes, but longer time will compact pellet more against the wall of the tube, making handling easier. The supernatant was then loaded onto a NucleoSpin Plasmid EasyPure Column in a collection tube. DNA was bound to the membrane by passing supernatant through via centrifugation for 1 min at 1,000-2,000 RCF and discard flow through. The column was then washed with 450 µl wash buffer (Buffer AQ , high ethanol content). This step was repeated to remove any residual wash buffer. The column was then transferred into a fresh microcentrifuge tube to collect the plasmid. 35 µl pre-warmed elution buffer (10 mM Tris-HCl, pH 8.5) was then added to the centre of the membrane and after 15 minutes of incubation at room temperature the elution was completed by centrifugation for 15 minutes. The column was then discarded.

2.1.2 Plasmid Concentration determination

Concentration of plasmid DNA was determined by measuring the absorbance of the sample at 260 nm (A260) using a NanoDrop ND-1000 instrument (Thermo Fisher Scientific, Waltham, Mass., USA) with the control software version 3.8.1. A blank measurement (mini-prep elution buffer only) was taken first. 2 µl of plasmid sample were then loaded and the A260 determined.

2.1.3 Bacterial Transformation

The same protocol was used for MACH1 (Thermo Fisher Scientific, Cat. No. C862003) cells and expression strains, with the difference being the medium used for incubation prior to plating

out. For MACH1 cells, the supplied SOC medium was used, for expression strains, 1x LB was used (See 2.5.1.1).

50 µl competent cells and 1.5 µl isolated plasmid DNA were added to a 1.5 ml Eppendorf tube and gently pipetted up and down three times using a P200 pipette. Cells were then incubated on ice for 15 min, followed by heat shock treatment (42°C for 45 seconds in a water bath) and further incubation on ice for 2 min. 100 µl of sterile medium were then added to the tube, gently mixed, and incubated for 1 h at 37°C in a stationary incubator. Cells were then plated out on LB agar (See 2.5.1.2) with appropriate selection antibiotics and the plates incubated overnight at 37°C in a stationary incubator.

2.1.4 Expression Vector Selection

Constructs were designed to facilitate crystallography as well as labelled and label-free biophysical and biochemical assays. To this end, expression vectors encoding for different affinity and solubility enhancing tags were used. Where a vector has a protease cleavage site for tag removal, it is indicated in **Table 4**.

Table 4: Overview of expression vectors used for protein production.

Vector name	Tag	Protease	Selection antibiotic	Note
pNIC28-Bsa4	N-terminal His_6	TEV	Kanamycin	(1)
pNIC-CTH0	C-terminal His_6	TEV	Kanamycin	
pNIC-ZB	N-terminal His_6, ZBasic solubility tag	TEV	Kanamycin	(2)
pSUMO-LIC	N-terminal His_6, SUMO solubility tag	SUMO	Kanamycin	
pNIC-Bio3	N-terminal His_6, C-terminal AviTag	TEV	Kanamycin	
pNIC-Bio3-LL	N-terminal His_6, C-terminal AviTag with long linker	TEV	Kanamycin	

His_6: hexahistidyl tag
TEV = Tobacco Etch virus protease
SUMO = SUMO protease
(1) AddGene Catalog No. 26103
(2) AddGene Catalog No 26107

2.1.4.1 Vector preparation

Vector stocks are kept in *E. coli* glycerol stocks (MACH-1). 5 ml overnight culture were grown as described above (2.1.1.1) and plasmid isolated via mini prep as described above (2.1.1.2). Vectors were then prepared for recombination by first digesting with appropriate restriction enzymes (**Table 5** and appendix **Table A2**). Reactions were set up in a final volume of 50 µl, with 5 µl isolated vector, 1.5 µl enzyme and the remainder made up with nuclease free water. If successful, this will excise the region between the two LIC sites, which contains the sacB gene. SacB (levansucrase) converts sucrose into a toxic product and thus acts as a selection marker for clones with successful insertion of PCR product (2.1.8).

Table 5: Conditions for vector restriction enzyme digest.

Enzyme	Buffers	Temperature (°C)	Time (h)	Supplier
BsaI-HF®v2 (20 units/µl)	CutSmart	37 [wb]	3	New England Biolabs
BseRI (4 units/µl)	CutSmart	37 [wb]	3	
BfuAI (5 units/µl)	3.1	50 [tc]	3	
BveI (5 units/µl)	FastDigest + Oligo	37 [wb]	3	Thermo Fisher Scientific

wb: water bath
tc: thermal cycler

2.1.5 Agarose gel electrophoresis

Agarose gel electrophoresis was either carried out using pre-case or in-house made gels.

Pre-case gels were E-Gel®96 1% Agarose (GP) (Invitrogen, Cat. No. G700801). Samples were 3 µl of sample were diluted in 12 µl water and loaded. 20 µl 2x diluted Low Range Quantitative DNA Ladder (Invitrogen, Cat. No. 12373031) were used as marker. Gels were run with the pre-set EG program on a Mother E-Base™ device (Invitrogen, Cat. No. EB-M03).

Cast gels were prepared by dissolving 1.5 % (w/v) agarose (Cleaver Scientific, Cat. No. CSL-AG100) in 1xTAE buffer(40 mM Tris base, 20 mM acetic acid, 1 mM EDTA, pH 8.5) by microwaving until boiling. Once all agarose was dissolved, the solution was left to cool on the bench until it could be safely touched with bare hands. 4 µl gel SYBR safe stain (Invitrogen, Cat. No. S33102) per 100 ml final volume were then added. Gels were run at 120 V for 30 minutes.

2.1.6 T4 DNA polymerase treatment

Vectors and inserts were both treated with T4 DNA polymerase (New England Biolabs, Cat. No. M0203L) to produce 12-14 bp overhanging ends for recombination. Reactions were set up with the same proportions of reagents for both inserts and opened vectors and scaled according to requirements. A 100 µl reaction consist of: 50 µl insert/vector, 10 µl 10x Buffer 2.1 (New England Biolabs, supplied with polymerase), 25 mM appropriate dNTP (Appendix **Table A2**), 100 mM DTT, 1 µl BSA (New England Biolabs, Cat. No. B9000S), 2.5 µl T4 polymerase (New England Biolabs, Cat. No. M0203L). Reactions were incubate using a thermal cycler for 30 min at 22°C, then 20 min at 75°C.

2.1.7 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was used to amplify sections of the protein of interest's gene to create different expression constructs, representing different domains of the protein and to optimise these for expression yield and stability, as well as suitability for assay development.

2.1.7.1 Primer plate

Primers pairs (forward and reverse primer, procured from Thermo Fisher) were prepared on a 96 well PCR plate. 10 µl of 100 µM stock of each primer were added to 180 µl nuclease free water.

2.1.7.2 Template plate

20 µl of 2.5 ng/ µl template were dispensed on a second PCR plate.

2.1.7.3 Master mix and reaction plate

21 µl PCR master mix per well were dispensed on a third plate.

If PCRs failed at first attempt, they were repeated with different master mixes, as indicated below, exploring polymerases with different fidelities and other additives.

Standard master mix (1x): 5 µl 5x Herculase II buffer (supplied with polymerase), 0.75 µl 10 mM dNTP mix, 0.25 µl Herculase II Fusion DNA Polymerase (Agilent, Cat. No. 600675), 15 µl nuclease free water (Invitrogen, Cat. No. AM9932).

Repeat 1 master mix (1x): 5 µl 5x Q5 reaction buffer (supplied with polymerase), 0.5 µl 10 mM dNTPs, 0.25 µl Q5® High Fidelity Polymerase (New England Biolabs, Cat. No. M0491L), 5 µl 5x Q5® high GC enhancer (supplied with polymerase), 9.25 µl nuclease free water.

Repeat 2 master mix (1x): 5 µl 5x Herculase II buffer, 0.75 µl 10 mM dNTP mix, 0.25 µl Herculase II Polymerase, 0.75 µl DMSO, 15 µl Nuclease Free Water

dNTP mix was prepared from 100 mM dNTP set (Invitrogen) with 10 mM each of dATP, dGTP, dCTP and dTTP.

From the prepared plates, 1.5 µl of primers and 2.5 µl template DNA were dispensed onto the corresponding wells on the reaction plate, which was then sealed with a PCR compatible plate seal (4titude, Cat. No. 4ti-0500).

2.1.7.4 PCR protocol

The reactions were carried out using either a C1000 Touch Thermal Cycler (Bio-Rad) or SimpliAmp Thermal Cycler (Life Technologies), based on availability, using the following program:

- 95°C, 10 min.
- (95°C, 30 s | 68°C, 30 s | 68°C 1-3 min) x 5 cycles
- (95°C, 30 s | 60°C, 30 s | 68°C 1-3 min) x 5 cycles
- (95°C, 30 s | 55°C, 30 s | 68°C 1-3 min) x 5 cycles
- (95°C, 30 s | 50°C, 30 s | 68°C 1-3 min) x 10 cycles
- 68°C, 10 min
- 15°C hold.

The extension time was adjusted to the length of the PCR product, allowing for 1 min per 1000 bp.

2.1.7.5 Product length verification

Length of PCR products was verified by agarose gel electrophoresis, either using the commercial 96 well E-Gel® system (Invitrogen, Cat. No. G700801) or in-house cast gels (1.5 % (w/v) agarose, with 1 µl per 10 ml gel SYBR safe stain, Invitrogen, Cat. No. S33102) and running 20 µl 2x diluted Low Range Quantitative DNA Ladder (Invitrogen, Cat. No. 12373031) as marker. In house gels were run at 120 V for 30 minutes.

2.1.7.6 DpnI treatment

If a product was amplified from a template carrying the same antibiotic resistance, it requires DpnI (New England Biolabs, Cat. No. R0176S) treatment to reduce template carryover. Unlike the template DNA, the PCR product will not carry methylation and will therefore not be affected by DpnI digest.

1 µl of 1:20 diluted DpnI (stock solution is 20 units/µl) was added to the appropriate wells on the reaction plate, which was then incubated for 1 h at 37°C.

2.1.7.7 PCR product purification

PCR products were then purified using a MultiScreen PCR_{96} purification plate (Merck, Cat. No. MSNU03050) and vacuum manifold (also Merck). Products were transferred onto the plate and the plates placed on top of the vacuum manifold. Vacuum was then applied until all liquid had passed through (manufacturer recommendation 5 – 10 minutes). Products were then washed with 100 µl/well nuclease free water and vacuum applied until dry. 50 µl/well TE buffer were then added, and the plate incubated for 10 min on the Themo MixMate plate shaker at 1000 RPM. Products were then gently resuspended by pipetting up and down with a multichannel and transferred to a fresh 96 well plate.

2.1.8 Ligation independent cloning

2.1.8.1 Annealing

Constructs were then completed via ligation independent cloning [151]. Purified inserts and opened vectors were first treated with T4 polymerase and appropriate dNTPs as described above (2.1.6). They were then annealed prior to transformation: On a fresh 96 well PCR plate, 1 µl treated plasmid and 3 µl treated insert were mixed, briefly spun down in a microplate centrifuge, and incubated for 2 h at room temperature.

2.1.8.2 Transformation

40 µl of MACH-1 cells were then added per well and the plate incubated on ice for 30 min. Afterwards, they were heat shocked at 42°C for 45 s, then briefly chilled on ice. 100 µl SOC medium (supplied with the cells) were added per well, the plate sealed with a breathable seal and incubated in a static incubator at 37°C for 1 h. Cells were then plated on LB agar, supplemented with the appropriate selection antibiotic for the vector (Appendix **Table A1**)

and 5% (w/v) sucrose, using sterile glass beads to spread the cells on the plate. Plates were incubated overnight incubation at 37°C in a stationary incubator.

2.1.8.3 Colony screen PCR

To test for the presence of the insert, colonies were picked from the agar plates and directly screening the cells for the insert. At the same time, the colonies were expanded to create glycerol stocks in case of success:

Each well of a sterile 96 well master block (Greiner Bio One, Cat. No. 780217) was filled with 1 ml LB medium with the appropriate selection antibiotic. At the same time, a new PCR plate was prepared. The following master mix recipe for one reaction was scaled up appropriately to cover each of the different vectors uses: 4.0 µl 5x reaction buffer (supplied with polymerase), 1.0 µl 10 mM primer mix (pLIC-F: TGTGAGCGGATAACAATTCC, pLIC-R: AGCAGCCAACTCAGCTTCC), 0.1 µl MyTaq DNA polymerase (Bioline, Cat. No. BIO-21106), 14.9 µl nuclease free water. If desired, more than one deep well block and PCR plate can be made up to screen multiple clones at the same time.

A colony of each transformation was then picked with a sterile 200 µl pipette tip and the corresponding well of the PCR reaction plate, followed by the master block, were inoculated. The master block was sealed with a breathable seal and the PCR plate with a PCR seal. The master block was placed in a shaking incubator at 37°C, 700 RPM, and the PCR plate in a pre-heated (94°C) thermal cycler. The program was as follows:

- 95°C, 10 min
- (95°C, 20 sec | 50°C, 20 sec | 72°C, 1-3 min*) x25 cycles
- 72°C, 5 min
- 15°C hold

*MyTaq DNA polymerase requires 30 sec per 1 kb.

PCR products were then analysed on a 1.5% agarose gel. For the additional base pairs added by the LIC primers, refer to appendix **Table A3**. Successful clones were then combined on a single plate: 120 µl were mixed with 30 µl sterile 60% (v/v) glycerol and mixed well. The plate was then sealed and stored at -80°C.

2.1.9 Glycerol stock plate

Once all successful clones have been added to the previous plate, it was thawed and cells once more grown in a master block with 1 ml LB (+appropriate antibiotic) per well. Wells were inoculated with a sterile pipette tip as before. After overnight incubation as described above, another glycerol stock plate was created by transferring 120 µl culture and mixing with 30 µl sterile 60% (v/v) glycerol and mixed well. The plate was then sealed and stored at -80°C.

2.1.10 Small scale test expression and purification

2.1.10.1 Expression

Constructs from 96 well high throughput cloning were assessed by small scale test expressions in *E. coli* BL21(DE3) (New England Biolabs, Cat No. C2527). Two 96 well master blocks with 1 ml TB medium supplemented with appropriate antibiotics were inoculated with expression strains. They were incubated for 5 h in a shaker incubator (Glas-Col Micro-Expression Shaker) at 37°C, 700 RPM, or until an OD_{600} of 2-3 was reached. OD600 was regularly measured. Once they reached 2-3, samples from the whole plate were diluted 1 part culture, 3 parts water (final volume 200 µl) in a flat, clear bottom 96 well microtiter plate. The OD_{600} was then measured using a plate reader. Once target OD was reached across the plate, culture plates were taken out of the incubator and left on the bench to cool for 30 min. Expression was then induced by addition of IPTG to a final concentration of 0.1 mM and the plates returned to the incubator at 18°C, 700 RPM overnight. Cells were sedimented the next day via centrifugation at 3000x RCF. Supernatant was decanted and plates dried upside down on tissue paper.

2.1.10.2 Purification

Cells were lysed using a combination of lysozyme, detergent and freeze thaw cycles. First, cells were resuspended in 200 µl/well lysis buffer (50 mM HEPES, 500 mM NaCl, 10% (v/v) glycerol, 10 mM imidazole, 50 Units benzonase, 0.5 mg/ml lysozyme, 0.1% (v/v) *n*-Dodecyl β-D-maltoside (DDM), 0.4 mM $MgSO_4$, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), pH 8.0) and resuspended using a MixMate® shaker (Eppendorf, Cat. No. 5353000537) at 1500 RPM. Plates were then placed into a -80°C freezer until completely frozen.

Lysates were then thawed in a room temperature water bath for 15 minute and resuspended using a shaking incubator (Glas-Col Micro-Expression Shaker, 10 min at 700 RPM). 3 µl/well

were transferred onto a PCR plate with 27 µl water and 10 µl sample buffer () and stored at 4°C. The lysates were then clarified via centrifugation at 3000 RCF for 10 min.

50 µl pre-equilibrated 50% Ni-NTA resin per well were transferred on a Millipore filter plate (1.2 µm pore size). The clarified lysate was then carefully transferred onto the filter plate and the plate placed into the shaker incubator at 18°C for 1 h at 400 RPM. The lysate was then passed through the resin using the vacuum manifold, taking care not to dry out the resin. The resin was then washed three times by passing 200 µl wash buffer (50 mM HEPES, 500 mM NaCl, 10% (v/v) glycerol, 25 mM imidazole, 0.5 mM TCEP, pH 7.5) through the resin.

On the filter plate, 40 µl elution buffer (50 mM HEPES, 500 mM NaCl, 10% (v/v) glycerol, 500 mM imidazole, 0.5 mM TCEP, pH 7.5) were added per well, the plate was sealed with a clear seal (Qiagen Tape Pad, Cat. No. 1018104) and a 96 well microtiter plate placed under the filter plate. Both were then incubated in the shaking incubator for 20 min at 18C, 400 RPM and the protein eluted by centrifuging the plate stack at 300 RCF for 3 min.

5 µl/well NuPAGE sample buffer (Invitrogen, Cat. No. NP0008) were pre-dispensed on a 96 well PCR plate and 15 µl sample then transferred onto the plate. Samples were then denatured by heating at 80°C for 10 min and then analysed by SDS-PAGE (2.10.1).

2.2 MYLIP construct design and cloning

The sequence of MYLIP was analysed via PSIPRED 4.0 and DISOPRED 3.0 on the PSIPRED Workbench (**http://bioinf.cs.ucl.ac.uk/psipred/**) [152]–[154].

In addition to constructs spanning the canonical FERM domain only, similar to those previously explored by Calkin et al. (M1-A273 and M1-Q344) [78], the disordered region C-terminal of the FERM domain was explored and a total number of 21 boundaries was chosen (See section 3.2). For each boundary, several vectors were chosen from the SGC vector collection.

Primers were ordered from Sigma and solubilised to 100 µM in 10 mM Tris–HCl buffer, pH 8.0 and stored at −20°C.

Cloning procedures were carried out as described above (2.1).

2.2.1 Site directed mutagenesis

Site directed mutagenesis (SDM) was carried out by Charline Giroud during my paternity leave to generate phospho-mimetic mutants of the MYLIP FERM domain for protein-peptide interaction studies.

SDM was carried out via PCR amplification of the construct of interest with mismatching primers. The primers in question carry the desired mutations and during the replication process, the mutation is incorporated into the PCR product.

Primers were designed using Agilent's web tool (**https://www.agilent.com/store/primerDesignProgram.jsp**) and the QuickChange XL kit from Agilent (Cat. No. 200516) was used according to instructions from the manufacturer, as outlined below:

2.2.1.1 Reaction set-up

- 5 µl of 10× reaction buffer
- 5–50 ng dsDNA template
- 125 ng forward primer
- 125 ng reverse primer
- 1 µl of dNTP mix (concentration not stated)
- Nuclease free water to a final volume of 50 µl
- 1 µl of PfuUltra HF DNA polymerase (2.5 U/µl)

2.2.1.2 PCR Protocol

The reactions were carried out using either a C1000 Touch Thermal Cycler (Bio-Rad) or SimpliAmp Thermal Cycler (Life Technologies), based on availability, using the following program:

- 95°C, 30 s.
- (95°C, 30 s | 55°C, 60 s | 68°C 1-3 min) x 12-18 cycles
- 15°C hold.

The number of cycles is dependent on the type of mutation introduced: Point mutations (12 cycles), single amino acid change (16 cycles), or multiple amino acid changes or insertions (18 cycles).

2.3 Enzymes of the ubiquitination cascade

All enzymes required for the full ubiquitination cascade were expressed in *E. coli* BL21 (DE3) cells procured from New England Biolabs (Cat. No. C2527).

2.3.1 Ubiquitin-activating enzyme (E1)

Literature research revealed a protocol to produce murine E1 in *E. coli* cells [155] and pET28-mE1 was a gift from Jorge Eduardo Azevedo (Addgene plasmid # 32534 ; http://n2t.net/addgene:32534 ; RRID:Addgene_32534). The exact molecular weight of the His_6 tagged protein was not stated, but translation of the sequence provided on the Addgene website revealed a 1078 amino acid long construct with a molecular weight of 119972.26 Da.

The construct was supplied as DH5alpha agar stabs in a pET-28b vector with kanamycin resistance cassette (to be supplemented at 50 µg/ml in culture medium). Cells were plated out on LB agar plates with 50 µg/ml kanamycin. After overnight incubation at 37°C, a few colonies were picked, and a 5 ml culture was inoculated and plasmid isolated via miniprep, as described above (2.1.1).

2.3.2 Ubiquitin-conjugating enzymes (E2s)

The E2 constructs (**Table 6**) were a kind donation from Zhuoyao Chen

Table 6: E2 constructs

Synonym 1	Synonym 2	Construct ID
Ubc1	UBC45	UBE2KA-c001
UBC3	UBC54	CDC34A-c001
UBcH5a	UBE2D1	UBE2D1A-c001
UBcH5b	UBC16	UBE2D2A-c001
UBcH5c	UBC19	UBE2D3A-c001
Ubc12	UBE2M	UBE2MA-c001

Out of the six constructs obtained, three were chosen for expression and purification based on the literature: UBE2D1/UBcH5a, UBE2D2UBcH5b, and UBE2D3/UBcH5c, which were registered in our database as UBE2D1A-c001, UBE2D2A-c001 and UBE2D3A-c001, respectively.

```
UBE2D1A-c001  MGSSHHHHHHSSGLVPRGS-ALKRIQKELSDLQRDPPAHCSAGPVGDDLFHWQATIMGPP
UBE2D2A-c001  MGSSHHHHHHSSGLVPRGSMALKRIHKELNDLARDPPAQCSAGPVGDDMFHWQATIMGPN
UBE2D3A-c001  MGSSHHHHHHSSGLVPRGSMALKRINKELSDLARDPPAQCSAGPVGDDMFHWQATIMGPN
              ****************** *****:***.** *****:*********:**********

UBE2D1A-c001  DSAYQGGVFFLTVHFPTDYPFKPPKIAFTTKIYHPNINSNGSICLDILRSQWSPALTVSK
UBE2D2A-c001  DSPYQGGVFFLTIHFPTDYPFKPPKVAFTTRIYHPNINSNGSICLDILRSQWSPALTISK
UBE2D3A-c001  DSPYQGGVFFLTIHFPTDYPFKPPKVAFTTRIYHPNINSNGSICLDILRSQWSPALTISK
              ** *********:************:****:**************************:**

UBE2D1A-c001  VLLSICSLLCDPNPDDPLVPDIAQIYKSDKEKYNRHAREWTQKYAM 165
UBE2D2A-c001  VLLSICSLLCDPNPDDPLVPEIARIYKTDREKYNRIAREWTQKYAM 166
UBE2D3A-c001  VLLSICSLLCDPNPDDPLVPEIARIYKTDRDKYNRISREWTQKYAM 166
              ********************:**:***:*::**** :*********
```

Figure 12: Protein sequences of UBE2D constructs used in this work.

2.3.3 Ubiquitin

The TwinStrep tagged ubiquitin construct was a kind donation from Mike Fairhead, supplied in form of mini-prepped expression vector (pNIC-NHSTIIT backbone, kanamycin resistance cassette). The extended tag was originally added to facilitate protein concentration determination via UV vis spectroscopy at 280 nm, since wild type mono-ubiquitin does not contain tryptophans. Another construct (RPS27AA-c001) was subsequently identified in the CMD collection after a search for sequence similarity and a MACH-1 glycerol stock carrying the plasmid was kindly retrieved from the collection by Korvus Wang of the Biotechnology group. RPS27A is a fusion gene of mono-ubiquitin at the N-terminus of the gene product and the 40S ribosomal protein 27a at the C-terminus. RPS27AA-c001 spans the boundaries of the mono-ubiquitin segment only. To determine protein concentration, bicinchoninic acid (BCA) assay had to be used.

```
TS-Ubi  MHHHHHHSSGASWSHPQFEKGGGSGGGSGGSAWSHPQFEKGSGVDLGTENLYFQSMQIFV  60
RPS27A  ---------------------------------MHHHHHHSSGVDLGTENLYFQSMQIFV  27
                                          ******************

TS-Ubi  KTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKES 120
RPS27A  KTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKES 87
        ************************************************************

TS-Ubi  TLHLVLRLRGGM   132
RPS27A  TLHLVLRLRGG-   98
        ***********
```

Figure 13: Sequence alignment of the ubiquitin constructs used in this work. TS-Ubi: His_6-TwinStrep tagged ubiquitin. RPS27A: RPS27AA-c001

2.4 LDLR adapter proteins

Expression constructs for known LDLR binders LDLR adapter protein 1 (LDLRAP1), Disability homolog-1 (DAB1), and Disability homolog-2 (DAB2) were designed, cloned, expressed, and purified by Charline Giroud while I was on paternity leave. Constructs are summarised in **Table 7**, while all details are listed in **Table A8** The DNA templates are listed in the appendix (**Figure A2**, **Figure A3**, **Figure A4**) and all experiments were carried out following the same methods as described in this work.

Table 7: Construct boundaries and vectors chosen for LDLR adaptor protein production.

Protein	Boundaries	Vectors
DAB1	23-174	pNIC28-Bsa4, pNIC-CTH0, pNIC-Bio3, pNIC-Bio3-LL
	20-175	pNIC28-Bsa4, pNIC-CTH0, pNIC-Bio3, pNIC-Bio3-LL
DAB2	33-181	pNIC28-Bsa4, pNIC-CTH0, pNIC-Bio3, pNIC-Bio3-LL
	33-191	pNIC28-Bsa4, pNIC-CTH0, pNIC-Bio3, pNIC-Bio3-LL
LDLRAP1	44-175	pNIC28-Bsa4, pNIC-CTH0, pNIC-Bio3, pNIC-Bio3-LL
	43-174	pNIC28-Bsa4, pNIC-CTH0, pNIC-Bio3, pNIC-Bio3-LL

2.5 Protein Expression

2.5.1 Media and antibiotics

All media were prepared with 18.2 MΩ water from an ELGA water system (Veolia) and then autoclaved, unless specified.

Selection antibiotics were made up to 1000x stock solutions:

Table 8: Selection antibiotic stock solutions

Antibiotic	Concentration	Source, Catalogue number
Kanamycin sulfate	50 mg/ml in water*	Fisher Bioreagents, 10031553
Chloramphenicol	34 mg/ml in 100% ethanol	AppliChem, A7495.0025
Streptomycin sulfate	100 mg/ml in water*	Sigma-Aldrich, S6501-5G

* 18.2 MΩ water

2.5.1.1 Lysogenic broth (LB)

LB was prepared from powder granule (Merck, Cat. No. 110285) stock: 25 g/l were dissolved in 18.2 MΩ H_2O and broth was sterilised via autoclaving.

2.5.1.2 LB agar plates

25 g/l LB powder stock (as above) and 15 g/l agar (Acros Organics, Cat. No. 40004000050) were dissolved in 18.2 MΩ H_2O and sterilised via autoclaving. Solidified agar was melted in the microwave, decanted in a smaller sterile bottle, and left on the bench until it could be safely touched with the naked hand. Medium was then supplemented with antibiotics (1:1000 from stocks, **Table 8**) as required and ca. 7 ml poured into 60 mm by 15 mm petri dishes (Merck, Cat. No. P5481-500EA).

2.5.1.3 Terrific broth (TB)

For 1 l final volume, 24 g yeast extract (Melford, Y20025-5000.0) and 12 g tryptone (Merck, 111931) and 4 ml glycerol (Fishe were dissolved in 950 ml 18.2 MΩ water and sterilised via autoclaving. 50 ml 20x TB salt stock was added before cultures were started.

2.5.1.4 TB salts (20x stock)

46.2 g KH_2PO_4 x $3H_2O$ (Sigma-Aldrich, P5504) and 250.8 g K_2HPO_4 (Honeywell, 60220) were dissolved in 1 l 18.2 MΩ water and sterilised via autoclaving.

2.5.1.5 SOC

Super Optimal broth with catabolite suppression (SOC) outgrowth medium was supplied with the *E. coli* BL21 DE3 cells purchased from New England Biolabs.

2.5.2 Expression strains

Expression strains of *E. coli* were transformed with isolated plasmid from the MACH-1 glycerol stocks above. Strains used were *E. coli* BL21(DE3)-R3-BirA for constructs with AviTag (produced in house by our biotech team) and *E. coli* BL21(DE3) (New England Biolabs, Cat. No. C2527H)

2.5.3 Transformation

Transformations were performed as described above (2.1.3), with no antibiotic in the LB medium.

2.5.4 Expression

Starter cultures in 1x LB with appropriate selection antibiotic(s) (10 ml for each litre of expression culture) were inoculated from glycerol stocks and incubated overnight in a shake incubator at 37°C and 200 RPM. Expression cultures (in multiples of 1 l TB in 2 l baffled flasks) were then inoculated with 10 ml starter culture per flask and incubated in a shaker incubator at 37°C at 180 RPM. Optical density at 600 nm (OD_{600}) was regularly measured. Once an OD_{600} of 1.4 – 1.8 was reached (ideally 1.8), cultures were moved to a cold room until the incubator was cooled down to 17°C. Expression was then induced by the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG, Fluorochem Ltd., Cat. No. M02726-5G) to a final concentration of 1 mM and the cultures incubated overnight.

Cells were harvested via centrifugation at 6000 RCF for 20 min at 4°C. Supernatant was discarded and disinfected by addition of Rely+ On™ Virkon® tablets according to the manufacturer's instructions (Lanxess) and incubation for at least 45 min. Cell pellets were then either immediately processed as detailed below or stored at -20°C until required.

2.6 MYLIP Protein purification

2.6.1 Buffers

Buffers used for protein purification were prepared at room temperature and stored at 4°C. Due to its short half-life[156], reducing agent dithiothreitol (DTT, D9779-5G, Sigma-Eldritch) was only added directly before use and stock was kept at -20°C. pH of imidazole stocks was adjusted to 8.0.

Table 9: Chromatography buffers for MYLIP protein purification.

Reagent	Lysis	Wash	Elution	Size exclusion
Tris-HCl (1 M stock)	50 mM			
NaCl (5 M stock)	500 mM			
DTT (1 M stock)	2 mM			
Glycerol	10% (v/v)			
Imidazole (2 M stock)	20 mM	50 mM	300 mM	-
pH	7.5			

2.6.2 Lysis

Cells were then resuspended in lysis buffer by agitation with a magnetic stirrer until the suspension was completely homogenous and 1 ml protease inhibitor cocktail (Merck, 535140-1SET) was added. Cells were then lysed via sonication (Vibra-Cell VCX750, Sonics & Materials, Newtown, CT, USA) on ice at 60% amplitude, 10 sec on/20 sec off, 20 minutes total time. Lysate was clarified via centrifugation (32,000 RCF, 60 min, 4°C) and the pellets discarded.

2.6.3 Immobilised Metal Affinity Chromatography (IMAC)

Immobilised metal affinity chromatography was performed using an ÄKTAxpress system (Amersham) with HisTrap FF 5ml columns (Cytiva, Cat. No. 17525501), wash prior to elution (base buffer with 50 mM imidazole) and step elution (base buffer with 300 mM imidazole). 8 ml peak fractions were collected and the first two were pooled and concentrated using 10 kDa MWCO Amicon concentrators to 5 ml (≤ 5% size exclusion chromatography column bed volume for optimal separation)

2.6.4 Protein concentration

To concentrate chromatography fractions, Amicon centrifugal concentrators (Merck) of appropriate molecular weight cut off (MWCO) were used. Concentrators were spun at 3700 RCF at 4°C in a standard benchtop centrifuge with swinging bucket rotor and equilibrated with the same buffer that the sample to be concentrated is in.

2.6.5 Size Exclusion Chromatography (SEC)

Size Exclusion Chromatography was also performed on an ÄKTAxpress system with a S75 16/600 HiPrep column (Cytiva). 8 ml peak fractions were collected.

2.6.6 TEV protease digest

For proteins where a tag was not required, the hexahistidyl tag was cleaved using Tobacco Etch virus protease: Protein sample was mixed with a 100 µL aliquot of TEV protease (produced in house, kindly gifted by Graciela Nunez-Alonso) and added to dialysis tubes (D-Tube Dialyzers, 71506-M, Merck) and dialysed at 4°C over-night against 1 litre of required buffer (e.g., protein storage/SEC buffer, assay buffer, ITC buffer…). Protease, cleaved tag, and un-cleaved protein were removed by passing the reaction through Ni-NTA resin on a gravity column and collecting the flow-through. To assess the efficiency of cleavage, bound protein was eluted (dialysis buffer supplemented with 300 mM imidazole). Flow through and eluted fractions were analysed via SDS-PAGE.

2.6.7 Protein concentration determination via UV_{280} absorbance

Concentration of proteins was determined by measuring the absorbance of the sample at 280 nm (A280) using a NanoDrop ND-1000 instrument (Thermo Fisher Scientific, Waltham, Mass., USA) with the control software version 3.8.1. Unless specified otherwise, all measurements were obtained in triplicate. A blank measurement (same buffer as the protein sample is in) was taken first. 2 µl of protein sample were then loaded and the A280 determined. Protein concentration was then automatically calculated by the software using the molar extinction coefficient ε and the molecular weight of the protein in kDa.

2.7 Murine E1 purification

For murine E1 constructs, different buffers were used, following the protocol developed by Carvalho et al.[155]. TCEP was substituted for DTT due to the short half-life of DTT.

2.7.1 Lysis

Lysis was performed via sonication as described above (2.6.2) and lysate clarification via centrifugation, also as described above. Lysis buffer: 50 mM Tris-HCl, 150 mM NaCl, 0.1% (w/v) Triton X-100, 1 mM TCEP, protease inhibitor (Merck, 535140-1SET), 1 mM EDTA, pH 8.0

2.7.2 IMAC

IMAC was performed on the bench using gravity columns (Bio-Rad Econo-Column, 2.5 x 20 cm, Cat. No. 7372522) and 5-7 ml settled column volume Nickel Sepharose resin (Cytiva, Cat. No. 17371201) equilibrated with lysis buffer. Clarified lysate was added in ca. 40 ml portions, resin was stirred up using a metal spatula. After the bed had been allowed to settle, the valve was opened and the flow through collected. After the entire lysate had been passed through the column, the resin was washed four times with 50 ml wash buffer (50 mM Na phosphate, 150 mM NaCl, pH 8.0), again agitating the bed and letting it settle before removing the buffer. Flow through was again collected separately. Elution was then performed with three times 10 ml elution buffer (50 mM Na phosphate, 150 mM NaCl, 100 mM imidazole, pH 8.0) and all three fractions collected. Following SDS-PAGE as described below (2.10.1) only the elution fractions were retained.

2.7.3 Ion Exchange Chromatography (IEX)

Elution fractions from IMAC were concentrated to 5 ml using 50 kDa MWCO Amicon concentrators and then washed five times by topping up with 10 ml IEX binding buffer (10 mM Tris-HCl, 1 mM EDTA, 1 mM TCEP at pH 8.0) and concentrating back to 5 ml. Ion Exchange Chromatography was then performed on an ÄKTAxpress system with a RESOURCE Q (Cytiva, Cat. No. GE17-1179-01) or Mono Q 5/50 column (Cytiva, Cat. No. GE17-5166-01) with 500 mM NaCl in the elution buffer. 2 ml fractions were collected.

Table 10: Ion Exchange Chromatography step details

Column	Column volume	Binding capacity per 1 ml resin	Gradient (% elution buffer)	Gradient length	Flow rate
RESOURCE Q	6 ml	45 mg BSA	0% – 100%	40 ml	1 ml/min
MonoQ	1 ml	50 mg *	0%-50% 50% - 100%	20 ml (20 CV) 2 ml (2 CV)	0.8 ml/min 0.8 ml/min

* No reference protein stated in product literature
CV: Column volume(s)

2.7.4 Size exclusion chromatography

Size Exclusion Chromatography was also performed on an ÄKTAxpress system with a S200 16/600 HiPrep column (Cytiva) with 20 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 1 mM TCEP, pH 8.0 as running buffer. 8 ml peak fractions were collected. Pooled fractions were then concentrated using 30 kDa MWCO Amicon concentrators.

2.8 Ubiquitin purification

Ubiquitin was purified following a protocol developed by Mike Fairhead, which is detailed below in sections **2.8.1** and **2.8.2**. Determination of protein concentration was performed either via UV_{280} absorbance (for constructs with tryptophans in the affinity tag, as described in section **2.6.7**) or Bicinchoninic acid assay (for constructs without tryptophans, see below in section **2.8.3**).

2.8.1 Lysis

Lysis was performed by sonication, as described above (2.6.2), in ice cold lysis buffer (20 mM HEPES, 500 mM NaCl, 5% (v/v) glycerol, 0.5 mM TCEP, pH 7.5).

2.8.2 IMAC

IMAC was performed using gravity columns as described above (2.7.2). Binding, wash, and elution buffer were created by adding imidazole from a 2 M stock solution at pH 8.0 to yield final concentrations of 30 mM (binding and 50 ml washes W1, W2), 50 mM (50 ml washes W3, W4) and 500 mM (25 ml elution).

2.8.3 BCA assay

Bicinchoninic acid (BCA) assay (Thermo Fisher, Cat. No. 23235) had to be used to determine the concentration of the RPS27AA-c001 expression, as this construct does not contain any tryptophans.

A BSA standard curve was prepared from the provided standard solutions in the kit, starting from 200 µg/ml, in 15 1:1 dilution steps down to 12.2 ng/ml. RSP27AA-p005 and the previous His_6-TwinStrep tagged ubiquitin were also diluted in the same manner. The reagent solution was prepared following the instructions in the kit. Protein dilutions were transferred onto a black, glass bottom 384 well microtiter plate (Greiner Bio One) in triplicate as technical replicates. Using an electronic multichannel pipette, reagent solution was added quickly onto the samples and standard curve. The plate was then sealed and incubated at 37°C for 2 hours.

After incubation, the absorbance at 562 nm was measured with the PHERAstar FSX plate reader. The standard curve was plotted with GraphPad PRISM and concentrations for sample wells where the absorbance fell in the linear range were interpolated.

2.9 UBE2D construct purification

E2 enzymes were expressed and purified following the same protocol as MYLIP constructs.

2.10 Quality control and protein identification

2.10.1 Gel electrophoresis

All fractions collected from IMAC and SEC were analysed via gel electrophoresis (NuPAGE Novex 4-12% Bis-Tris Midi Protein Gels, WG1403BOX, Life Technologies) in MES running buffer (NuPAGE MES SDS Running Buffer, NP00002-02, Life Technologies). 10 µl of fraction was mixed with 10 µl 2x sample loading buffer (NuPAGE LDS Sample Buffer (4x), NP0008, Life Technologies) and 10 µl were loaded on the gel. 5 µl Precision Plus All Protein Blue molecular weight marker (Bio-Rad, 1610373) were loaded. If samples from more than one purification or purification step were loaded on one gel, marker was run between for visual separation and more accurate molecular weight estimation. Gels were run for 40 min at 170 V (constant).

Gels were briefly washed with water and then stained with Instant Blue Protein Stain (Stratech Scientific Ltd, ISB1L-EXP-1L) for one hour. After staining, gels were rinsed and then soaked in water for at least one hour to remove excess stain and then imaged using a Bio-Rad Gel Doc EZ Imager (Bio-Rad, Hercules, California, USA; Catalogue number 1708270).

Continuous fractions displaying bands of expected molecular weight were then pooled and concentrated to ≥300 µM using 10 kDa MWCO Amicon concentrators as above. Concentrated protein was then split into 200 µl aliquots and stored at -80°C.

2.10.2 Intact mass liquid chromatography and mass spectrometry (LC/MS)

Maintenance and calibration of the equipment and production of buffers were carried out by the mass spectrometry facility.

For identification by intact mass, samples were analysed using liquid chromatography on an Agilent 1290 uHPLC system (Agilent Technologies) followed by mass spectrometry on a 6530 electrospray ionisation quadrupole time-of-flight (QTOF) mass spectrometer (Agilent Technologies).

Samples were first diluted to 0.02 mg/ml in solvent A (0.1% formic acid in ultra-high purity water (Millipore)). 50 µl of sample were then injected onto a guard column (2.1 mm x 12.5 mm Zorbax 5 µm 300SB-C3) at 40°C in a column oven at a flow rate of 1.0 ml/min at 90% solvent A, 10% solvent B (0.1% formic acid in LC-MS grade methanol (Chromasolve)). The column was then washed with a linear gradient from 10% to 80% solvent B for 35 seconds. Elution was performed isocratically at 95% solvent B and the column was re-re-equilibrated post-run at 10% solvent B.

Eluate was then passed to the QTOF, operating in positive ion mode at a capillary voltage of 4000 V, nebuliser pressure at 60 PSI, and drying gas flow rate of 12 l/min at 350°C. The fragmentor of the ion optics section was operated at 250 V, the skimmer at 60 V and the octupole RF at 250 V.

Data were analysed using the MassHunter Qualitative Analysis software (Version B.07.00, Agilent).

2.10.3 Thermal stability

Baseline thermal stability of the proteins was assessed using differential scanning fluorimetry using a NanoTemper Prometheus NT.48 (NanoTemper Technologies GmbH) with standard capillaries (PR-C002) before freezing and after thawing. Samples were diluted to 5 µM in SEC buffer, loaded as described in section 2.12.5 and run from 20°C to 95°C with a temperature ramp of 3°C/min at 60% intensity of light source. Further details in section 2.12.5.

2.11 Peptides

Peptides were purchased from LifeTein, LLC.

Table 11: List of all peptides from LDLR family

Cat No	Name	N-terminal addition	Sequence	C-terminal addition	Start	Stop	Length
LT25462	LDLR$_{(808-826)}$		LLWKNWRLKNINSINFDNP		808	826	19
LT25463	biotin-LDLR$_{(808-826)}$	biotin-	LLWKNWRLKNINSINFDNP		808	826	19
LT110654	LDLR$_{(811-834)}$		KNWRLKNINSINFDNPVYQKTTED		811	834	24
LT110655	LDLR$_{(811-834)}$-K-biotin		KNWRLKNINSINFDNPVYQKTTED	-K-biotin	811	834	24
LT25697	LDLR$_{(811-859)}$-K-biotin		KNWRLKNINSINFDNPVYQKTTED EVHICHQDGYSYPSRQMVSLEDDV	-K-biotin	811	858	48
LT36021	Cy5.5-C-LDLR$_{(811-833)}$	Cy5.5-C-	KNWRLKNINSINFDNPVYQKTTE		811	833	23
LT26395	biotin-APP$_{(745-767)}$	biotin-	EERHLSKMQQNGYENPTYKFFEQ		745	767	23
LT26396	biotin-VLDLR$_{(820-842)}$	biotin-	RNWQHKNMKSMNFDNPVYLKTTE		820	842	23
LT26397	biotin-ApoER2$_{(849-871)}$	biotin-	RNWKRKNTKSMNFDNPVYRKTTE		849	871	23
LT112707	APP$_{(745-767)}$		EERHLSKMQQNGYENPTYKFFEQ		745	767	23
LT112708	VLDLR$_{(820-842)}$		RNWQHKNMKSMNFDNPVYLKTTE		820	842	23
LT112709	ApoER2$_{(849-871)}$		RNWKRKNTKSMNFDNPVYRKTTE		849	871	23

N-terminal additions: biotin conjugated to primary amine; Cy5.5 fluorophore coupled to additional cysteine via maleimide crosslinker.
C-terminal additions: biotin coupled to side chain amine on additional lysine.

Table 12: List of additional peptides from LDLR super-family

Cat No	Name	N-terminal addition	Sequence	C-terminal addition	Start	Stop	Length
LT117864	LRP1$_{(4456-4479)}$-K-btn		HQRMTNGAMNVEIGNPTYKMYEGG	-K-biotin	4456	4479	24
LT117865	LRP1$_{(4490-4513)}$-K-btn		ADFALDPDKPTNFTNPVYATLYMG	-K-biotin	4490	4513	24
LT117867	LRP2$_{(4490-4513)}$-K-btn		ADFALDPDKPTNFTNPVYATLYMG	-K-biotin	4490	4513	24
LT117866	LRP1B$_{(4545-4568)}$-K-btn		DLKGPLTAGPTNYSNPVYAKLYMD	-K-biotin	4545	4568	24
LT117867	LRP2$_{(4581-4604)}$-K-btn		NLFKRKSKQTTNFENPIYAQMENE	-K-biotin	4581	4604	24
LT117868	LRP3$_{(2841-2864)}$-K-btn		LCDGDNDCGDNSDENPTYCTTHTC	-K-biotin	2841	2864	24
LT117869	LRP4$_{(1752-1775)}$-K-btn		SKFTDPGMGNLTYSNPSYRTSTQE	-K-biotin	1752	1775	24
LT117870	ITGB1$_{(766-789)}$-K-btn		FEKEKMNAKWDTGENPIYKSAVTT	-K-biotin	766	789	24
LT117871	ITGB1$_{(778-798)}$-K-btn		GENPIYKSAVTTVVNPKYEGK	-K-biotin	778	798	21
LT117872	APLP1$_{(626-649)}$-K-btn		EEQQLRELQRHGYENPTYRFLEER	-K-biotin	626	649	24
LT117873	INSR$_{(982-1005)}$-K-btn		RQPDGPLGPLYASSNPEYLSASDV	-K-biotin	982	1005	24

C-terminal additions: biotin coupled to side chain amine on additional lysine.

2.12 Biochemical/biophysical assays

2.12.1 Peptide displacement assays

The peptide displacement assays used in this work are based on the interaction of a protein or domain of a protein of interest and a peptide fragment (ca. 20 amino acids in length) of its

interaction partner and the readout of the assay is dependent on the proximity of the two biomolecules.

Protocols and background shared between all three types of assays are discussed in this section while particulars are discussed in the corresponding sections.

In the case of both AlphaScreen® and HTRF assays, protein and peptide are both coupled to a donor and acceptor moiety that will produce detectable light if the two moieties are in sufficient proximity for an energy transfer to occur. If the interaction between protein and peptide is disrupted, the transfer between donor and acceptor moiety cannot occur and thus no light is emitted. In this study, proteins are tagged with a hexahistidyl tag which is also used for the IMAC step of protein chromatography. All peptides were purchased from LifeTein, LLC. (Somerset, NJ, USA) and are tagged with biotin on either the N or C terminus.

To determine optimal concentration of assay components, titration experiments were performed:

AlphaScreen: Donor and acceptor moieties at equal concentrations, newly determined for each batch of reagents.

Optimal protein and peptide concentrations, namely the EC_{90} for both, were determined by titrating both components against each other along 1:1 serial dilution. (For an example, see **Figure 14A**.)

All assays are performed in a final volume of 20 µl on 384 well assay plates (ProxiPlate®-384 Plus, PerkinElmer, Cat. No. 6008289) or a final assay volume of 10 µl on 1536 well assay plates (Greiner Bio-One, WG1403BOX). Plates were read using a PHERAstar FS or FSX plate reader with the appropriate optics module (BMG Labtech, Offenburg, Germany).

Reagents were either dispensed with electronic multichannel pipettes or Multidrop Combi liquid dispensers with small volume steel nozzle cartridges (Thermo Fisher Scientific, Cat. No. 5840330).

In the final screening assays, reagents were prepared as 2x concentrated working solutions and added in equal volumes. Assay plates were then incubated between one and 12 hours. All assay formats were stable enough to allow for lengthy incubation times to fit around equipment booking considerations.

2.12.1.1 AlphaScreen

In this system, the donor moiety is a micro bead (roughly 275 nm in diameter[157]) coated with streptavidin which will bind to biotinylated molecules (protein, peptides or small molecule compounds) and produces a singlet oxygen ($^1O_2$) upon excitation with light at 680 nm. The acceptor moiety is a micro bead coated with a nickel chelate which binds poly-histidine tagged proteins and will emit light between 520 – 620 nm upon receiving the singlet oxygen (**Figure 14B**). The half-life and diffusion rate of singlet oxygen in the assay buffer limits the effective distance between the donor and acceptor bead pair to roughly 200 nm[157], [158].

AlphaScreen Histidine (Nickel Chelate) Detection Kits were purchased from PerkinElmer (Catalogue number 6760619). Assay buffer was 25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), 150 mM sodium chloride, 0.1% bovine serum albumin and 0.05% CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate), pH 7.4.

2.12.1.2 Homogenous Time Resolved Förster Resonance Energy Transfer (HTRF)

This system utilises energy transfer between two fluorophores, a donor and acceptor fluorophore[159]. The donor moiety in this case is Europium cryptate, where the 3^+ ion of the lanthanide Europium is chelated in the fluorophore trisbipyridine cryptate. The donor moiety is the allophycocyanin (APC) based fluorophore XL665 which is conjugated to streptavidin. In contrast to enzyme-linked immunosorbent assays (ELISA), the assay is homogenous, i.e. no immobilisation of any components on the wall of the assay plate well is required and all components of the system are in solution.

Upon excitation at 320 – 340 nm, the fluorophore is excited from the base S_0 to the S_1 state for a brief period. When falling back into the base state, the fluorophore emits a virtual photon with a wavelength of 615 nm that can excite other fluorophores in proximity. XL665 is capable of being excited at this wavelength and upon returning to its own base state emits light at 665 nm (**Figure 14C**). The donor fluorophore has a relatively long luminescence lifetime (µs to ms), which can be exploited to increase the signal-to-noise ratio by starting the detection only after most of the background fluorescence upon initial excitation has subsided. Hence, the assay is referred to as a being time resolved.

To minimise the effect of variations on the assay plate (e.g. due to well to well variations and errors in liquid dispensing), the FRET signal (i.e. the intensity of fluorescence emitted by the acceptor (I_A) is commonly normalised against the intensity of fluorescence emitted by the donor (I_D) (**Equation 1**).

$$R = \frac{I_A}{I_D} * 10000$$

Equation 1

Europium (Eu^{3+}) cryptate conjugated anti-His_6 antibodies were purchased from Pekin Elmer (LANCE Eu-W1024 Anti-6xHis, Cat. No. AD0111), Streptavidin-XL665 from Cisbio (Cat. No. 610SAXLB).

Assay buffer was 20 mM Tris (tris(hydroxymethyl)aminomethane), 20 mM sodium chloride (NaCl), 0.05% TWEEN-20, 0.05% bovine serum albumin (BSA), pH 7.0.

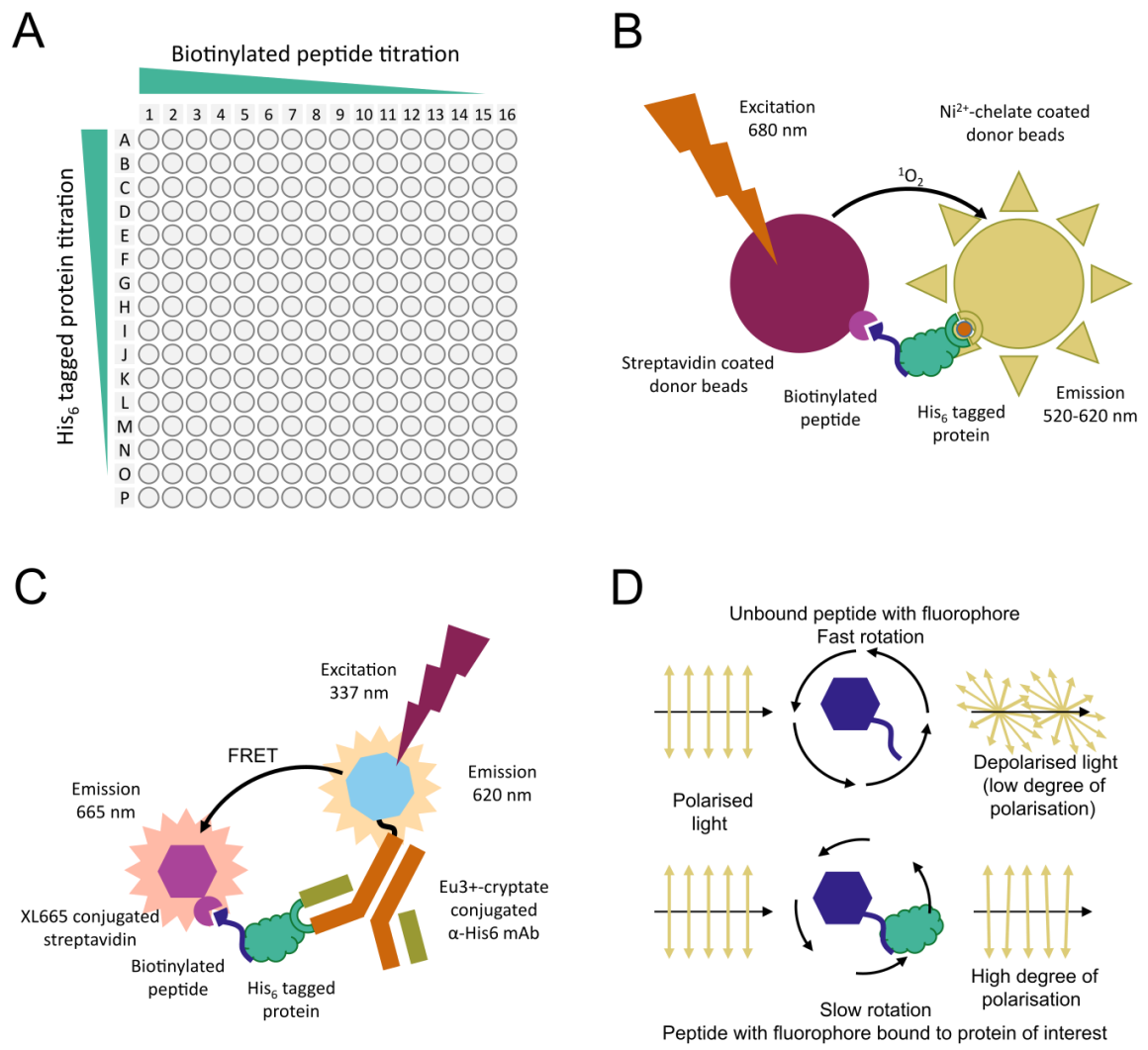


Figure 14: Peptide displacement assays. **A:** 2D titration of biotinylated protein and His_6-tagged peptide. **B:** AlphaScreen. **C:** Homogenous time resolved Förster resonance energy transfer (HTRF). **D:** Fluorescence polarisation.

2.12.1.3 Fluorescence Polarisation Assay

In this assay, the smaller interaction partner, namely the peptide (**Table 11**), is coupled with a Cy5.5 fluorophore (excitation maximum at 675 nm, emission maximum a 694 nm) and the difference in polarised and non-polarised emission of the fluorophore after excitation is measured: Fluorophores on unbound peptides have a higher degree of rotational freedom than fluorophores on bound peptides and thus the fluorescence has a lower degree of polarisation (**Figure 14D**). The readout given by the plate reader is the fluorescence

polarisation (**Equation 2**) or fluorescence anisotropy (**Equation 3**)[160], where $I_{\parallel}$ is the fluorescence intensity parallel to the excitation light plate and $I_{\perp}$ is the fluorescence intensity perpendicular to the excitation light plane. Experiments were carried out in phosphate buffered saline (Merck Life Science UK Limited, Cat. No. P4417-50TAB) supplemented with 1 mM TCEP.

$$FP = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}}$$

Equation 2

$$FA = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}}$$

Equation 3

2.12.1.4 LDLR-PTB cooperativity

The experiments described in this section (LDLR-PTB-cooperativity) were performed by Charline Giroud while I was on paternity leave.

To investigate whether the phospho-tyrosine binding domains of known LDLR adapter proteins have any influence on the association of the LDLR peptide and the FERM domain of MYLIP.

Following the same principle as the other protein-peptide titrations, biotin-VLDLR$_{(820\text{-}842)}$ biotin-ApoER2$_{(849\text{-}871)}$, biotin-APP$_{(745\text{-}767)}$ was titrated against two constructs of DAB1 and DAB2, each, in AlphaScreen and HTRF. Peptide was titrated from 200 nM to 0.2 nM in 10 1:1 steps. PTB domains were titrated from 150 nM to 2.34 nM in 1:1 steps.

AlphaScreen buffer was 20 mM Na phosphate, 100 mM NaCl, 0.1%(w/v) BSA, 0.05% TWEEN-20 (v/v), pH 7.5. Protein-peptide titrations were first incubated for one hour, then 10 µl/well AlphaScreen donor and acceptor beads were added at 1:800 dilution. Plates were then read after 1 h using the PHERAstar FSX.

HTRF buffer was the same as for AlphaScreen. Protein-peptide titrations were again incubated for 1 h prior to addition of reagents. Eu^{3+}-cryptate anti-6his antibody and XL-SA665were diluted at 1:1600. 10 µ/well were added and plates incubated at RT for 1 h before reading plates in the PHERAstar FSX.

HTRF protein-peptide titrations for MYLIP constructs and receptor tail peptides were then repeated as described above, but in presence of a set concentration of PTBs with cleaved hexahistidyl tags.

To investigate whether the generated HTRF signal is specific, MYLIP with cleaved hexahistidyl tag was then titrated into MYLIPA-p018 (250 nM), receptor peptides (62.5 nM) and DAB1/2 PTB (cleaved tag) (200 nM). Plates were processed as described above.

2.12.2 Microscale Thermophoresis (MST)

The basis for this technique is the change in mass and hydrodynamic radius of a particle due to binding an interaction partner[161]. When heat is applied to a small area in a fluid, the resulting heat gradient induces a heat and mass flow in the fluid. The movement of fluorescent particles along with this flow, leads to a drop in fluorescence. If the fluorescent particle is bound to a ligand, its mass and hydrodynamic radius will change, leading to a change in the rate of movement. The dissociation constant K_D between protein and ligand can be determined by fitting a sigmoidal dose response curve to the fluorescence intensity (in percent of initial fluorescence) at the end of the application of heat.

The experiments were carried out on a Monolith NT.155 instrument in “standard” capillaries (MO-K022, NanoTemper Technologies) within the parameters set out by the manufacturer. To determine optimal fluorophore conditions, a serial dilution of peptide was performed and a concentration in the middle of the recommended window of fluorescence intensity was chosen. The same peptide purchased for fluorescence polarisation assays was used.

To determine the K_D of the peptides to proteins, a serial dilution of protein was prepared on a polypropylene 384 well plate (Greiner Bio-One) as a 2x working solution, covering a concentration range of 10x and 0.1x the estimated K_D. To each concentration point, a 2x working solution of the peptide was then added and the two solutions mixed. Capillaries were then loaded via capillary action by inserting the capillaries into the wells of the plate and turning the plate sideways. Data analysis was then performed with the provided analysis software.

2.12.3 Biolayer Interferometry (BLI)

Experiments were carried out with a FortéBio Octet RED^{384} instrument using Super-Streptavidin sensors (FortéBio, Cat. No. 18-5070). All steps were carried out on Tilted-well

(TW384) Microplates (FortéBio, Cat. No. 18-5076) and sensors that were re-used were stored in the provided racks, using a black 96 well flat bottom plates (Greiner Bio-One, Cat. No. 655209) as buffer reservoirs.

2.12.3.1 Sensor loading

For all experiments, biotinylated peptides or proteins were immobilised on Super-Streptavidin sensors (ForteBio, Cat. No.). Sensors were first dipped in assay buffer (25 mM HEPES, 150 mM NaCl, 0.05% BSA, 0.05% TWEEN-20, pH 7.0) to equilibrate for 60 seconds at a stir rate of 1000 RPM, then into 20 µM peptide in assay buffer for 300 seconds (association), followed by 60 seconds in assay buffer to wash off any unbound material (dissociation). If the sensor surface was not saturated (i.e., the sensograms reached a plateau) after the first loading cycle, it was repeated until saturation was observed.

2.12.3.2 K_D determination

For K_D determination, a serial dilution of ligand was prepared from around 0.1x – 10x of the expected K_D. Sensors were moved through "baseline" (equilibrate in a fresh assay buffer well for 60 seconds), "association" (in ligand solution for 60 seconds to records binding of ligand), and "dissociation" (return to same buffer well as before for 120 seconds to record dissociation of ligand). In each cycle, a new baseline/dissociation well is used.

2.12.3.3 Data analysis

Data were analysed using the provided "Data Analysis 11.0" software and GraphPad PRISM.

Briefly, sensograms of each cycle were baseline corrected. The relative response signal intensity over baseline at the plateau of the association step was determined by fitting the data to a one phase association curve. Plateaus of each step were then plotted against the corresponding concentration in PRISM and a one site specific binding model (**Equation 4**) was fitted to the datapoints.

$$Y = B_{max}\frac{X}{(K_d + X)}$$

Equation 4

2.12.4 Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry was used to determine the stoichiometry and dissociation constants of protein-protein, protein-peptide, and protein-small molecule compound interactions.

All experiments were performed using a Nano ITC Standard Volume microcalorimeter (TA Instruments) with a cell volume of 1 ml and a 250 µl volume injection syringe.

Experimental parameters were as follows:

1 Injection of 4 µl, followed by 30 injections of 8 µl. The purpose of the first injection is to displace any air bubbles that could have been introduced at the meniscus at the bottom of the syringe while inserting it into the instrument and to remove any volume of fluid from the opening of the syringe where diffusion during the equilibration time could have led to a localised change in concentration of the ligand. The concentrations in cell and syringe were kept at a 1:10 ratio, generally at 30 µm to 300 µm. The component with the lower solubility was placed in the cell. Default stir rate and temperature were 350 rpm and 25°C, respectively. Concentrations, temperature and stir rate were adjusted for optimal results after initial experiments gave an indication of the behaviour of the interaction studied.

Data were analysed using the proprietary NanoAnalyze software from the manufacturer. Integrated heats of injection were fitted to an independent model.

Between experiments, the sample cell was washed with 50 ml ITC buffer, 50 ml 18.2 MΩ water and then filled with ITC buffer prior to the addition of sample.

The syringe was washed 5 times with ITC buffer, then flushed with copious amounts of 18.2 MΩ water.

2.12.5 Differential Scanning Fluorimetry (DSF)

2.12.5.1 Thermal Unfolding

Thermal stability of proteins was assessed by differential scanning fluorimetry using a Prometheus NT.48 system (NanoTemper), which measures fluorescence (330 nm and 350 nm) and light scattering. Data were analysed using the provided PR.ThermControl software.

Fluorescence intensity depends on the number of tryptophans and tyrosines in a protein, so the intensity of the light source must be adjusted to achieve a high enough signal and gather

reliable information (manufacturer recommended window is 8000 to 18000 counts). Samples were diluted to appropriate concentration (10 µM) and loaded into capillaries (nanoDSF grade standard capillaries, PR-C002, NanoTemper) via capillary action.

2.12.5.2 Isothermal Measurements

Stability of biomolecules was assessed using isothermal measurements with the same device and samples were prepared the same way. The provided PR.TimeControl software was used. Samples were prepared and loaded as described above and fluorescence and light scattering were recorded over 90 min at 25°C.

2.13 Enzymatic assays – substrate ubiquitination

Enzymatic assays were performed to detect the conjugation of ubiquitin subunits to the substrate peptides.

2.13.1 Detection of LDLR ubiquitination via Western blot

In the first instance, the assay described by Martinelli et al [106] was recreated with Cy5.5 labelled $LDLR_{(811-833)}$ (**Table 11**) and full length MYLIP ($MYLIP_{FL}$) produced in this study was compared to commercially obtained MYLIP/IDOL (AMS Biotechnology, Abingdon, UK, Catalogue number 71208, Lot# 130521.HF). Imaging was performed on a LI-COR Odyssey DLx imaging system (Lincoln, Nebraska, USA) and images processed with Image Studio Lite (LI-COR).

Enzymatic reactions were set up in 1.0 ml volumes in 1.5 ml microcentrifuge vials and carried out in an improvised water bath: A pipette tip box was filled with water, pre-warmed to 37°C, onto which a tube floater was placed. The bath was then placed into a small benchtop incubator to maintain temperature over the incubation time.

The components of the reactions were as follows: 5 mM ATP, 200 nM MUBE1, 250 nM UBE2D1, 50 nM $MYLIP_{FL}$ (MYLIPA-p020 or commercially sourced), 15 µM ubiquitin, 15 µM Cy5.5-$LDLR_{(811-833)}$. Reaction buffer was 25 mM Tris-HCl, 150 mM NaCl, 1 mM TCEP, 3 mM $MgCl_2$, at pH 8 (adjusted at room temperature). ATP was added last as the reaction starter. The 0 min time point was taken before addition of ATP.

At 0, 5, 30 and 60 minutes, and after over-night incubation, 10 µl samples for SDS-PAGE were taken and the reactions were stopped by adding the samples onto 14 µl stop/quenching

solution (711 mM DTT, 142 mM EDTA), pre-dispensed onto a 384 well polypropylene microtiter plate (200 µl square wells). DTT reduces the catalytic cysteines in the system to stop any transfer of activated ubiquitin, while EDTA chelates any free Mg^{2+} to stop further activation of ubiquitin.

Western blotting was kindly performed by Andreas Damianou using the following protocol:

Samples were separated using Tris–glycine SDS-PAGE gradient (4–15% acrylamide) gels (Bio-Rad, Cat. No. 4561085) (running buffer: 25 mM Tris base, 192 mM glycine, 2 % (w/v) SDS, pH 7.6) and transferred onto nitrocellulose membranes (Membranes) using Mini Trans-Blot® Cell (Bio-Rad, Cat. No. 1703930) (transfer buffer: 20 mM Tris base, 200 mM glycine, 20% (v/v) methanol). Membranes were then blocked, incubated with anti-ubiquitin antibody, secondary antibody. Membranes were imaged using the LI-COR Odyssey using the 700 nm channel (Cy5.5 labelled peptide) and 800 nm channel (anti-ubiquitin antibody).

2.13.2 Detection of LDLR ubiquitination via gel fluorescence

These reactions were set up as above (2.13.1) Control reactions with absent assay components were incorporated: Three t=0 measurements without E1 (UBE1), E2 (UBE2D) and-E3 (MYLIP), respectively, as well as a full time-course without substrate.

SDS-PAGE was performed as described above (2.10.1) and the gel was imaged on the LI-COR Odyssey using the 700 nm channel.

2.13.3 HTRF based ubiquitination assay – discrete timepoints

This assay is based on the formation of a FRET fluorophore pair (**Figure 14C**) between Eu^{3+}-cryptate labelled ubiquitin and streptavidin-XL665 binding to the biotinylated receptor tail peptides after the peptide has been poly-ubiquitinated.

In the first instance, the assay was performed following the publication example given by the manufacturer[162]: 200 nM biotinylated substrate peptide was incubated with 300 nM UBE2D, 150 nM $MYLIP_{(Full\ length)}$, 200 nM ATP, and a 20:500 nM mixture of Eu^{3+}-cryptate labelled ubiquitin and unlabelled ubiquitin. Five reactions were set up with four different peptides $LDLR_{(811-833)}$-K-biotin, biotin-$VLDLR_{(820-842)}$, biotin-$ApoER2_{(849-871)}$, biotin-$APP_{(745-767)}$ and $LDLR_{(811-833)}$-K-biotin again in the absence of ATP (negative control). The reaction buffer was the same as above (25 mM Tris-HCl, 150 mM NaCl, 1 mM TCEP, 3 mM $MgCl_2$, pH 8.0).

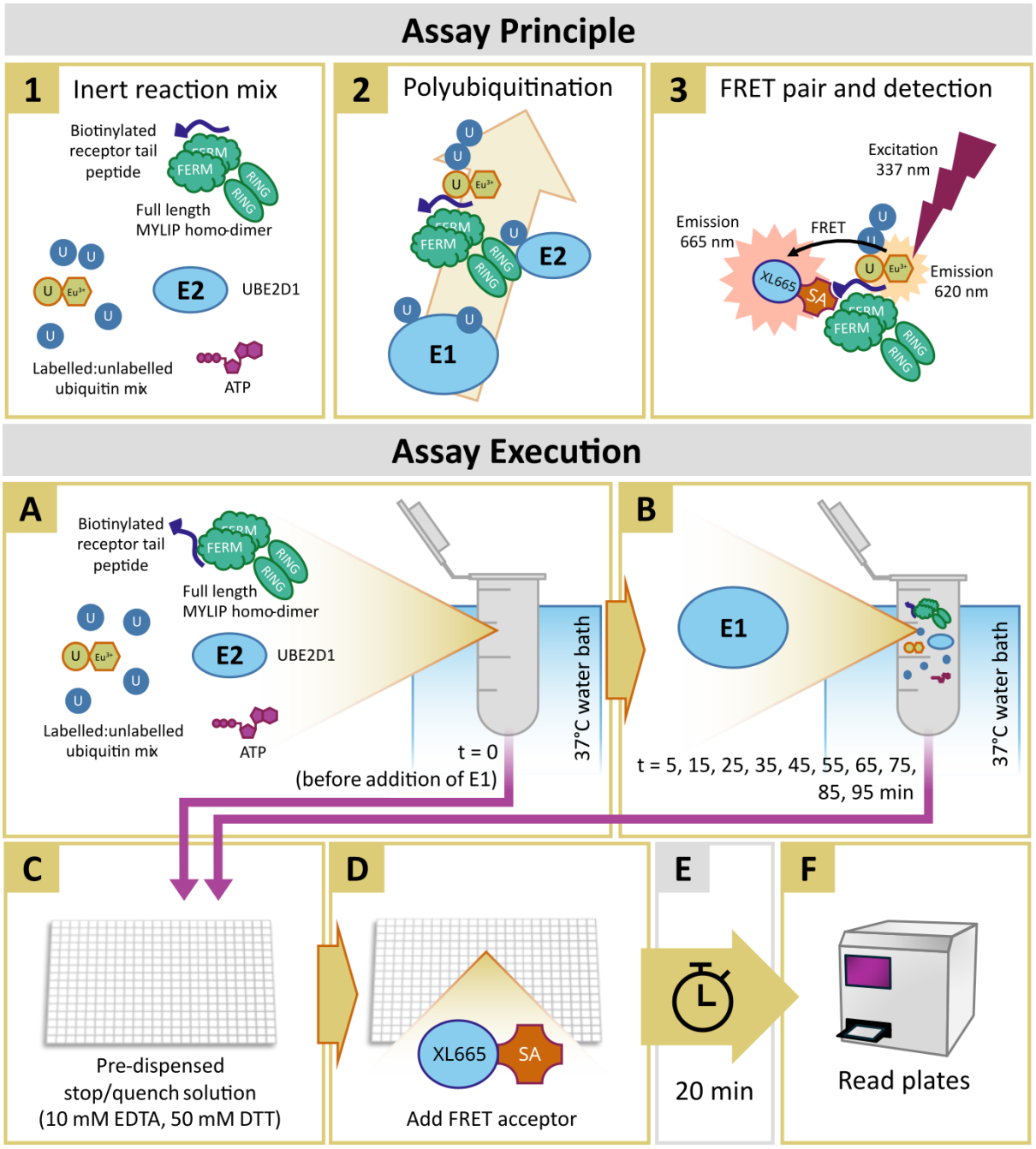


Figure 15: Ubiquitination assay with discrete timepoints **Top:** Assay principle. **1:** Reaction mix is inert before addition of E1. **2:** Polyubiquitination of substrate peptide commences after addition of E2. **3:** After quenching of ubiquitination cascade, FRET acceptor (streptavidin-XL665) is added and binds to the biotinylated substrate peptide. FRET can be observed after excitation of the donor fluorophore if the distance between donor and acceptor is sufficiently small. **Bottom:** Assay execution **A:** Combining assay reagents (biotinylated substrate peptide, full length MYLIP, UBE2D1, labelled and unlabelled ubiquitin, and ATP; t = 0 timepoint sample taken before addition of E1. **B:** Starting reaction via addition of E1; samples taken at indicated timepoints. **C:** 384 well assay plate with pre-dispensed stop/quench solution. **D**: Addition of FRET acceptor (streptavidin-XL665). **E:** Incubation for 20 minutes. **F:** Read plates with PHERAStar FS plate reader.

The reactions were performed at room temperature in 2 ml microcentrifuge tubes and started with the addition of 8 nM E1. At set time points (0, 5, 15, 25, 35, 45, 55, 65, 75, 85, 95 minutes) 3x 10 µl were removed and transferred onto a 384 well plate with pre-dispensed stop solution (10 mM EDTA, 50 mM DTT; 5 µl/well). Time point t = 0 min was taken before the addition of E1.

After the last time point, 5 µl/well detection solution (200 nM FRET acceptor streptavidin-XL665 (Perkin Elmer)) was added, and the plate was incubated at room temp for 20 minutes. The plate was then read with the standard HTRF protocol compatible with the Eu^{3+}-crypate/streptavidin-XL665 fluorophore pair on the PHERAstar FS plate reader.

2.13.4 Continuous HTRF based ubiquitination assay

In the first instance, the assay was conducted to determine optimal mUBE1 concentration with biotinylated substrate peptides ($LDLR_{(811-833)}$, $VLDLR_{(820-842)}$, $ApoER2_{(489-871)}$, $APP_{(745-767)}$, **Table 11**).

The reactions are performed on a 384 well ProxiPlate (Perkin Elmer) in a 20 µl final assay volume and FRET signal is measured every five minutes. Reaction mixes were prepared in the same buffer as before: 20 nM Eu^{3+}-cryptate labelled ubiquitin, 500 nM unlabelled ubiquitin, 200 nM substrate peptide, 300 nM UBE2D1, 150 nM $MYLIP_{FL}$, 200 nM ATP, 200 nM streptavidin-XL665. They were pre-dispensed at 10 µl per well and the reaction was started by adding 10 µl starter mix (0, 1, 2, 4, 8 or 16 nM mUBE1 in reaction buffer). Plates were then read using a BMG Labtech PHERAstar FS or FSX plate reader, depending on availability.

2.13.4.1 Optimisation of HTRF based ubiquitination assay

Concentrations of E1 were varied from 0 to 40 nM and concentrations of Eu^{3+}-cryptate-ubiquitin were varied from 10 to 80 nM in a checkerboard fashion (**Figure 16**). Other assay components were used as described above (2.13.4). Peptides tested were biotin-$LDLR_{(811-833)}$, $VLDLR_{(820-842)}$-K-biotin, $ApoER2_{(489-871)}$-K-biotin and $APP_{(745-767)}$-K-biotin, as well as biotin-$LDLR_{(811-833)}$ in absence of ATP as a negative control. The experiment was conducted in continuous mode and for comparison of results, the final FRET ratios were used.

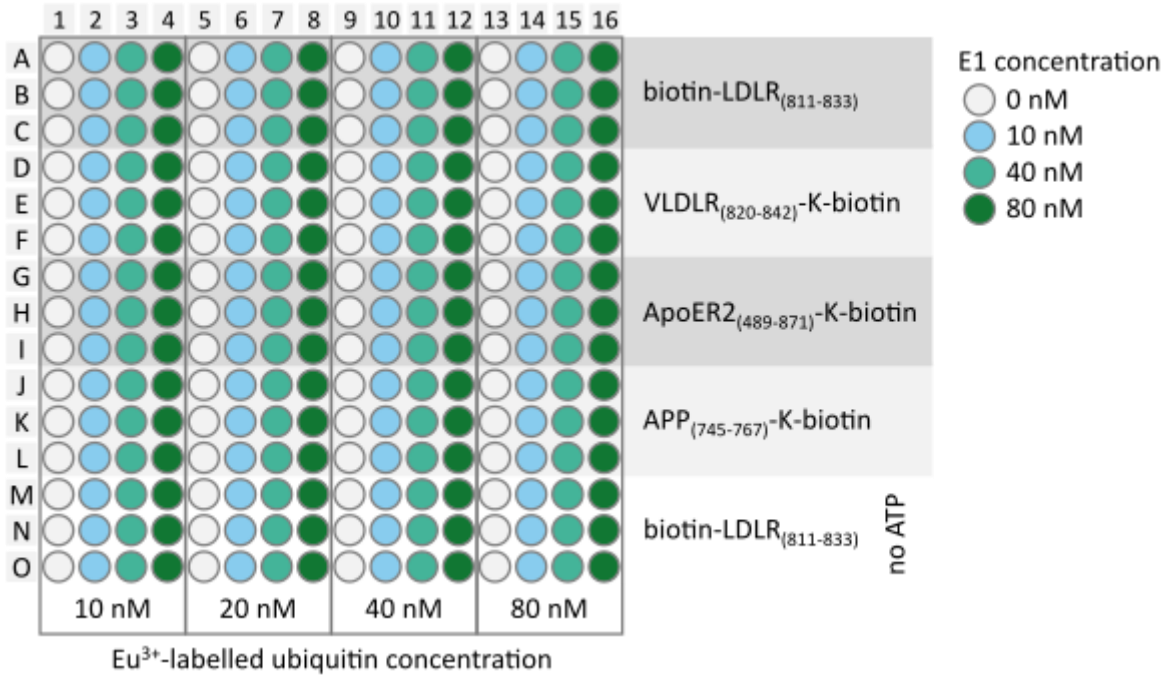


Figure 16: Checkerboard titration layout for E1 and Eu^{3+}-cryptate-ubiquitin concentration optimisation.

2.13.5 MUBE1-p002 batch validation

In the first instance, a gel fluorescence-based assay was performed as described above (2.13.2) with appropriate control reactions (each leaving out one of the components) and all three UBE2D isoforms that were expressed. A deliberately high concentration was used to ensure visibility of product formation. Assay concentrations were as follows: E1 250 nM, E2 200 nM, MYLIPA-p022 50 nM, human mono-ubiquitin 15 µM, Cy5.5-LDLR(811-834) 0.25 µM, ATP 1 mM.

For the plate based HTRF assay, the assay concentration of the first batch had been determined as described above (2.13.4.1). Using the previously established conditions, another assay was carried out where a dilution series of E1 was tested (80 nM, 40 nM, 20 nM, 10 nM, 5 nM, 2.5 nM, 1 nM, no enzyme as negative control). This was carried out as a continuous assay on a 384 well microtiter plate (ProxiPlate) in the PHERAstar FSX plate reader with readings taken every five minutes.

Table 13: reaction scheme for MUBE1-p002 validation via gel fluorescence assay.

	UBE2D1A-p001							UBE2D2A-p001							UBE2D3A-p001						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
E1	-	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+	+	+
E2	+	-	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+	+
E3	+	+	-	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+
Ubi	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+
Pep	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+
ATP	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+

2.13.6 Testing additional members of LDLR superfamily

Additional members of the LDLR superfamily were selected to for investigation as neo-substrates of IDOL. Peptides containing the NPxY endocytic motive were chosen. While IDOL requires a lysine adjacent to the NPxY motif to ubiquitinate a peptide, not all peptides had one.

Table 14: List of additional peptides from LDL super-family.

Name	Sequence	NPxY	Adjacent lysine
LRP1$_{(4456-4479)}$-K-btn	HQRMTNGAMNVEIGNPTY**K**MYEGG	-	→
LRP1$_{(4490-4513)}$-K-btn	ADFALDPD**K**PTNFTNPVYATLYMG	+	←
LRP2$_{(4490-4513)}$-K-btn	ADFALDPD**K**PTNFTNPVYATLYMG	+	←
LRP1B$_{(4545-4568)}$-K-btn	DL**K**GPLTAGPTNYSNPVYA**K**LYMD	+	↔
LRP2$_{(4581-4604)}$-K-btn	NLF**K**R**K**S**K**QTTNFENPIYAQMENE	+	↑
LRP3$_{(2841-2864)}$-K-btn	LCDGDNDCGDNSDENPTYCTTHTC	+	
LRP4$_{(1752-1775)}$-K-btn	S**K**FTDPGMGNLTYSNPSYRTSTQE	+	
ITGB1$_{(766-789)}$-K-btn	FEKE**K**MNAKWDTGENPIYKSAVTT	+	→
ITGB1$_{(778-798)}$-K-btn	GENPIY**K**SAVTTVVNP**K**YEG**K**	+	↔,[x]
APLP1$_{(626-649)}$-K-btn	EEQQLRELQRHGYENPTYRFLEER	+	
INSR$_{(982-1005)}$-K-btn	RQPDGPLGPLYASSNPEYLSASDV	+	

→ Lysine is downstream
← Lysine is upstream
↔ Lysines upstream and downstream
[x] Lysine is inside NPxY motif

2.13.6.1 Competition assay with non-biotinylated peptide

Non-biotinylated LDLR(811-834) was pre-dispensed from DMSO stock using the Echo acoustic liquid handler to final assay concentrations from 400 µM to 190 nM. The same assay buffer as before was used: 25 mM Tris-HCl, 150 mM NaCl, 1 mM TCEP, 3 mM MgCl2, at pH 8 (adjusted at room temperature).

Reagents were then added stepwise in three mixes:

Mix 1 (10 µl/well): 80 nM Eu3+-cryptate ubiquitin, 420 nM unlabelled ubiquitin, 300 nM UBE2D1-p001, 150 nM MYLIPA-p022, 200 nM ATP, 200 nM SA-XL665

Mix 2 (5 µl/well): 80 nM MUBE1-p002.

Mix 3 (5 µl/well): 200 nM LDLR(811-834)-K-biotin.

Upon dispensing of the last mix, the plate was immediately read on the PHERAstar FSX plate reader with the HTRF time course protocol (one read every 5 minutes).

2.14 MYLIP autoubiquitination assay

2.14.1 MUBE1A-p003.2 batch validation

MUBE1A-p003.2 was validated for the auto-ubiquitination assay at discrete time points, carried out on a 384 well master block deep well plate (Greiner Bio-One, Cat. No. 781270). Reactions were set up as the following master-mixes with three technical replicates.

Complete reaction mix: 16 nM MUBE1A-p003.2, 300 nM UBE2D1A-p001, 150 nM MYLIPA-p022, 20 nM Eu3+-cryptate ubiquitin, 480 nM TwinStrep-ubiquitin, 200 nM ATP in a final volume of 450 µl, split over three wells on 384 well master block.

Mix 1 (E1, E2, E3, ATP): Complete reaction mix.

Mix 2 (E2, E3, ATP): Omission of E1

Mix 3 (E1, E3, ATP): Omission of E2

Mix 4 (E1, E2, ATP): Omission of E3

Mix 5 (E1, E2, E3): Omission of ATP

Stop/quenching solution: 50 mM DTT, 10 mM EDTA in a final volume of 5000 µl.

Detection solution: 200 nM SA-XL665, in a final volume of 2000 µl.

ATP was added simultaneously at the end to start the reaction. 5 µl/well stop/quenching solution were pre-dispensed on a 384 well ProxiPlate and at set timepoints (0 min = before addition of ATP, 1 min, 2 min, 5 min, 10 min, 15 min, 20 min then every 10 minutes until 80 min), 10 µl were transferred to the corresponding wells on the ProxiPlate. 5 µl/well of

detection solution were then added and incubated for a further 30 min before the plate was read on the PHERAstar FSX. Data was then plotted.

2.14.2 Compound dispensing

Compound stocks were dispensed using a Labcyte Echo555 acoustic liquid dispenser (Beckman Coulter) from low dead volume (LDV) plates (Beckman Coulter, LP-0200). The dispenser dispenses in discrete 2.5 nl droplets, making it necessary to create intermediate dilutions to dispense full dose response curves. To minimise sample use these were also created on LDV plates. All transfers were backfilled with DMSO to ensure equal DMSO content in every well to keep any potential effects of DMSO on the assay consistent over the plate.

DMSO for backfilling, creation of intermediate dilutions and DMSO reference wells was dispensed from 384 well polypropylene plates (Beckman Coulter, PP-0200)

The Echo system is only capable of calculating final assay concentrations from one stock concentration. If compounds on a plate were of a different concentration than the majority, this would later be corrected manually in the analysis.

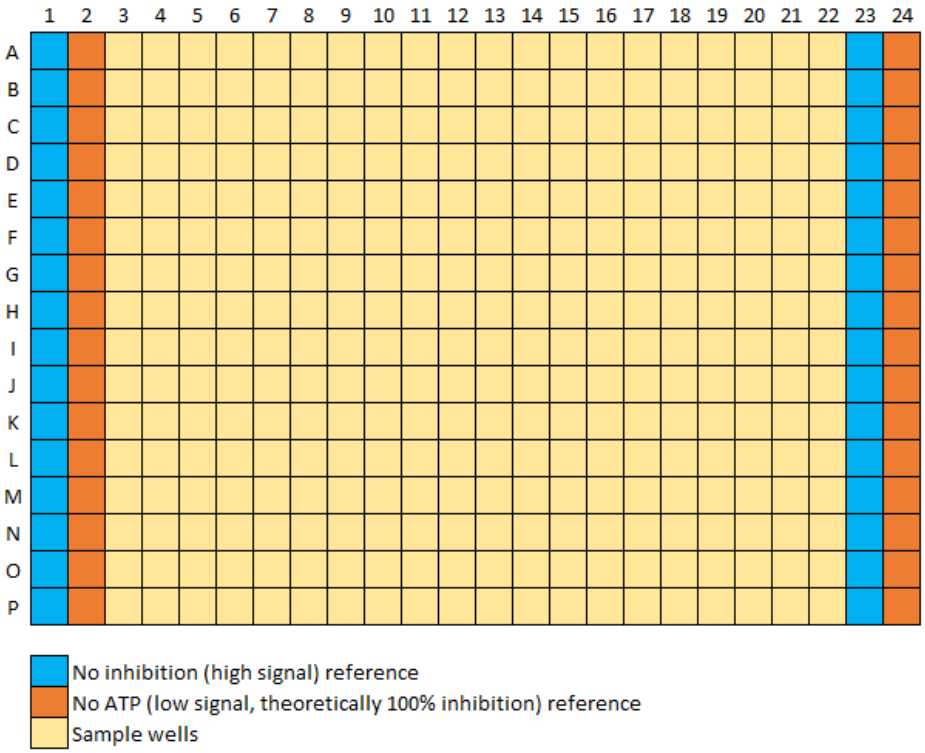


Figure 17: MYLP auto-ubiquitination assay pilot screen plate layout.

2.14.3 Pilot screen

Assay buffer was the same buffer as above (2.13.1).

The reaction's components were divided into four master mixes, all prepared in assay buffer ahead of beginning the experiment (**Table 15**). Mix 3 was prepared during the reaction incubation due to the short half-life of DTT in solution. Plates were read using the BMG PHERAstar FSX plate reader. Data were analysed using the Endpoint Single Dose workflow of BBQ.

Table 15: Master mixes for MYLIP autoubiquitination pilot screen

Mix 1	Enzymes and substrates	16 nM MUBE1A-p003, 300 nM MUBE2D1A-p001, 150 nM MYLIPA-p022 20 nM ubiquitin-Eu3+-cryptate (FRET donor) 480 nM ubiquitin-TwinStrep-His_6
Mix 2	Start solution	200 nM ATP
Mix 3	Stop/quench solution	50 mM DTT, 10 mM EDTA
Mix 4	Detection/FRET acceptor	100 nM SA-XL665 (FRET acceptor)

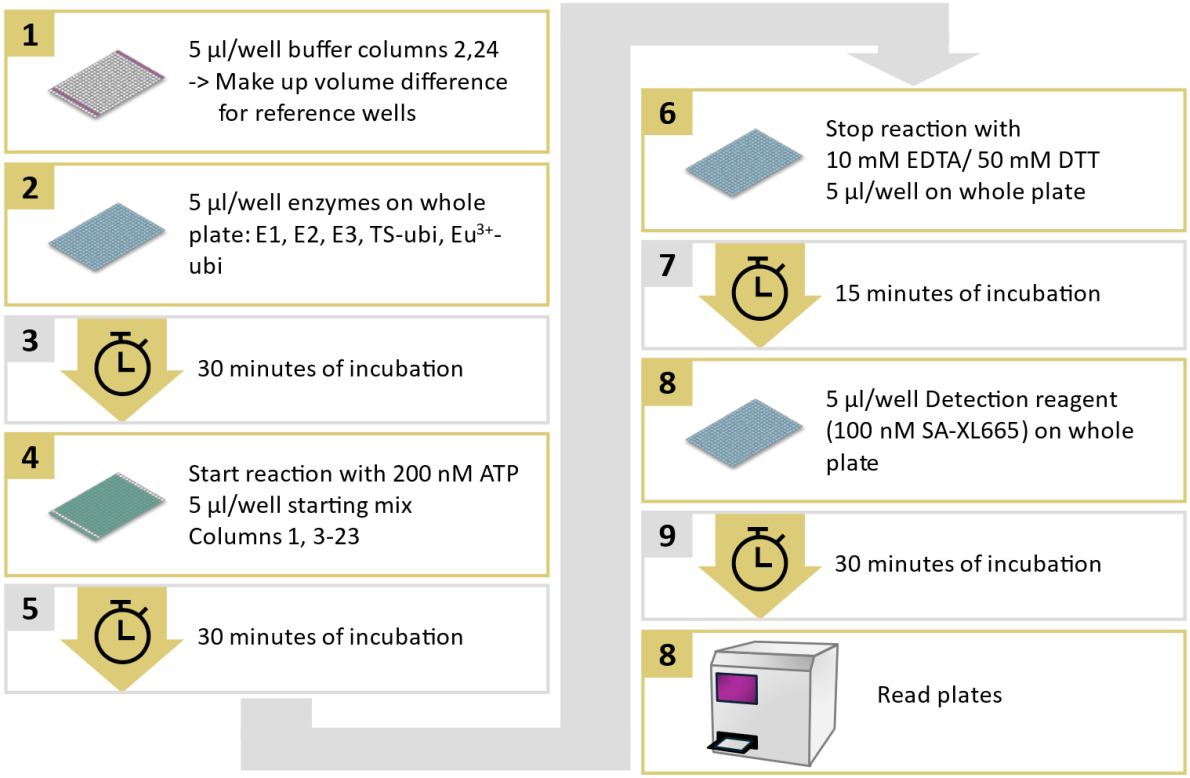


Figure 18: MYLIP auto-ubiquitination assay pilot screen assay steps.

Chapter 3 - Results of Protein Production and Assay Development

3.1 Introduction

Calkin et al.[78] demonstrated a dose-dependent readout in a fluorescence polarization assay where isolated MYLIP FERM domain (M1-273) was titrated into a set concentration of BODIPY labelled LDLR family receptor tail peptides. They were able to determine K_D values for $LDLR_{(811-833)}$, $VLDLR_{(820-842)}$, and $ApoER2_{(849-871)}$ of 26 µM, 18 µM, and 8 µM, respectively. Based on these findings, it should be possible, if challenging, to establish a protein-peptide interaction/peptide displacement assay amenable to high throughput screening (as discussed in 1.4.3.1 and 2.12.1). In the first instance, the results of Calkin at al. were to be reproduced using 2D titrations (serial dilution of MYLIP FERM domain vs serial dilution of receptor tail peptides) in available assay technologies, followed by validation using orthogonal biophysical methods. Once validated, the assay would then be optimised for high-throughput screening. To test the robustness of the assay and to generate initial results, a small sub-set of a diversity library would then be screened with the assay(s) developed. Additionally, an enzymatic assay measuring the ubiquitination of receptor tails was to be developed as another orthogonal method.

To enable the experimental part of this work, various protein constructs were designed, cloned, expressed, and purified: MYLIP constructs to establish a protein-peptide-interaction assay, full length MYLIP for enzymatic studies and assay development, as well as the remaining components of the ubiquitination cascade (E1, E2 and ubiquitin).

3.2 Protein construct design and production

The following section shows the results of the design, cloning, expression and purification of different MYLP constructs as well as the other proteins of the ubiquitination cascade to support the development of the various biochemical and biophysical assays required for this work. The overwhelming majority of proteins expressed and purified satisfactorily and could be used in the efforts to confirm the results from the literature and subsequent efforts in assay development.

3.2.1 MYLIP construct design, cloning and test expressions

After inspecting sequence analysis using DISOPRED workbench and post-translational modification sites deposited on PhosphoSitePlus (**Figure 20**), 21 boundaries were chosen to explore the increasingly disordered region following the canonical FERM domain(**Table 16**, **Figure 20**) and 90 MYLIP constructs were designed in total (Appendix, **Table A4**, Primers in **Table A5**, **Table A6**). The first batch of constructs (c001-c048) has been cloned by Leela Shrestha of the unit's Biotechnology team as part of the normal pipeline (Constructs cloned by me, or during my paternity leave by Charline Giroud are clearly indicated.) In the first instance, the first batch of 47 MYLIP FERM domain underwent small scale text expression in *E. coli* BL21(DE3) cells (**Table A7, Figure 21**) to determine relative expression levels and to inform which constructs to carry forward for full scale expression with the constructs shown in **Table 17** having been selected. These spanned the canonical FERM domain (terminating between A273 and F300) as well as parts of the extended FERM domain (terminating at V335).

```
________10 ________20 ________30 ________40 ________50 ________60
MLCYVTRPDA VLMEVEVEAK ANGEDCLNQV CRRLGIIEVD YFGLQFTGSK GESLWLNLRN
________70 ________80 ________90 _______100 _______110 _______120
RISQQMDGLA PYRLKLRVKF FVEPHLILQE QTRHIFFLHI KEALLAGHLL CSPEQAVELS
_______130 _______140 _______150 _______160 _______170 _______180
ALLAQTKFGD YNQNTAKYNY EELCAKELSS ATLNSIVAKH KELEGTSQAS AEYQVLQIVS
_______190 _______200 _______210 _______220 _______230 _______240
AMENYGIEWH SVRDSEGQKL LIGVGPEGIS ICKDDFSPIN RIAYPVVQMA TQSGKNVYLT
_______250 _______260 _______270 _______280 _______290 _______300
VTKESGNSIV LLFKMISTRA ASGLYRAITE THAFYRCDTV TSAVMMQYSR DLKGHLASLF
_______310 _______320 _______330 _______340 _______350 _______360
LNENINLGKK YVFDIKRTSK EVYDHARRAL YNAGVVDLVS RNNQSPSHSP LKSSESSMNC
_______370 _______380 _______390 _______400 _______410 _______420
SSCEGLSCQQ TRVLQEKLRK LKEAMLCMVC CEEEINSTFC PCGHTVCCES CAAQLQSCPV
_______430 _______440
CRSRVEHVQH VYLPTHTSLL NLTVI
```

Figure 19: Amino acid sequence of full length MYLIP.

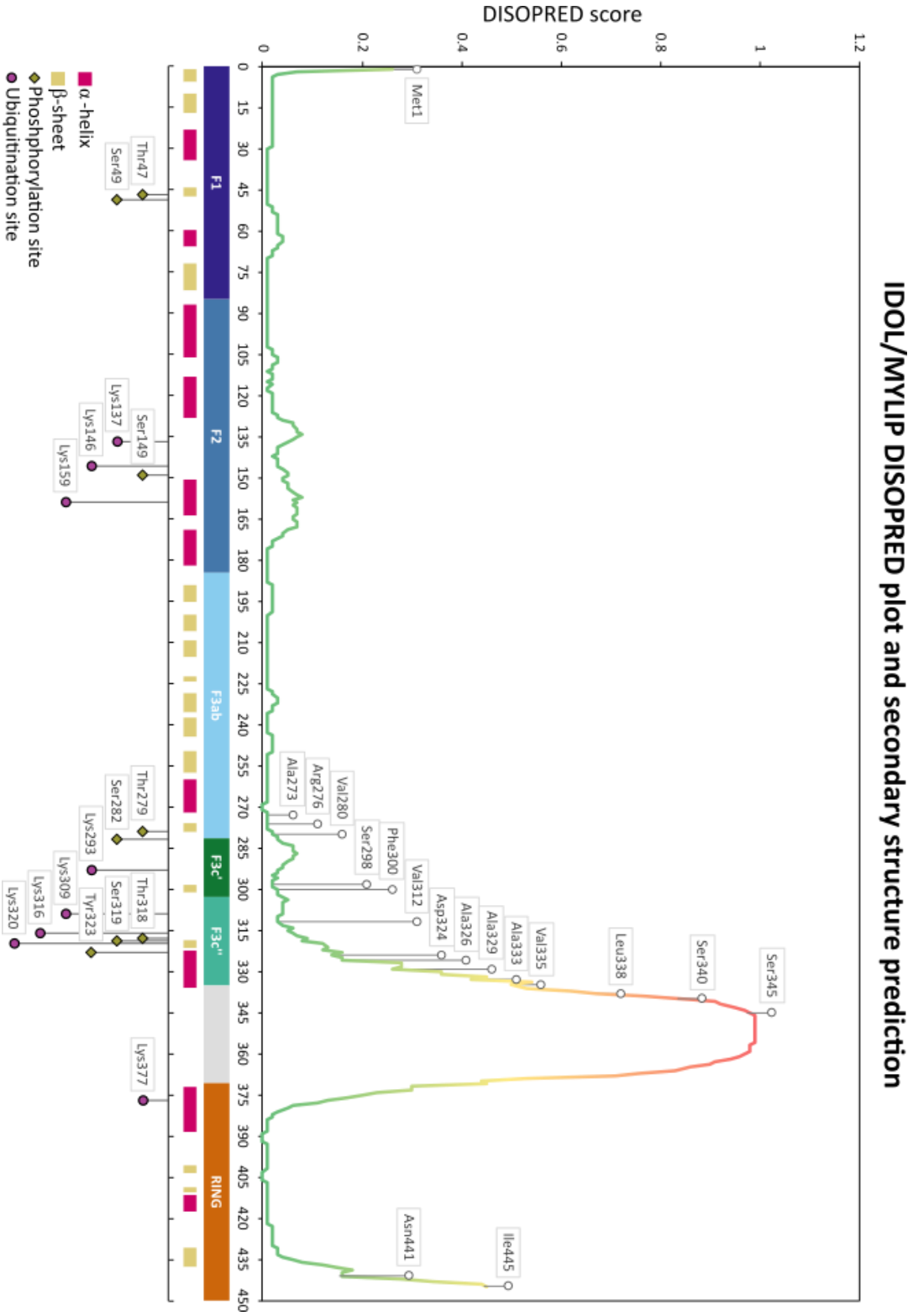


Figure 20: Sequence analysis of MYLIP. DISOPRED plot and secondary structure prediction are shown. Domains are annotated according to currently solved structures[77], [106]. Also indicated are post translational modification sites as accessed from PhosphoSitePlus. MYLIP auto-ubiquitination sites (K293, K309, K316 or K320) were not included in PhosphoSitePlus dataset.

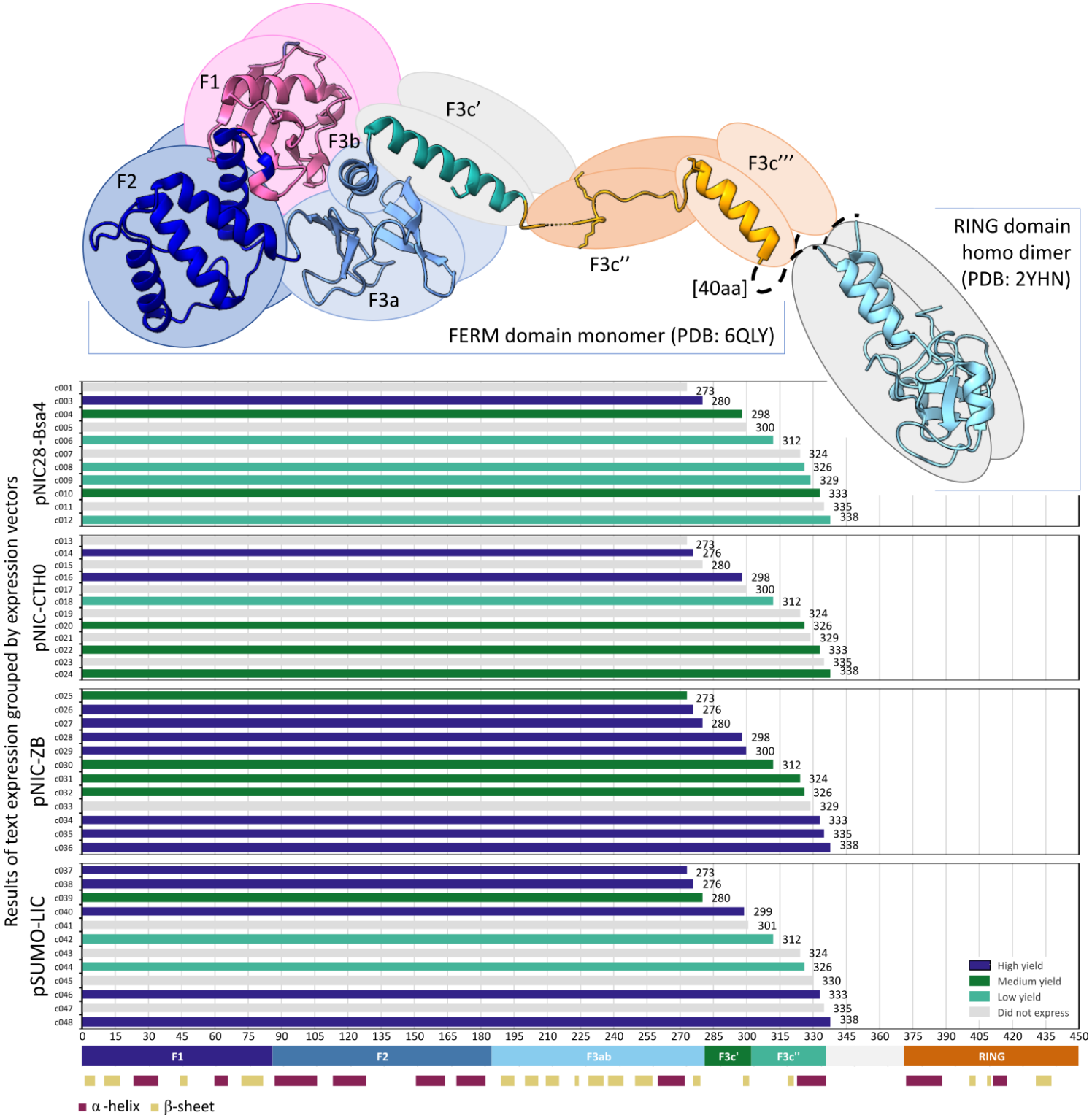


Figure 21: Results of small-scale test expression mapped against structure (top panel) and domain/secondary structure map (bottom panel) of MYLIP. Constructs on the bar graphs are grouped by the expression vector used. Relative expression levels as measured by band intensity are indicated with colours. Dark blue: high yield, dark green: medium yield, light green: low yield, grey: did not express.

Table 16: List of boundaries and vectors used for MYLIP constructs

	Start	Stop	Vectors
1	Met1	Ala273	B, C, Z, S
2	Met1	Arg276	B, C, Z, S
3	Met1	Cys277	B, C, Z, S
4	Met1	Thr279	B, C, Z, S
5	Met1	Val280	B, C, Z, S
6	Met1	Ala283	B, C, Z, S
7	Met1	Ser298	B, C, Z, S
8	Met1	Phe300	B, C, Z, S
9	Met1	Val312	B, C, Z, S
10	Met1	Asp324	B, C, Z, S
11	Met1	Ala326	B, C, Z, S
12	Met1	Ala329	B, C, Z, S
13	Met1	Ala333	B, C, Z, S
14	Met1	Val335	B, C, Z, S
15	Met1	Leu338	B, C, Z, S
16	Met1	Ser340	B, C, Z, S
17	Met1	Ser345	B, C, Z, S
18	Met1	Asn441	B, C, Z, S
19	Met1	Ile445	B, C, Z, S, 3
20	Glu183	S345	B, C, Z, S, 3
21	Glu183	Ile445	B, C, Z, S, 3

B = pNIC28-Bsa4; C = pNIC-CTH0; Z = pNIC-Zb; S = pSUMO-LIC, 3 = pNIC-Bio3, L = pNIC-Bio3-LL

3.2.1.1 Site directed mutagenesis

This work has been carried out by Charline Giroud during my paternity leave. Three constructs were chosen (MYLIP1A-c016 (Met1->Ser 298), MYLIP1A-c029 (Met1->Phe300) and MYLIP1A-c035 (Met1->Val335). One primer pair was designed to introduce two mutations, T279D and S282D, to mimic phosphorylation. Primers, sequencing, and purification results are listed in the appendix.

3.2.2 Full scale MYLIP protein expression and purification

Most proteins that were selected for production expressed well, with only a few being unstable or not expressing at all (**Table 17**). Representative examples of purification results

(chromatograms, SDS-PAGE, thermal unfolding, concentration determination) are shown in this section and all other results are shown in the appendix (**Figure A7** - **Figure A17**). Only two showed modifications or degradation/truncations (MYLIPA-p008: M1-A273, pSUMO-LIC, **Figure 24**; MYLIPA-p014: Ma1-F300, pNIC-ZB, **Figure 25**). Full length MYLIP (MYLIPA-p020, M1-I445, pNIC28-Bsa4, **Figure 26**) showed too many contaminants to perform DSF. Size Exclusion fractions A3 onwards had bands that roughly correlated with the size to be expected of the isolated FERM domain, consistent with reported differential localisation of the domains of MYLIP in cells, hinting at spontaneous cleavage. SDS-PAGE bands highlighted on **Figure 26** were cut out and submitted for tryptic digest and MS/MS at the mass spectrometry facility of our department. Two of the bands were identified as MYLIP, with one having run at ca 50 kDa on the gel (correct full-length construct, 52463.3) and one at 30 kDa on the gel. The other two submitted bands were identified as contaminants (**Table 18**).

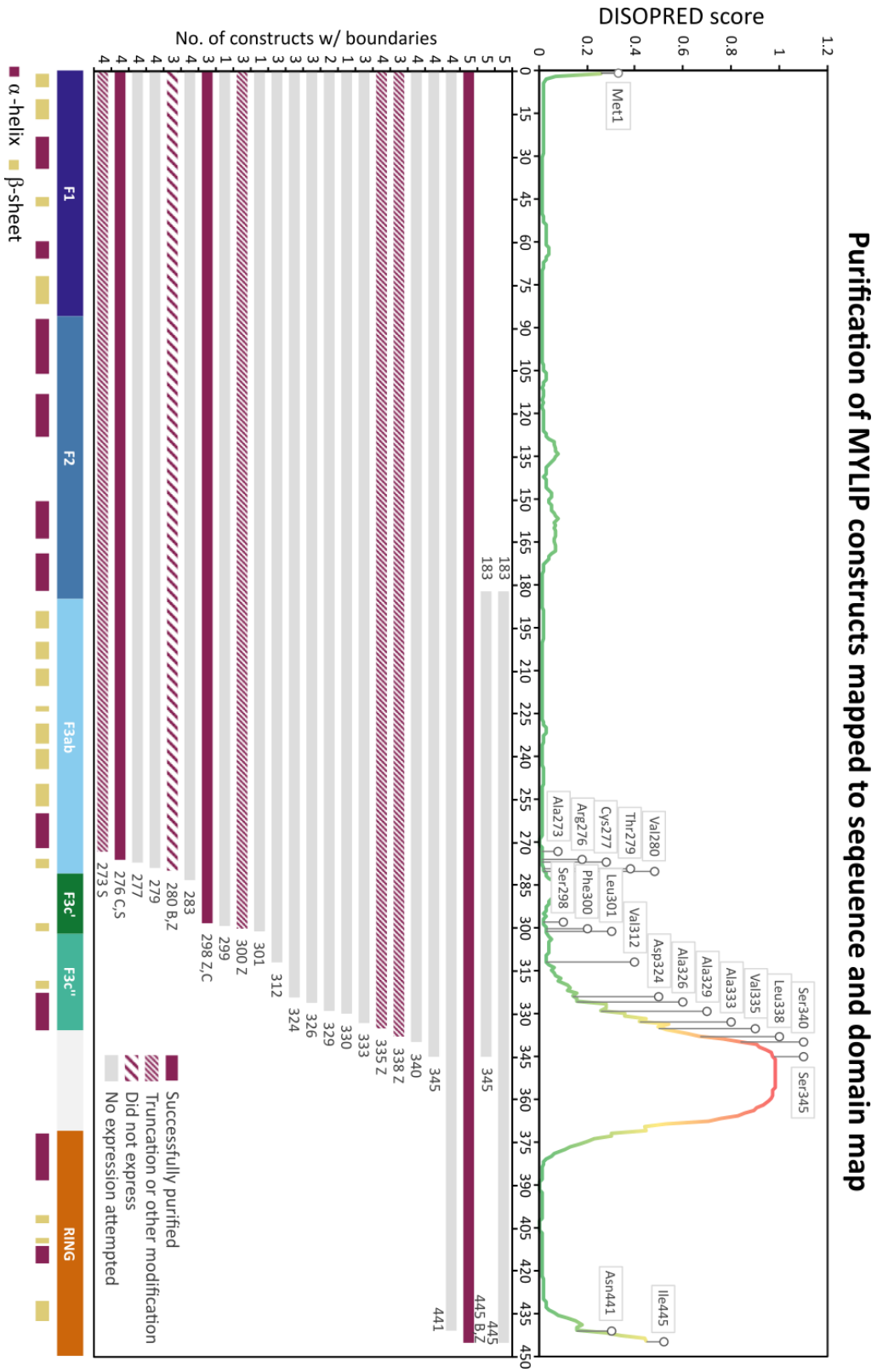


Figure 22: Construct boundaries of designed MYLIP constructs (**Table 16**)mapped to DISPORED plot and domain map (see **Figure 20**). Construct boundaries are shown as bar graph in the middle panel and are colour coded for i. no expression attempted, ii. did not express, iii. expressed but truncated or otherwise modified and iv. successfully purified. For the purified constructs, the expression vectors are indicated (B: pNIC28-Bsa4, C: pNIC-CTH0, Z: pNIC-Zb)

Table 17: Expressed and purified MYLIP constructs.

Purification	Yield (mg/l)	Construct	Boundaries	Vector	Notes
MYLIPA-p001	0	MYLIPA-c003	M1-V280	pNIC28-Bsa4	†
MYLIPA-p002	27.7	MYLIPA -c014	M1-R276	pNIC-CTH0	
MYLIPA-p003	17.5	MYLIPA -c016	M1-S298	pNIC-CTH0	
MYLIPA-p004	12	MYLIPA -c028	M1-S298	pNIC-ZB	❄,™
MYLIPA-p005	8.7	MYLIPA -c029	M1-F300	pNIC-ZB	
MYLIPA-p006	21.9	MYLIPA -c027	M1-V280	pNIC-ZB	
MYLIPA-p007	0	MYLIPA -c038	M1-R276	pSUMO-LIC	†
MYLIPA-p008	11.9	MYLIPA -c037	M1-A273	pSUMO-LIC	™
MYLIPA-p009	0	MYLIPA -c035	M1-V335	pNIC-ZB	‡
MYLIPA-p010	0	MYLIPA -c036	M1-L338	pNIC-ZB	‡
MYLIPA-p011	32.2	MYLIPA -c014	M1-R276	pNIC-CTH0	
MYLIPA-p012	6.3	MYLIPA -c016	M1-S298	pNIC-CTH0	
MYLIPA-p013	11.6	MYLIPA -c027	M1-V280	pNIC-ZB	
MYLIPA-p014	5	MYLIPA -c029	M1-F300	pNIC-ZB	™
MYLIPA-p015	8	MYLIPA -c088	M1-S298	pNIC-CTH0	Ⓟ,©
MYLIPA-p016	0	MYLIPA -c089	M1-F300	pNIC-ZB	Ⓟ,©, †
MYLIPA-p017	0	MYLIPA -c090	M1-V335	pNIC-ZB	Ⓟ,©, †
MYLIPA-p018	5.9	MYLIPA -c088	M1-S298	pNIC-CTH0	Ⓟ
MYLIPA-p019	N/A	MYLIPA -c088	M1-S298	pNIC-CTH0	Ⓟ, d
MYLIPA-p020	2.7	MYLIPA -c049	M1-I445	pNIC28-Bsa4	T
MYLIPA-p022	3.75	MYLIPA -c049	M1-I445	pNIC28-Bsa4	

†: Did not express
‡: Precipitated after tag cleavage
❄: Precipitated after freeze-thaw.
™: Truncation or other modifications.
Ⓟ: Phosphomimetic mutant (T279D-S282D)
©: Expressed and purified by Charline Giroud.
d: dummy entry for tag-cleaved aliquots of MYLIPA-p018.
T: Identity confirmed via LC/MS after tryptic digest.

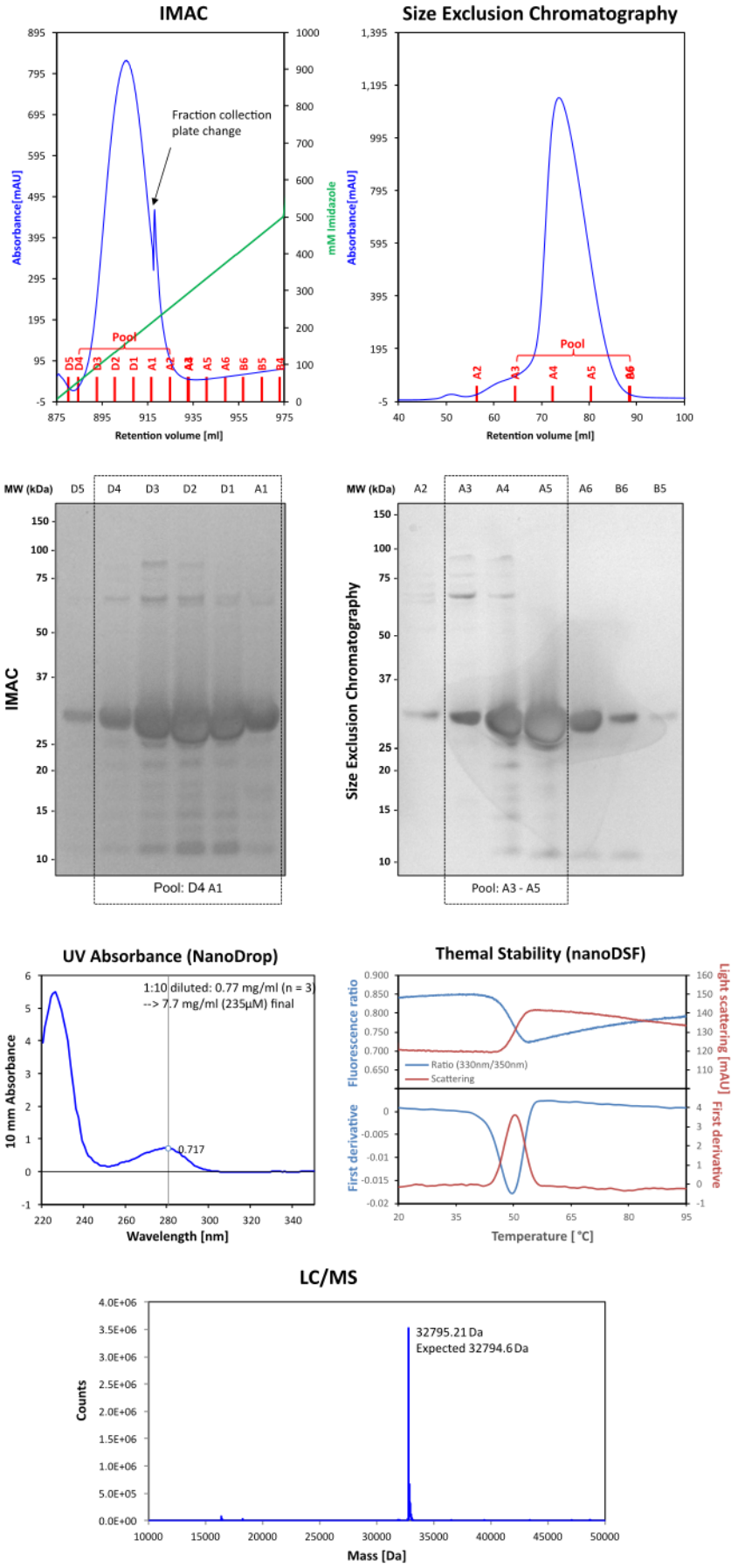


Figure 23: Purification details of MYLIPA-p002(c014:M1-R276;pNIC-CTH0)

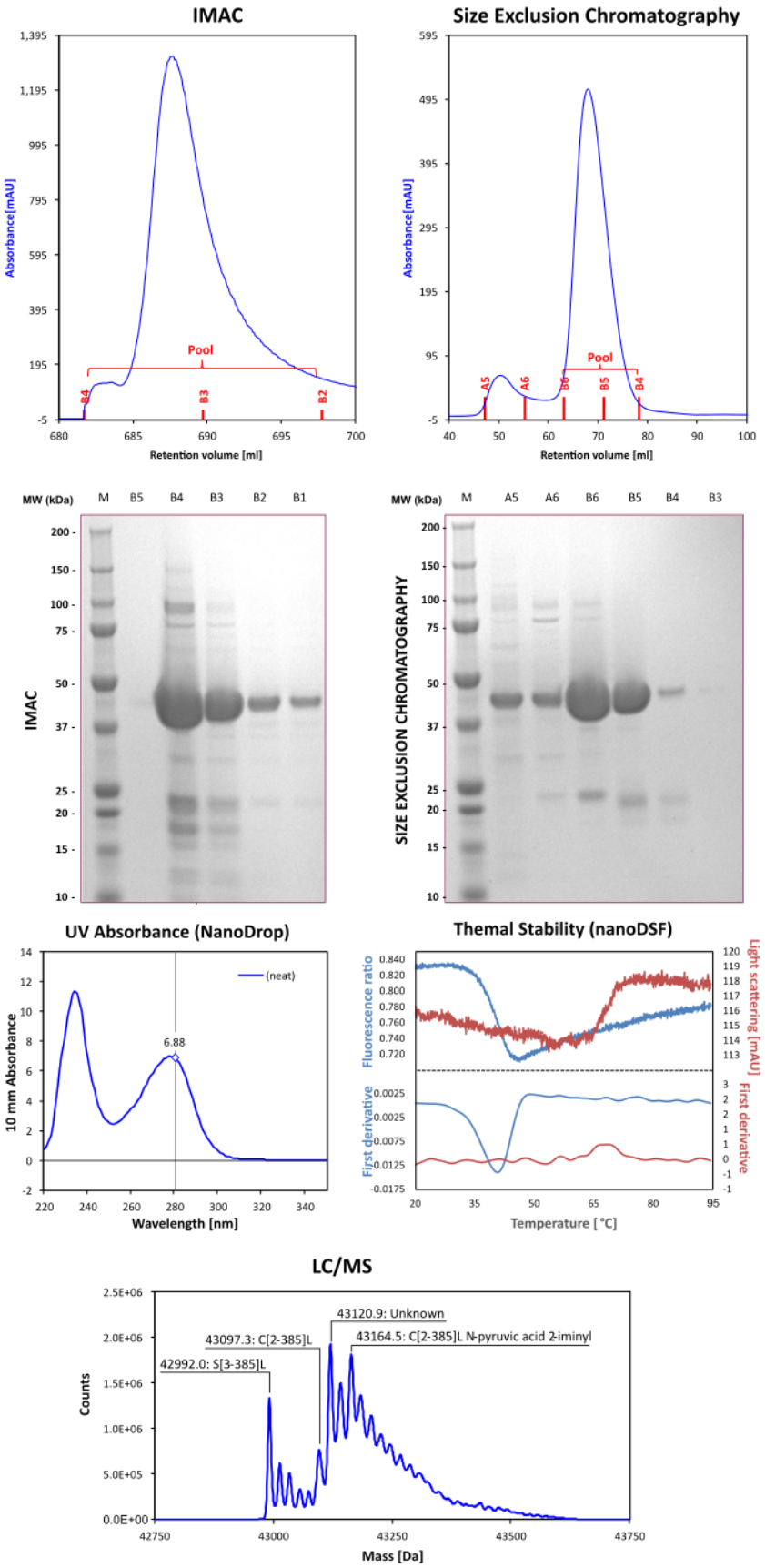


Figure 24: Purification details of MYLIPA-p008(c037:M1-A273;pSUMO-LIC).

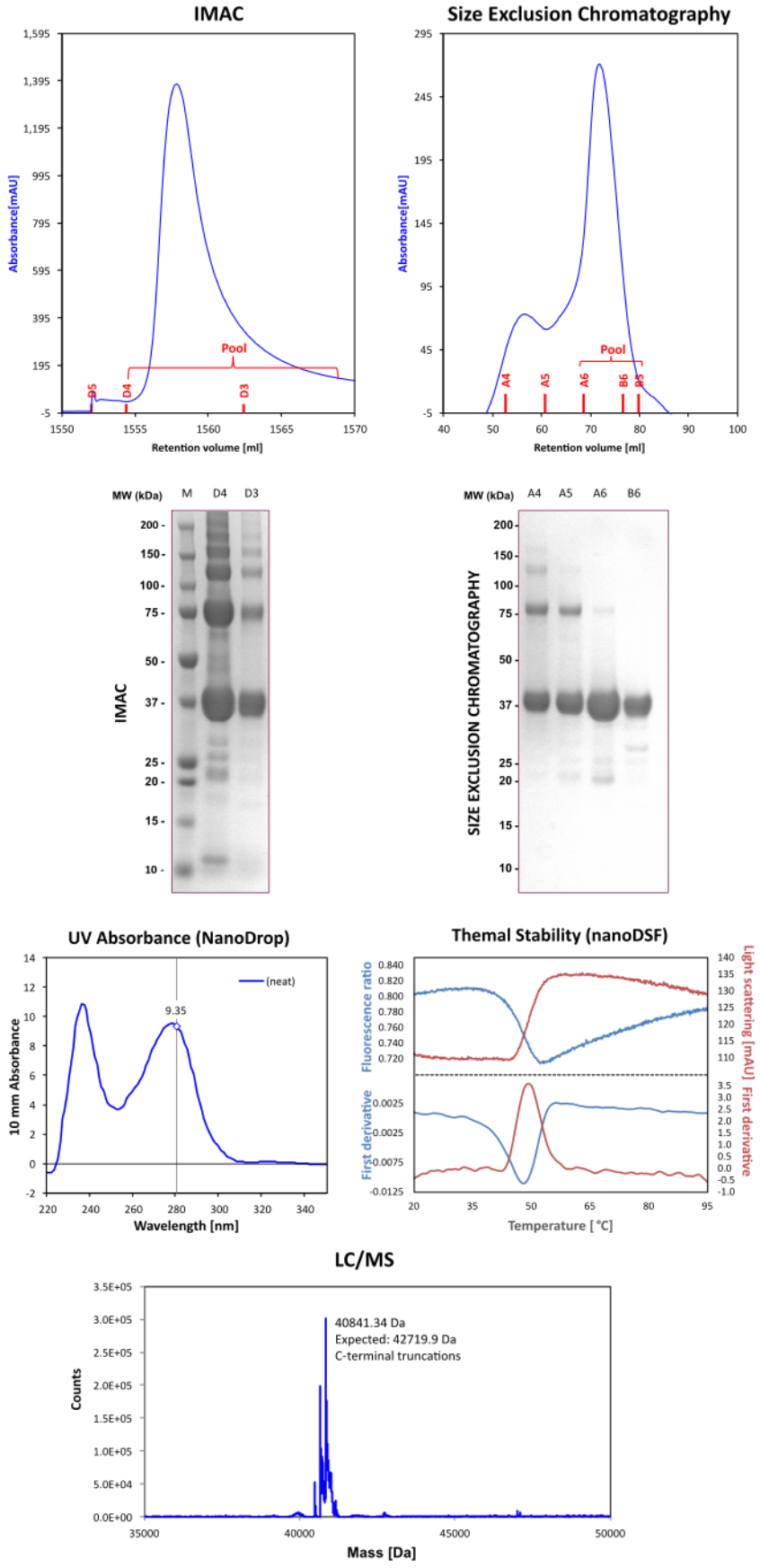


Figure 25: Purification details for MYLIPA-p014(c029:M1-F300;pNIC-ZB).

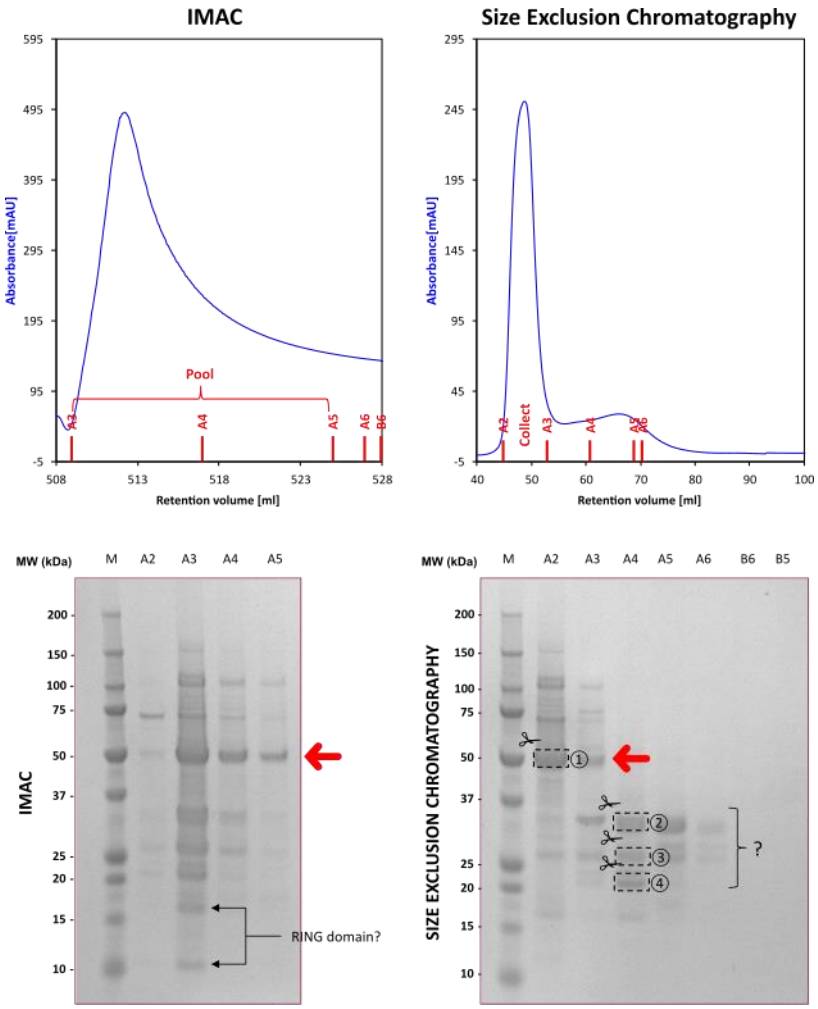


Figure 26: Purification details for MYLIPA-p020(c049:M1-I445;pNIC28-Bsa4). Intact mass spectrometry and DSF were not performed due to low purity. Marked bands on SDS-PAGE after size exclusion were cut out and submitted to mass spectrometry after tryptic digest. Identity for Bands 1 and 2 was confirmed as MYLIP.

Table 18: MASCOT results for submitted bands from MYLIPA-p020 SDS-PAGE after size exclusion chromatography.

Band	Database	Mascot Score	Title	Annotation
1	CMD constructs	209	MYLIPA-c007	Target protein only
	Swissprot sequence (top 5)	188	E3 ubiquitin-protein ligase MYLIP OS=Homo sapiens GN=MYLIP P	
		157	Trypsin OS=Sus scrofa PE=1 SV=1	
		107	Anionic trypsin-1 OS=Rattus norvegicus GN=Prss1 PE=1 SV=1	
		64	ANIONIC TRYPSIN-1 OS=RATTUS NORVEGICUS GN=PRSS1 PE=1 SV=1	
		61	Bifunctional protein GlmU OS=Streptococcus pneumoniae seroty	
2	CMD constructs	153	MYLIPA-c001	
	Swissprot sequences (top 5)	153	E3 ubiquitin-protein ligase MYLIP OS=Homo sapiens GN=MYLIP P	
		135	Trypsin OS=Sus scrofa PE=1 SV=1	
		70	Tyrosine--tRNA ligase OS=Thermococcus kodakarensis (strain A	
		50	WD repeat-containing protein 44 OS=Rattus norvegicus GN=Wdr4	
		48	Heat-inducible transcription repressor HrcA OS=Lactococcus l	
3	CMD constructs	70		Contaminating proteins only, target is not detected
	Swissprot sequences (top 5)	80	Peptidyl-prolyl isomerase cwc27 OS=Emericella nidulans (stra	
		80	30S ribosomal protein S3 OS=Photorhabdus luminescens subsp.	
		68	FKBP-type peptidyl-prolyl cis-trans isomerase SlyD OS=Escher	
		65	Cell division protein FtsZ OS=Francisella tularensis subsp.	
		62	CELL DIVISION PROTEIN FTSZ OS=FRANCISELLA TULARENSIS SUBSP.	
4	CMD constructs	49	JMJD3A-c028	Contaminating proteins only, target is not detected
	Swissprot sequences (top 5)	248	cAMP-activated global transcriptional regulator CRP OS=Esche	
		170	Trypsin OS=Sus scrofa PE=1 SV=1	
		73	TRYPSIN OS=SUS SCROFA PE=1 SV=1	
		56	ARGININOSUCCINATE LYASE OS=STREPTOCOCCUS UBERIS (STRAIN ATC	
		56	CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR SUBUNIT 1 O	

3.2.3 Ubiquitin cascade enzyme production

3.2.3.1 Ubiquitin activating enzyme (E1)

Three purifications were performed in total. During the IEX step of the first purification, the air sensor on the AKTA system developed a fault and air was pumped over the column. Surprisingly, this batch was still active (**Figure 53**). The third purification yielded two peaks with the same protein eluting at different conductivities: Same molecular weight estimated by SDS-PAGE (**Figure 30**, 120 kDa). The first peak is most likely aggregated protein, and the second peak was chosen for validation and use in the assays.

Table 19: Summary of murine E1 purifications.

Batch	Total	Concentration	mM NaCl at peak	mS/cm at peak
MUBE1A-p001	42 mg	5.2 mg/ml	265 mM	2.34
MUBE1A-p002	8.4 mg	5.23 mg/ml	300 mM	14.81
MUBE1A-p003.1	1.1 mg	2.8 mg/ml	110 mM	2.47
MUBE1A-p003.2	6.2 mg	2.1 mg/ml	210 mM	10.2

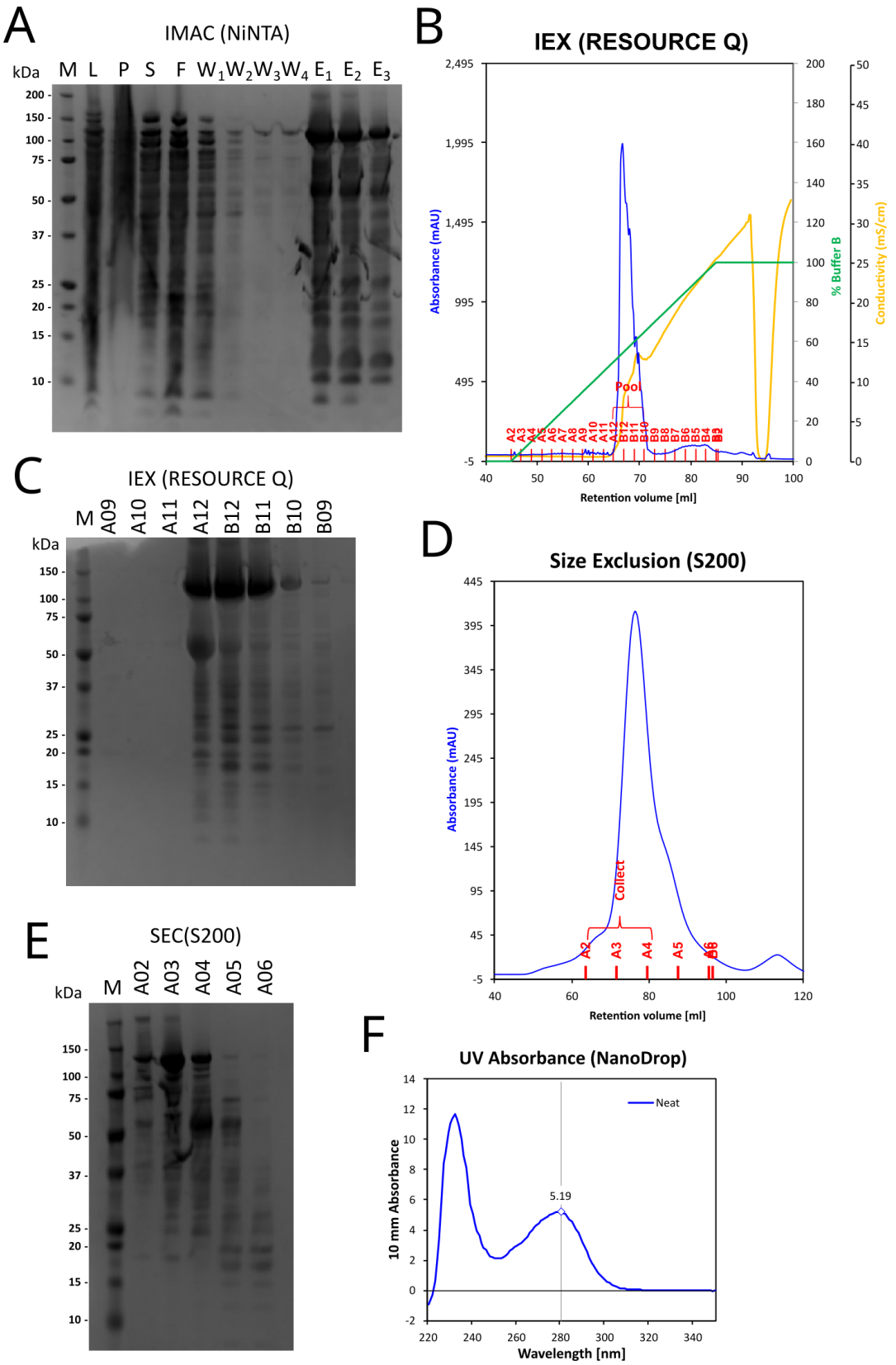


Figure 27: Purification details of MUBE1A-p001. **A**,**C**,**E**: SDS-PAGEs after chromatography steps. M: marker. L: Lysate, S: Supernatant, P: Pellet, F: Flow through, W_{1-4}: washes. E_{1-3}: Elution. **B**, **D**: Chromatograms of chromatography steps. Pooled fractions indicated. **F**: Nanodrop UV spectrum of pooled and concentrated final fractions.

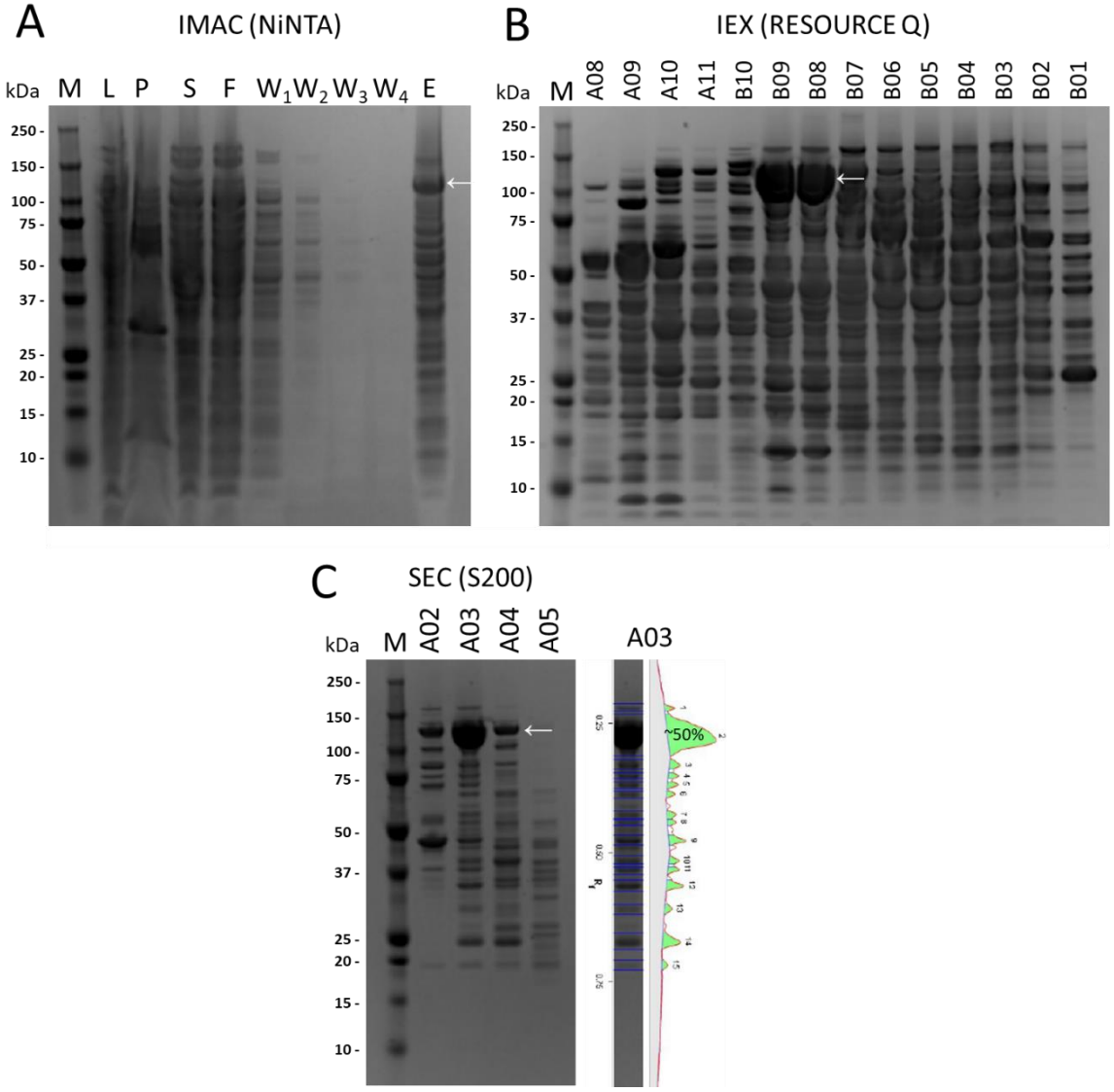


Figure 28: SDS-PAGEs of MUBE1A-p002 chromatography steps. **A:** IMAC. M: Marker, L: Lysate, P: Pellet, S: Supernatant, F: Flow-through. W_{1-4}: Washes. E: Elution with 100 mM imidazole. **B:** Ion-exchange chromatography with Resource Q. **C:** Size Exclusion chromatography with S200, Lane for A03 was analysed using the Bio-Rad imager software to detemine approximate purity of the protein.

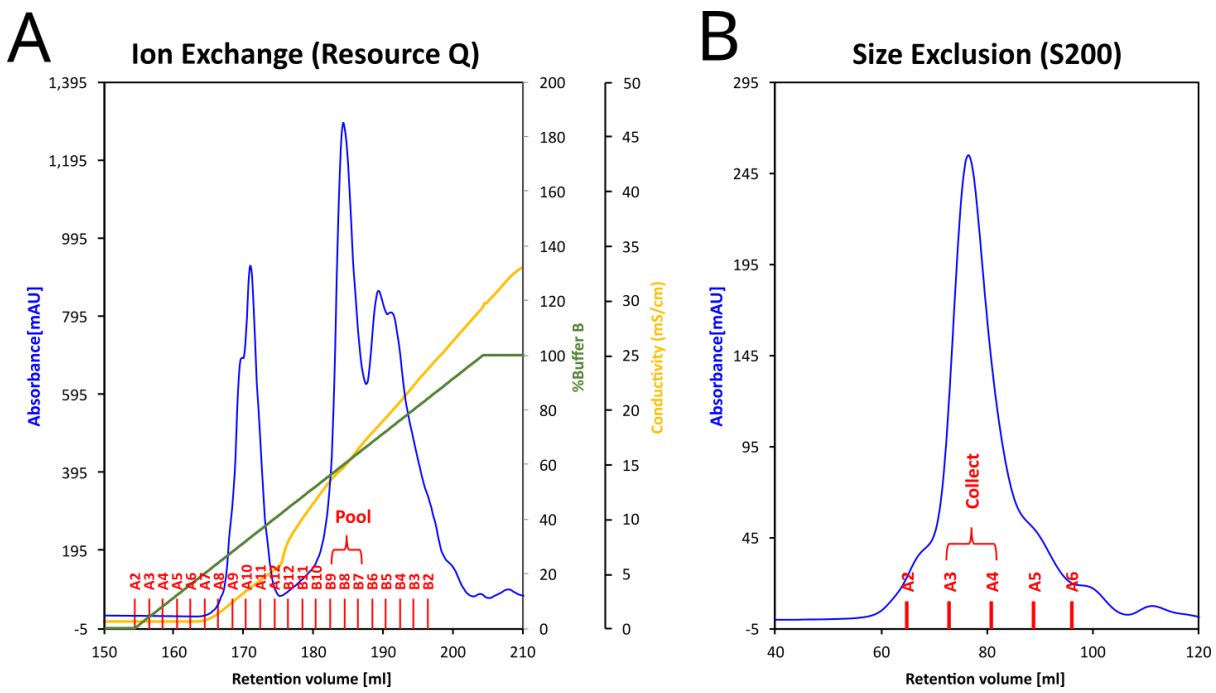


Figure 29: Chromatograms for MUBE1A-p002 purification. **A:** Ion exchange chromatography (Resource Q) **B:** Size exclusion chromatography. Pooled fractions are indicated in each chromatogram.

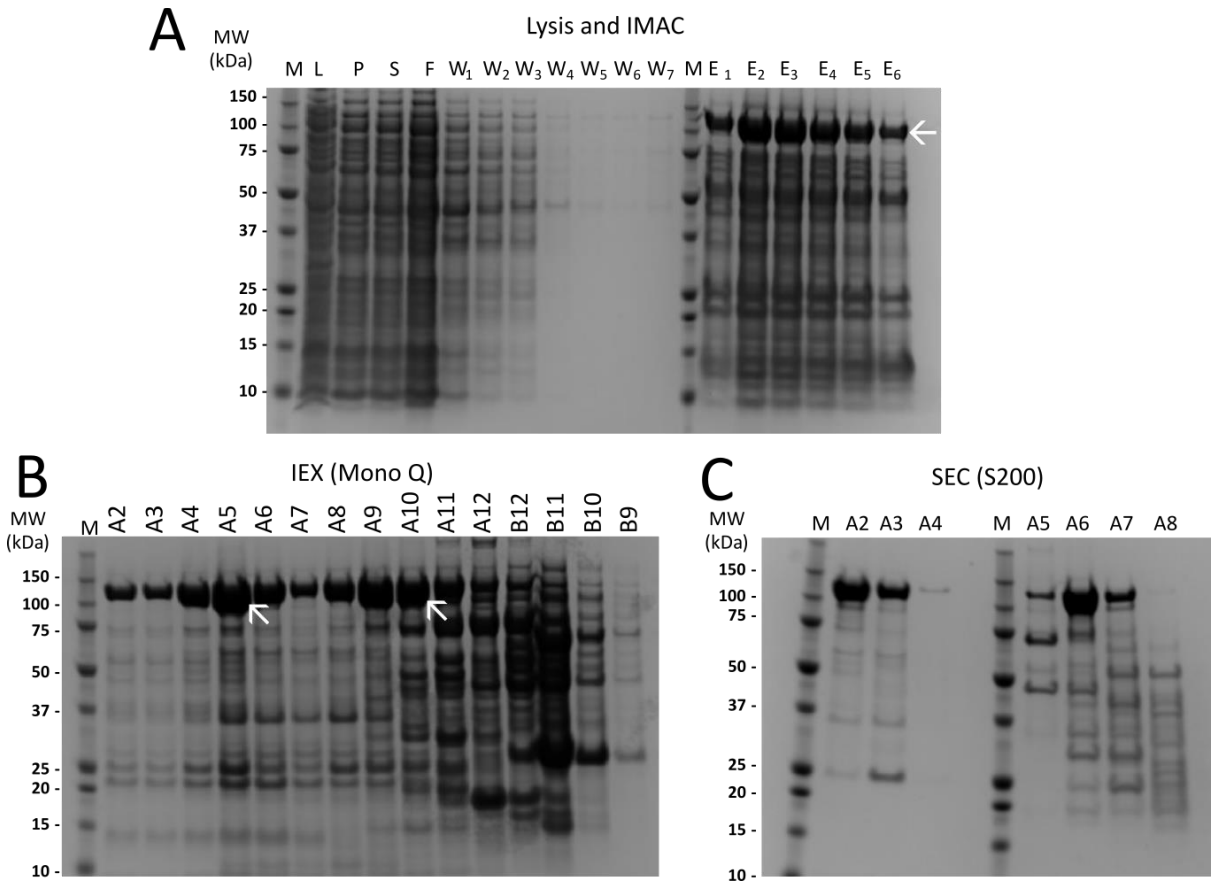


Figure 30: SDS-PAGES after chromatography steps for MUBE1A-p003. **A:** Lysis and IMAC. M: Marker. L: Lysate. P: Pellet. S: Supernatant. F: Flow-through. W_{1-7}: Washes. E_{1-6}. **B:** Ion-exchange chromatography (MonoQ). **C:** Size exclusion chromatography (S200). Bands corresponding to expected mass of 120 kDa are indicated with white arrow on **A**, **B**.

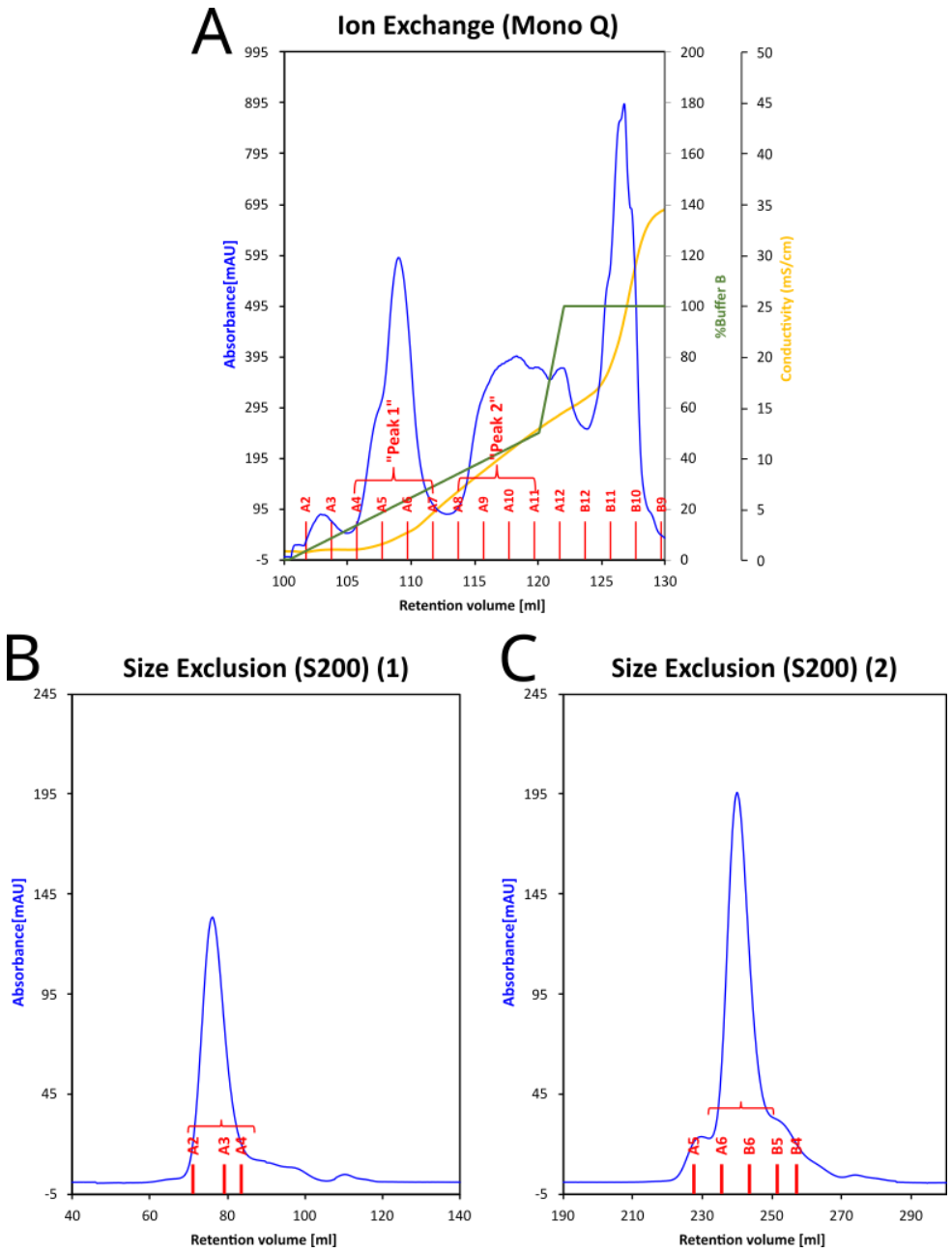


Figure 31: Chromatograms from MUBE1A-p003 purification. **A:** Ion Exchange Chromatography with MonoQ column. **B**, **C**: Size exclusion chromatography of two peaks collected from IEX.

3.2.3.2 Ubiquitin-conjugating enzymes (E2)

All three E2 constructs (UBE2D1A-c001, UBE2D2A-c001 and UBE2D3A-c001 expressed at excellent purity (**Figure 32**, **Figure 33** and **Figure 34**)) and good yields **Table 20**.

Table 20: Summary of E2 purification yields.

Purification ID	Total yield (mg)	Concentration (mg/ml)
UBE2D1A-p001	25.0	3.3
UBE2D2A-p001	29.4	7.0
UBE2D3A-p001	10.6	10.0

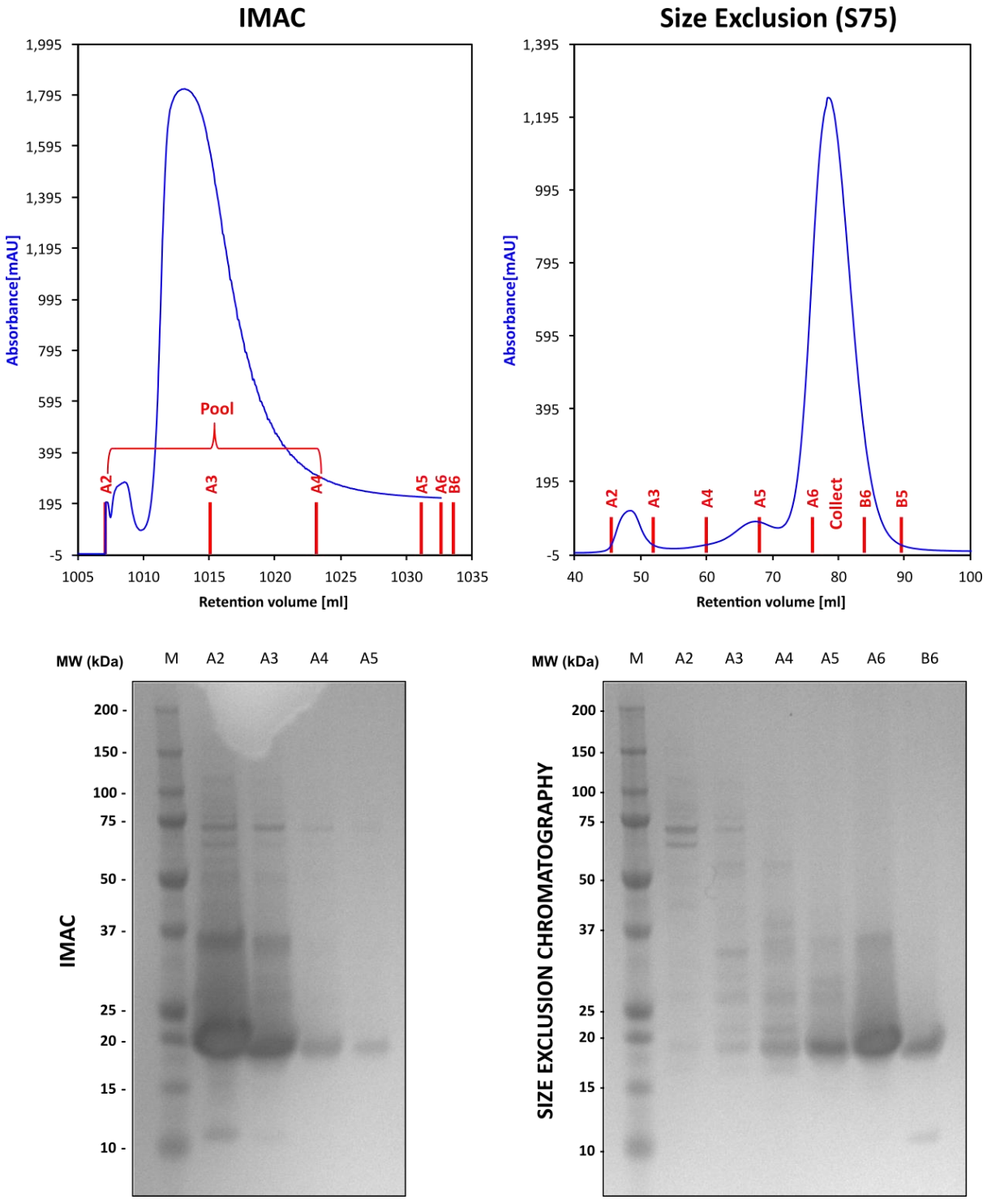


Figure 32: UBE2D1A-p001 purification details: IMAC and Size exclusion chromatography (S75) chromatograms and SDS-PAGE analysis of collected fractions. Pooled fractions are indicated.

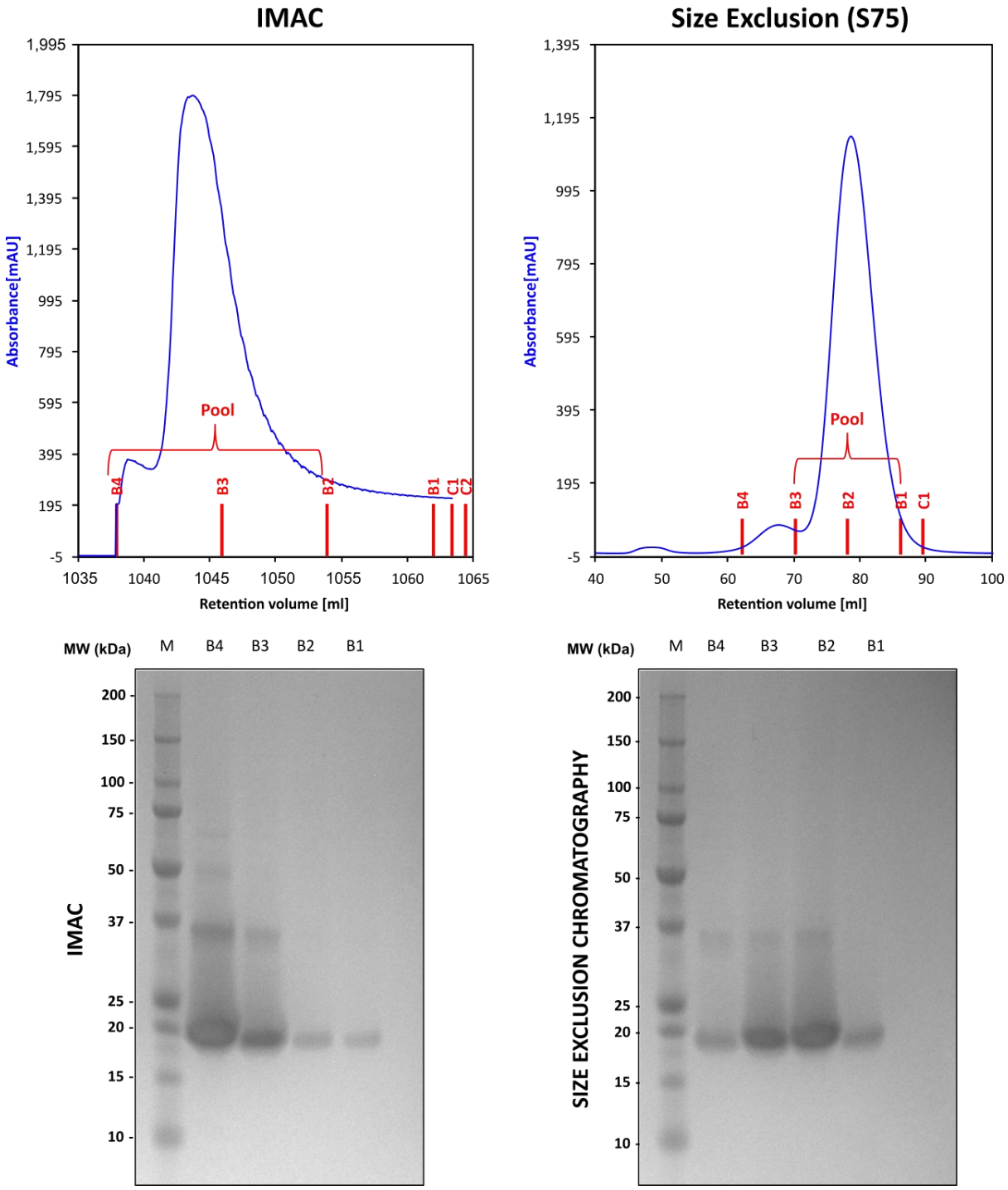


Figure 33: Purification details for UBE2D2A-p001: IMAC and Size exclusion chromatography (S75) chromatograms and SDS-PAGE analysis of collected fractions. Pooled fractions are indicated.

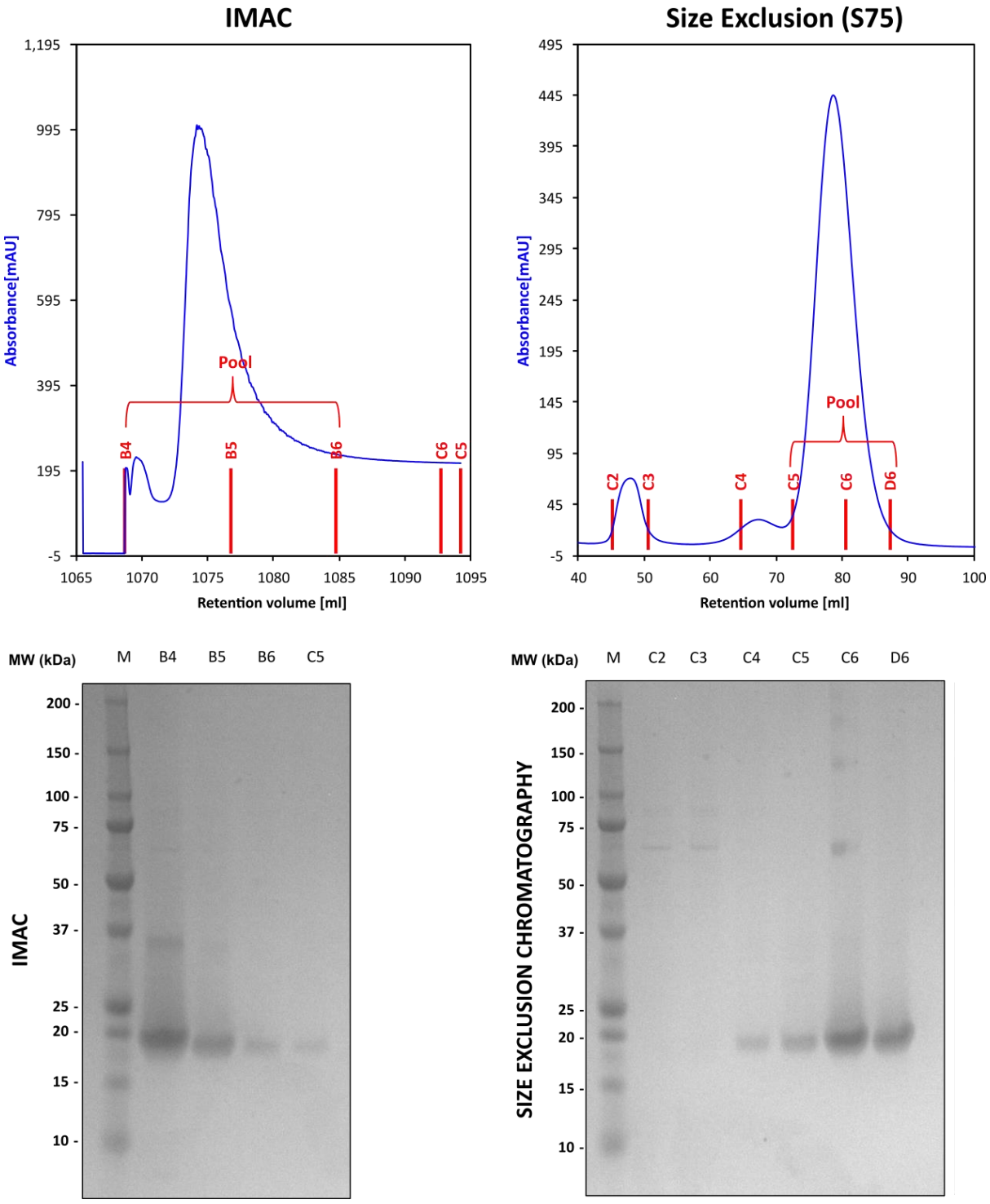


Figure 34: Purification details of UBE2D3A-p001: IMAC and Size exclusion chromatography (S75) chromatograms and SDS-PAGE analysis of collected fractions. Pooled fractions are indicated.

3.2.3.3 Ubiquitin

The human mono-ubiquitin gifted by Mike Fairhead expressed well. Cleavage of the affinity tag was attempted for the first ubiquitin batch. Tag cleavage (first erroneously with TEV, then Thrombin) was unsuccessful and the batch was discarded. The second batch was expressed and purified without attempting to cleave the tag. Following SDS-PAGE, washes W_2 and W_3 were combined with the elution fractions. A total of 50.8 mg of protein were obtained from two litres of culture. It was concentrated to a final concentration of 13.36 mg/ml or 922 µM at excellent purity (**Figure 35**). The intact mass was 14346.6 Da, consistent with a loss of N-terminal methionine. RPS27AA-c001 also expressed extremely well. A total protein concentration of 13 mg/ml or 1.2 mM. The previous His_6-TwinStrep-mono-ubiquitin was also tested, and the concentration was 12.23 mg/ml, compared to 13.4 mg/ml as determined by A_{280}.

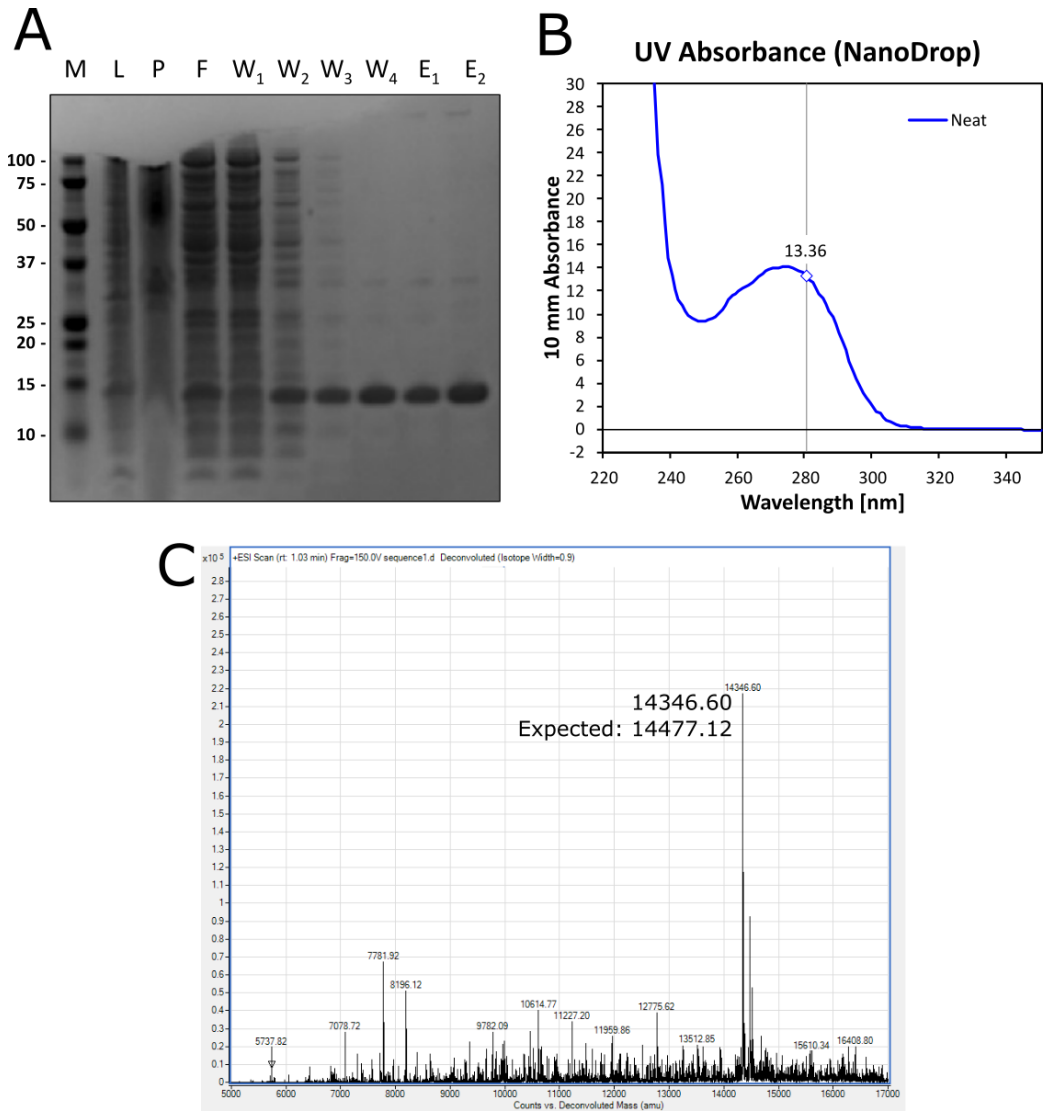


Figure 35: Details of second human mono-ubiquitin purification. **A:** SDS-PAGE (M: Marker; L: Lysate; P: Pellet; F: flow-through; W_1,W_2:) **B:** Nanodrop UV vis spectrum **C:** Deconvoluted MS spectrum

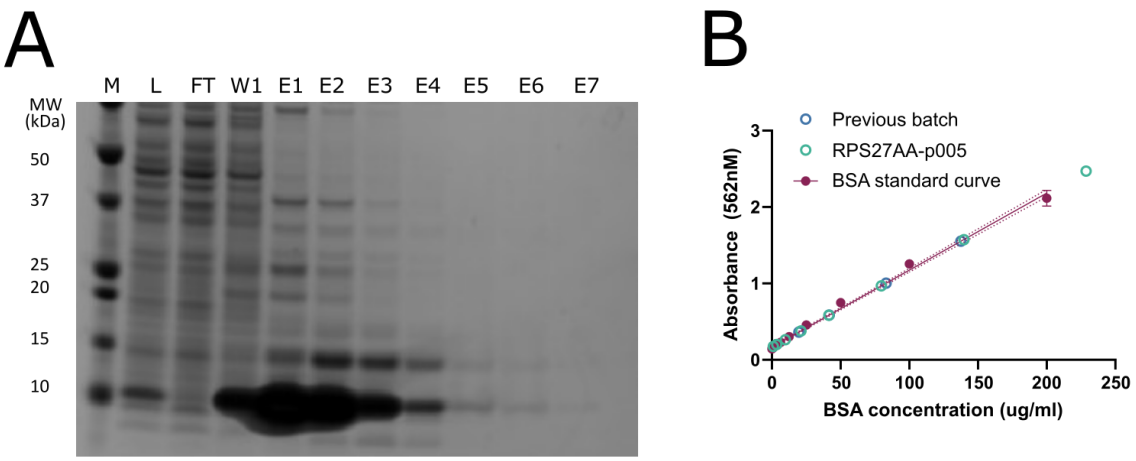


Figure 36: Purification details for RPS27AA-p005. **A:** SDS-PAGE of LYSIS and IMAC. M: Marker, L: Lysate, FT: Flow-through, W_1: Wash. E_{1-7}: Elutions. **B:** Determination of concentration via BCA assay. RPS27AA-p005: 12.65 mg/ml. Previous His6-TwinStrep-mono-ubiquitin: 12.23 mg/ml.

3.2.4 LDLR adapter protein construct design and production

3.2.4.1 Construct design

One set of boundaries was chosen for DAB1 and DAB2, based on solved crystal structures of the mouse homolog (1NTU, 1NTV, and 1M7E, respectively) (Appendix **Table A8** - **Table A12**). This work was carried out by Charline Giroud during my paternity leave.

3.2.4.2 Protein expression and purification

LDLRAP1A-c001 and LDLRAP1A-c002 did unfortunately not express. The DAB constructs tested did express well and were of correct mass. DAB1A-p003/p004 (DAB1A-c002) showed loss of N-terminal methionine. (Appendix **Figure A18** - **Figure A29**)

3.2.5 Summary and Discussion

Pleasingly, most of the constructs that were initially chosen for production based on the test expression yields expressed well and only a few suffered from truncations or were not stable enough to survive an initial freeze-thaw cycle without precipitation (**Table 17**). When the relative test expression yields were grouped by expression vector and mapped structure/domain map of full length MYLIP (**Figure 21**), a few trends are discernible:

i. pNIC28-Bsa4 is the least efficient vector used in this work.

ii. pNIC-CTH0, with its C-terminally located His6 tag, favours constructs with either entirely missing or complete F3c'' region. Regions F3c' and F3c'' are comprised of alpha-helices that are not tightly packed against the clover-leaf structure of the canonical FERM domain.

iii. pNIC-ZB with its highly charged ZBasic tag is highly effective at stabilizing the FERM domain constructs.

iv. pSUMO-LIC gives expression across the entire range of the FERM domain, if not all constructs, but also seems to favour constructs with either entirely missing or complete F3c'' region.

Martinelli et al.s SAXS data and crystal structure indicate that the F3c'' a-helix is folded over the peptide binding site comprised of F3ab and F3c'. In this case, it is conceivable that a fully intact or completely absent F3c'' helix, as opposed to a partial helix, is more favourable for the stability of the protein. In the case of the C-terminal affinity tags, steric hindrance due to the partially folded over helix may well affect the ability to purify the protein, even if it does express.

Considering the bias towards the pNIC-ZB vector, it is worth noting that the proteins that precipitated after tag cleavage were expressed in the pNIC-ZB vector, and constructs spanning residues 1-335 were not stable in any of the other vectors, indicating a general instability issue with these boundaries. Constructs spanning residues 1-338 were stable in the other vectors, with low, moderate, and high levels of protein in the test expressions for pNIC28-Bsa4, pNIC-CTH0 and pSUMO-LIC, respectively. It is also worth noting that the text expression protocol uses DDM to aid lysis and this could artificially inflate the apparent yield of proteins with sub-optimal solubility.

All remaining proteins/enzymes of the ubiquitin cascade (E1, E2s, ubiquitin) seemed to express well with good yields and purities.

Table 21: Expressed and purified phosphotyrosine binding domain constructs.

Purification	(mg/ml)	Construct	Boundaries	Vector	Notes
LDLRAP1A-p001					†
LDLRAP1A-p002					†
DAB1A-p001	15.54	DAB1A-c001	23-174	pNIC-Bsa4	
DAB1A-p002	0.95	DAB1A-c001	23-174	pNIC-Bsa4	X1
DAB1A-p003	21.6	DAB1A-c002	23-174	pNIC-CTH0	
DAB1A-p004	6.83	DAB1A-c002	23-174	pNIC-CTH0	X2
DAB2A-p001	9.2	DAB2A-c001	33-181	pNIC-Bsa4	
DAB2A-p002	1.57	DAB2A-c001	33-181	pNIC-Bsa4	X3
DAB2A-p003	10.94	DAB2A-c002	33-181	pNIC-CTH0	
DAB2A-p004	10.53	DAB2A-c002	33-181	pNIC-CTH0	X4

†: Did not express
X1: TEV cleaved portion of DAB1A-p001
X2: TEV cleaved portion of DAB1A-p003
X3: TEV cleaved portion of DAB2A-p001
X4: TEV cleaved portion of DAB2A-p003

3.3 Binding of MYLIP FERM domain to LDLR family receptor tail peptides

The following section shows the results of the protein-peptide titrations in the protein-protein interaction assays described in section 2.12.1. with the goal of recapitulating the published MYLIP FERM domain – receptor tail interaction. True binding could not be observed for any of the protein-peptide combinations investigated, either for the wild type MYLIP FERM domain or the phosphomimetic mutant (**Table 22**). It was then attempted to investigate possible co-operative effects of MYLIP and the phosphotyrosine binding domains (PTBs) of Dab-1 and Dab-2 after verifying that they bind to the receptor tail peptides. While the PTBs bound the tested receptor tail peptides, no co-operative effect could be observed with either wild type FERM domains or the phosphomimetic mutant.

If not indicated otherwise, all experiments were performed with the standard buffers listed in materials and methods. For all experiments, longest incubation time (upper left) data is shown; O/N refers to overnight.

Table 22: Overview of protein-peptide-titrations in plate-based assay technologies.

	FERM domain boundaries (and vector)				
	M1-R276	M1-S298	M1-A273	M1-S298	M1-S298^{phos}
Peptide	pNIC-CTH0	pNIC-CTH0	pSUMO-LIC	pSUMO-LIC	pNIC-CTH0
LDLR$_{(811-859)}$-K-btn	A^{p}	H^{x}/A^{p}			
LDLR$_{(811-834)}$-K-btn	H^{x}		H^{x}	H^{x}	
Cy5.5-LDLR$_{(811-833)}$		F^{x}			
btn-VLDLR$_{(820-842)}$	A^{x}	H^{x}			H^{x}
btn-ApoER2$_{(849-871)}$	A^{p}	H^{x}			
btn-APP$_{(745-767)}$	A^{x}	H^{x}			

H: HTRF, A: AlphaScreen, F: Fluorescence polarisation

^{x}: negative result, ^{P}: positive result

^{phos}: Phosphomimetic mutant

3.3.1 Titrations of MYLIP constructs against LDLR family receptor tails in HTRF, AlphaScreen and Fluorescence polarisation

3.3.1.1 HTRF titrations of MYLIP constructs against LDLR$_{(811-859)}$-K-biotin

A

1h

LDLR(811-859)-K-biotin (nM)

MYLIPA(1-298 C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	0.4883	0.2441	-
4000	131	135	129	131	129	124	127	125	128	131	125	131	128	128	130	119
2000	135	129	132	127	127	123	127	126	119	123	122	125	127	123	135	128
1000	141	134	124	134	126	125	131	123	133	129	120	125	129	129	127	124
500	148	131	135	131	140	122	143	131	130	139	124	129	131	126	133	121
250	157	140	145	135	140	123	147	136	126	134	131	131	130	125	129	124
125	158	147	149	151	155	140	140	136	134	134	129	131	131	128	132	130
62.5	168	150	150	139	155	143	153	141	135	138	130	130	135	128	130	125
31.25	157	152	147	155	159	141	153	144	132	133	129	141	131	127	131	130
15.625	152	148	150	149	151	137	151	146	148	135	137	136	134	132	132	129
7.8125	150	149	142	147	153	124	146	139	141	142	127	129	125	131	131	130
3.90625	145	134	142	138	142	132	146	136	131	134	127	135	133	130	131	124
1.953125	140	141	135	142	140	129	138	129	133	128	128	126	127	132	130	124
0.976563	130	141	133	136	136	131	134	132	128	136	132	130	130	127	134	132
0.488281	135	141	135	138	138	132	136	128	131	136	127	131	127	124	131	128
0.244141	135	135	131	130	137	129	133	123	127	131	128	134	138	124	133	124
-	145	137	142	136	138	131	136	127	128	135	127	134	128	134	132	120

B

O/N

LDLR(811-859)-K-biotin (nM)

MYLIPA(1-298 C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	0.4883	0.2441	-
4000	169	200	191	182	184	182	178	194	187	180	175	166	175	168	168	166
2000	195	171	194	173	175	176	176	185	173	172	175	169	191	168	175	170
1000	181	170	181	174	174	181	177	191	171	172	173	158	171	162	167	172
500	188	176	184	178	179	182	180	181	177	181	168	159	169	169	167	170
250	178	170	193	175	175	191	178	187	177	169	169	160	162	162	170	182
125	189	178	196	193	184	158	188	200	192	180	181	162	168	188	167	182
62.5	182	176	194	187	190	196	187	201	181	175	174	162	165	156	167	185
31.25	201	189	209	200	199	198	193	205	186	173	183	161	170	162	168	179
15.625	204	242	202	192	193	197	190	203	189	174	174	163	167	163	167	170
7.8125	192	184	199	195	191	190	191	200	178	177	176	155	167	165	162	177
3.90625	192	179	199	187	186	187	194	196	178	177	176	166	162	158	168	173
1.953125	177	179	183	177	173	174	206	187	174	171	171	155	170	163	164	175
0.976563	180	166	181	159	160	168	164	172	169	173	166	163	158	155	165	169
0.488281	172	162	173	172	165	158	160	175	151	171	175	157	159	154	164	175
0.244141	171	168	170	158	157	163	160	162	159	161	162	158	160	157	159	169
-	178	162	190	162	158	154	168	183	146	204	175	169	170	175	178	184

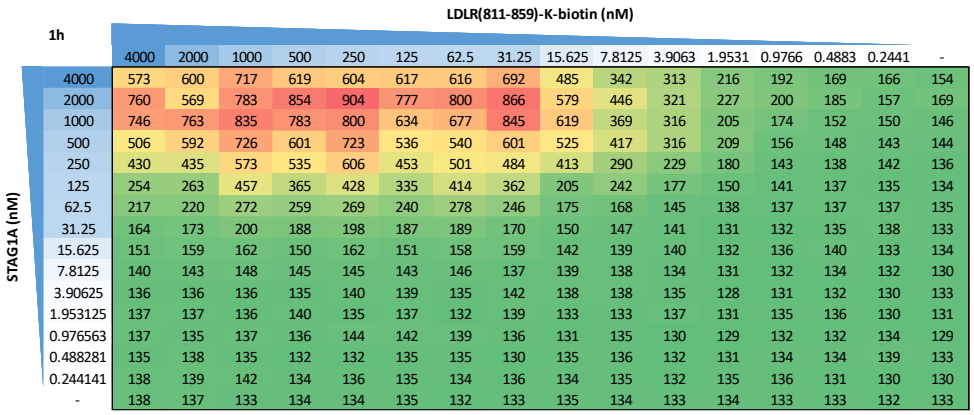
C

O/N

LDLR(811-859)-K-biotin (nM)

MYLIPA(1-296 N-His6-ZB) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	0.4883	0.2441	-
4000	581	384	408	354	389	400	384	392	381	336	296	258	255	247	267	270
2000	434	436	495	481	507	549	552	609	501	361	282	232	226	198	226	223
1000	573	554	587	550	636	660	717	765	513	363	242	199	184	178	186	190
500	529	569	590	613	648	686	731	617	411	308	197	179	175	170	178	191
250	405	473	466	495	535	549	590	545	358	263	186	173	167	162	175	180
125	334	394	455	213	473	496	496	458	307	214	184	171	172	175	167	173
62.5	294	309	368	390	397	387	386	348	248	193	174	164	170	164	176	178
31.25	260	298	310	333	348	323	345	281	208	178	174	164	167	166	173	176
15.625	238	253	279	272	271	266	272	233	205	174	168	163	160	162	169	179
7.8125	221	248	252	275	253	238	239	219	184	172	169	170	165	162	169	176
3.90625	202	210	217	219	220	212	209	190	174	171	163	156	164	163	167	182
1.953125	197	200	195	193	200	198	192	182	171	163	161	163	164	162	164	177
0.976563	180	172	188	174	190	180	178	175	162	159	165	165	155	164	167	176
0.488281	163	175	180	190	177	173	169	180	166	171	161	164	167	165	169	173
0.244141	183	173	180	168	170	178	183	172	171	174	172	165	167	165	167	173
-	178	174	167	160	165	171	158	163	173	183	179	172	176	178	179	180

Figure 37: HTRF titration of LDLR$_{(811-859)}$-K-biotin against MYLIP. **A:** MYLIPA$_{(1-298\ C\text{-}His6)}$ in sodium phosphate buffer (20 mM sodium phosphate, pH 7.5, 100 mM NaCl, 0.05% TWEEN-20, 0.1 mg/ml BSA). **B:** MYLIPA$_{(1-298\ C\text{-}His6)}$. **C:** MYLIPA$_{(1-296\ N\text{-}His6\ ZB)}$. **D:** MYLIPA$_{(1-298\ N\text{-}His6\ ZB)}$.

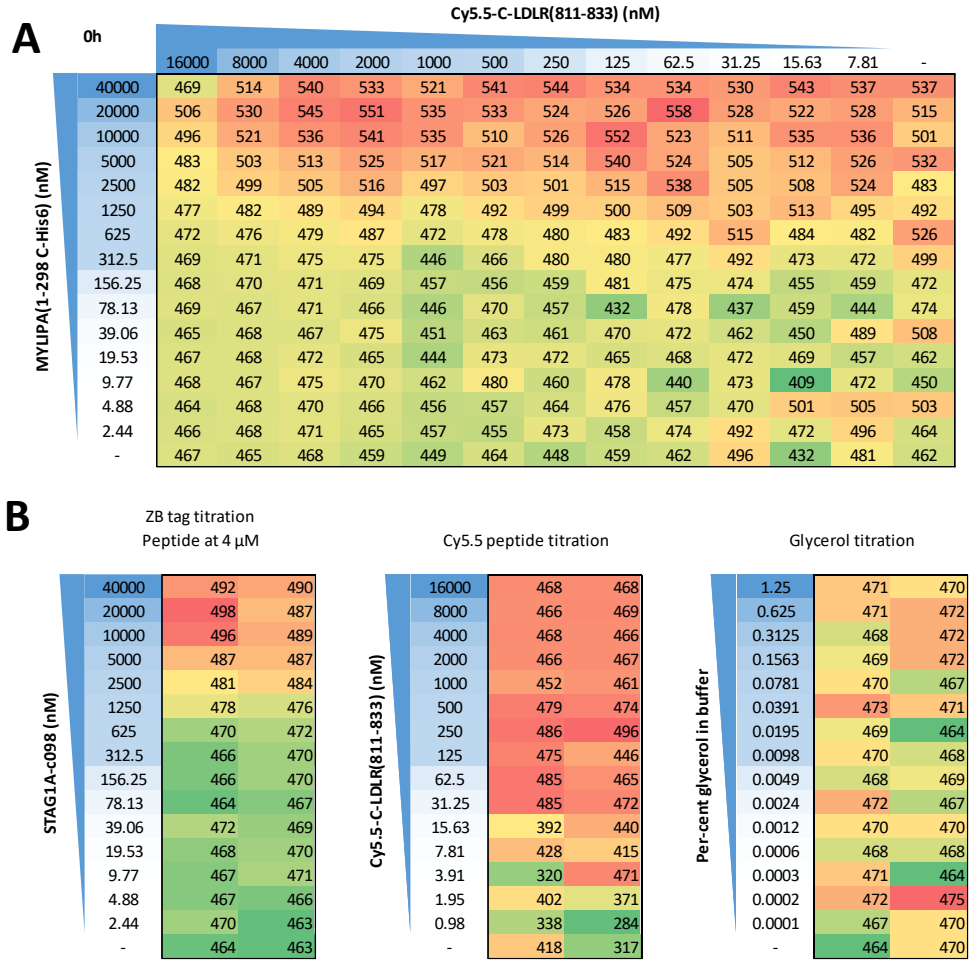


1h

STAG1A (nM) \ LDLR(811-859)-K-biotin (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	0.4883	0.2441	-
4000	573	600	717	619	604	617	616	692	485	342	313	216	192	169	166	154
2000	760	569	783	854	904	777	800	866	579	446	321	227	200	185	157	169
1000	746	763	835	783	800	634	677	845	619	369	316	205	174	152	150	146
500	506	592	726	601	723	536	540	601	525	417	316	209	156	148	143	144
250	430	435	573	535	606	453	501	484	413	290	229	180	143	138	142	136
125	254	263	457	365	428	335	414	362	205	242	177	150	141	137	135	134
62.5	217	220	272	259	269	240	278	246	175	168	145	138	137	137	137	135
31.25	164	173	200	188	198	187	189	170	150	147	141	131	132	135	138	133
15.625	151	159	162	150	162	151	158	159	142	139	140	132	136	140	133	134
7.8125	140	143	148	145	145	143	146	137	139	138	134	131	132	134	132	130
3.90625	136	136	136	135	140	139	135	142	138	138	135	128	131	132	130	133
1.953125	137	137	136	140	135	137	132	139	133	133	137	131	135	136	130	131
0.976563	137	135	137	136	144	142	139	136	131	135	130	129	132	132	134	129
0.488281	135	138	135	132	132	135	135	130	135	136	132	131	134	134	139	133
0.244141	138	139	142	134	136	135	134	136	134	135	132	135	136	131	130	130
-	138	137	133	134	134	135	132	133	135	134	133	134	133	133	132	133

Figure 38: HTRF control experiment with unrelated, ZB tagged, protein STAG1A. STAG1A titrated against LDLR$_{(811-859)}$-K-biotin.

3.3.1.2 Fluorescence Polarisation assay



A

0h

MYLIPA(1-298 C-His6) (nM) \ Cy5.5-C-LDLR(811-833) (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	-
40000	469	514	540	533	521	541	544	534	534	530	543	537	537
20000	506	530	545	551	535	533	524	526	558	528	522	528	515
10000	496	521	536	541	535	510	526	552	523	511	535	536	501
5000	483	503	513	525	517	521	514	540	524	505	512	526	532
2500	482	499	505	516	497	503	501	515	538	505	508	524	483
1250	477	482	489	494	478	492	499	500	509	503	513	495	492
625	472	476	479	487	472	478	480	483	492	515	484	482	526
312.5	469	471	475	475	446	466	480	480	477	492	473	472	499
156.25	468	470	471	469	457	456	459	481	475	474	455	459	472
78.13	469	467	471	466	446	470	457	432	478	437	459	444	474
39.06	465	468	467	475	451	463	461	470	472	462	450	489	508
19.53	467	468	472	465	444	473	472	465	468	472	469	457	462
9.77	468	467	475	470	462	480	460	478	440	473	409	472	450
4.88	464	468	470	466	456	457	464	476	457	470	501	505	503
2.44	466	468	471	465	457	455	473	458	474	492	472	496	464
-	467	465	468	459	449	464	448	459	462	496	432	481	462

B

ZB tag titration
Peptide at 4 µM

STAG1A-c098 (nM)		
40000	492	490
20000	498	487
10000	496	489
5000	487	487
2500	481	484
1250	478	476
625	470	472
312.5	466	470
156.25	466	470
78.13	464	467
39.06	472	469
19.53	468	470
9.77	467	471
4.88	467	466
2.44	470	463
-	464	463

Cy5.5 peptide titration

Cy5.5-C-LDLR(811-833) (nM)		
16000	468	468
8000	466	469
4000	468	466
2000	466	467
1000	452	461
500	479	474
250	486	496
125	475	446
62.5	485	465
31.25	485	472
15.63	392	440
7.81	428	415
3.91	320	471
1.95	402	371
0.98	338	284
-	418	317

Glycerol titration

Per-cent glycerol in buffer		
1.25	471	470
0.625	471	472
0.3125	468	472
0.1563	469	472
0.0781	470	467
0.0391	473	471
0.0195	469	464
0.0098	470	468
0.0049	468	469
0.0024	472	467
0.0012	470	470
0.0006	468	468
0.0003	471	464
0.0002	472	475
0.0001	467	470
-	464	470

Figure 39: Fluorescence polarisation assay development. **A:** Protein-peptide titration of MYLIPA$_{(1-298\ C\text{-}His6)}$ against Cy5.5-LDLR$_{(811-833)}$ **B:** Control experiments with ZB tagged protein, glycerol content and total fluorescence (Cy5.5 peptide only titration). Numbers shown are milli-polarisation values.

3.3.1.3 AlphaScreen – MYLIPA(1-276 C-His6) and MYLIPA(1-298 C-His6) titrations against receptor tail peptides.

A

LDLR(811-859)-K-biotin (nM)

MYLIPA(1-276 C-His6) (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
16000	443.4	303.9	446.2	406.1	422.3	406.3	417.4	367.5	346.2	309.3	234.1	117.4	60.8	26.6	12.2	5.9
8000	582.9	602.5	603.6	597.3	571.1	557.5	525.7	494.3	435.6	368.9	226.4	111.2	56.5	21	8.7	3.1
4000	592.4	611.4	604.8	565.1	539.7	511.3	474.1	444.7	409.6	318.4	198	103	53.3	23.8	9.7	3.9
2000	502.6	474	486.1	473.5	446.9	431.2	390.8	348	306.3	262.5	160.3	91.9	52.2	22.5	10.5	3.7
1000	226.8	278	315.6	328.2	321.8	311.8	279.7	255.3	216.1	205.1	125	73.3	41.7	18.9	12.2	3.7
500	95.8	113.3	130.1	137.7	134.2	214.5	240.1	213.7	311.7	249.5	156.4	103.8	59.7	33.8	19.7	4.6
250	313.1	400.8	523.8	625.1	622.1	627.5	641.8	591.1	574.7	461.2	289.1	197.4	131.3	74.4	43.8	7.4
125	339.7	460.4	576.8	748.1	806.1	836.8	861.6	874.7	810.5	681.1	441.3	285.3	196	107	61	9.3
62.5	356.5	478.5	673.1	780	759.1	928.4	983.1	978	917.2	818.5	559.2	384.7	266.8	158.9	86.4	7.3
31.25	316.8	377.9	504.4	591.1	692.2	815.6	832.6	852.6	840.3	761.4	518.4	333.6	262.4	159.4	95.1	7.1
15.625	7	925.4	444.8	512.4	620.5	687.1	703.4	691.7	662.3	605.5	390.7	275.7	189.7	116.5	64.9	6.4
7.8125	239	288.5	351	441.8	518.5	567.6	559.8	542.8	545.1	491	322.8	209.5	145.1	75.4	42.6	4.9
3.90625	223	259.9	335	402	456.7	444.5	435	414.6	370.2	336.8	218.5	140.8	85.3	49.1	25.2	3.6
1.953125	203.3	244.7	315.6	375	389	369.7	328.3	275.5	249.5	219.8	138.5	89.9	48.3	25.5	13.6	4.5
0.976563	212.5	233.5	296	363.7	365.3	310.8	224.3	181.5	135.3	112.6	72.8	38.3	27.6	12.1	6.4	2.9
-	183.2	210.8	278.6	359.5	335.4	212.4	86.6	30.1	9.3	5.3	3.2	3	4.1	3.1	3.2	3.1

B

biotin-VLDLR(820-842) (nM)

MYLIPA(1-276 C-His6) (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
16000	3.2	3.9	3.5	3.8	3.4	3.1	3.3	2.6	3.6	3.8	3.5	6	7.7	7	5.7	4.6
8000	3.3	2.9	2.8	3.5	2.9	2.8	2.7	2.8	2.9	2.9	3.1	3.6	3.5	2.9	2.6	3.7
4000	3.4	3.1	3.3	3.2	3.8	2.9	3.1	3.9	3.3	3.1	4	2.7	2.9	3.8	3.2	3.4
2000	3.5	3.2	3.4	3.6	3.2	3.1	2.6	2.3	4.1	2.4	3.1	3.6	3.7	3.9	2.6	2.6
1000	2.8	2.5	2.4	3	2.6	3	2.8	2.8	2.3	2.2	2.5	2.9	2.3	2.7	2.7	2.9
500	4.7	3.7	4	6.1	3.6	3.9	2.7	3.6	3	3.8	3.1	3.4	2.8	2.9	2.3	3.4
250	29.4	35.1	38.8	36.7	42	42.4	32.2	35.5	30.3	26.1	20.6	7.2	5	4.7	5.1	4.1
125	46.7	74.1	96.6	114.4	128.1	134.7	120.5	135.3	116.8	102.4	68.2	28.6	9.9	4.8	5.3	5.3
62.5	62.1	93.1	139.1	158.1	180.5	194	193.4	195.5	176.6	158.7	118	37.2	7	4.4	4.3	4.8
31.25	62.3	92.5	126.9	148.5	169.7	194.3	195.6	192.7	191.4	177.3	138.2	39.3	9.3	3.1	4.5	3.8
15.625	66.3	93.6	126.3	146.1	167	179.7	188.2	188.3	183.2	174.6	132.7	46	12.1	4.7	3.6	3.4
7.8125	53	61.1	77.5	98.1	113.2	124.3	132.2	129.4	138.7	135.7	112.6	323.2	11.4	3.6	3.4	4.9
3.90625	32.9	37.8	47.8	53.8	67.6	70.9	77.7	74.7	76.9	75.6	68.9	31	10.2	4.2	2.4	4.5
1.953125	21.2	17.5	19.7	24.4	27.8	29	30.6	32.2	31.8	30.7	26.5	19.8	8.9	3.9	3.3	3.4
0.976563	8.7	8.9	7.6	9.3	9.3	10.3	10.8	12.1	10.2	9.9	7.9	6.8	4.4	3.9	3.7	3.4
-	4.4	3.2	3.2	3.9	3.2	3	4.2	3.6	2.6	3.5	3.5	3	3.3	2.7	4.4	3.6

C

biotin-APP(745-767) (nM)

MYLIPA(1-276 C-His6) (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
16000	75.3	69.7	68.5	74	66.9	68.2	72.8	67	63.6	59.7	54.6	30.5	10.1	8.9	6.5	4.8
8000	72.6	74.9	73.9	72	74.4	71.1	68.3	67.2	70	62.3	59.8	44.2	10.3	3.4	3.1	4.3
4000	51.9	54.6	50.9	53.1	45.8	53.4	49.3	48.1	57.6	65.9	50.6	22.8	4	4	3.6	2.8
2000	15.9	17.9	17.8	19.6	18	19	23.3	25.9	22.4	18.6	20	6.2	3.2	3.5	4.2	3.6
1000	5.2	4.5	5.1	4.1	4.9	5.8	4.9	4.7	5.4	4.6	3.7	4.4	3	2.8	3.2	2.5
500	5.4	6	4.6	4.6	5.5	3.8	5.5	5.6	3.8	5	5.3	3.1	3.2	3.1	3.1	3
250	5	6.3	5.6	6.3	5	5.2	6.1	6.1	6.8	7.4	4.8	6.6	5.1	4.3	4.6	5.8
125	5.5	5.2	5.8	6.4	6	5.5	6.4	6.1	6	4.9	6.3	4.6	5.6	5.5	4.4	6.1
62.5	4.8	5.5	5.9	5.1	4	4.6	5.7	4.7	5	4.7	3.9	4.4	4.4	4.5	4.5	5.1
31.25	4	3.5	4.9	3.5	4	3.8	4	3.8	3.2	3.3	3.6	4.4	4.1	5.1	3.8	4.7
15.625	4	4.2	4.3	4.1	4.1	4.3	3.2	4.3	3.9	3.3	4.1	3.6	4.1	3.3	3.9	4.6
7.8125	3.1	3.9	3	3.1	3.9	3.3	3.7	2.9	3.1	4	4	3	3.4	3.5	4.5	3.2
3.90625	3.2	3.2	3.4	3.5	3	3.5	2.9	3.4	3	3.2	2.3	3.2	3.9	3	2.9	2.7
1.953125	2.9	2.9	3.1	3.1	3.2	2.9	2.3	2.9	2.7	2.6	3.3	2.5	2.5	3.6	2.8	3.4
0.976563	2.7	3.1	2.6	3.2	3.4	2.6	2.6	3.1	2.8	2.7	2.8	3.3	3.1	3.7	2.7	2.3
-	4.4	4.4	2.4	3.1	2.9	2.8	3	2.6	2.6	3.2	3.8	3	2.6	3.5	3.5	2.8

D

biotin-ApoER2(849-871) (nM)

MYLIPA(1-276 C-His6) (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
16000	3.4	3.3	3	3.2	3.1	3.2	3.4	3	3.3	5.5	10.2	11.9	9.9	5.7	5.7	5.9
8000	2.9	2.8	3	3.5	3.1	3.7	2.6	3.5	3.7	3.9	14.6	10.3	4.1	3.6	3.8	2.6
4000	3.6	3	3.2	4.3	3.3	3.1	3.5	3.5	2.8	4.4	15.4	17.1	4.4	3.2	2.6	3.1
2000	4.2	3.5	2.1	2.8	3.7	2.7	3	3.4	3.2	4.3	23.7	21.2	5.5	2.8	3	2.9
1000	3.3	2.7	3.3	2.9	3.1	2.9	3.2	3.2	2.9	3.1	12	15.5	6.9	3.3	2.5	1.9
500	3	3.2	3.3	4	4.3	4.8	4.5	5.5	5.8	5.9	14.6	51.3	22.3	7.8	5.8	3.1
250	3.7	6.9	11.9	18.1	26.9	35.6	50.6	72	100.1	139.9	236.3	264.1	143.3	80.4	39.7	4.2
125	7.6	13.5	17.9	27.1	45.7	71.7	103.6	156.4	224.4	316.6	423.4	457.9	285	166	84.3	5.1
62.5	11.6	18.3	31.8	50.3	69	106.9	143.8	188.7	278.1	389.2	588.8	579.7	373.2	197.9	88.9	3.6
31.25	13.4	24.6	50.2	70.8	99.3	115.4	141.3	176.3	241.7	382.9	584	588.5	358.3	155.8	70	4.4
15.625	14	30.2	58.6	92.6	117.3	112.9	110.1	146.9	183.6	300.1	435	503	282.9	133.2	55.4	3.5
7.8125	14.4	28.4	61.7	110.3	143.8	137.6	124	107.9	124.9	196	452.5	334.1	208.8	81.7	34.4	4.5
3.90625	10.1	22	62.5	121	146.8	112.8	86	71	69.7	93.3	175.2	187.4	83.8	37.6	21.7	3.7
1.953125	7.8	14.6	49.8	108.2	120.2	85.2	57	34.3	32.8	35.8	38.3	55.4	32.8	16.4	8.9	4
0.976563	5.8	9.9	31.9	78.7	73.7	50.9	24.9	16.5	11.5	10.2	15.4	12.9	9.7	6.5	5	2.2
-	3.5	3.5	3.2	2.1	2.9	1.7	3.4	3.1	3.2	3.5	3.9	3.4	3.7	3.6	3.6	3

Figure 40: AlphaScreen titration of MYLIPA(1-276 C-His6) against receptor tail peptides. **A:** LDLR(811-859)-K-biotin. **B:** biotin-VLDLR(820-842). **C:** biotin-APP(745-767). **D:** biotin-ApoER2(849-871).Signal is in thousands.

3.3.1.4 AlphaScreen – MYLIPA(1-276 C-His6) and MYLIPA(1-298 C-His6) x LDLR(811-859)-K-biotin vs LDLR(811-834)(not biotinylated)

LDLR(811-834) (nM)

	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
MYLIPA(1-276 C-His6) (30 nM) LDLR(811-859)-K-biotin (30 nM)	545.8	479.8	430.0	403.0	425.1	427.1	438.7	464.2	445.3	444.3	446.9	466.0	443.5	450.7	449.2	498.9
	539.2	526.3	451.5	363.3	402.2	426.8	417.0	447.5	438.6	425.5	429.0	444.4	413.3	421.0	367.3	477.8
	519.0	460.3	396.9	354.0	375.6	382.4	375.7	382.5	398.2	395.4	410.6	421.4	403.3	399.1	374.8	420.3

LDLR(811-834) (nM)

	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	0.4883	0.2441	-
MYLIPA(1-276 C-His6) (30 nM) LDLR(811-859)-K-biotin (30 nM)	353.6	349.7	331.7	340.1	345.6	372.6	375.7	388.4	373.4	394.7	404.8	401.2	416.0	401.4	378.7	451.9
	369.4	419.5	356.4	323.4	329.3	350.5	320.8	366.6	371.0	378.3	377.5	382.3	377.5	375.8	336.9	278.0
	385.2	315.1	367.7	301.0	359.1	362.9	306.0	342.4	356.0	382.7	363.3	359.4	375.6	377.5	327.7	314.9

LDLR(811-834) (nM)

	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
MYLIPA(1-298 C-His6) (10 nM) LDLR(811-859)-K-biotin (30 nM)	444.9	390.2	304.2	231.3	267.1	264.0	258.9	280.0	277.4	282.1	286.4	283.7	276.4	288.4	275.8	248.3
	449.9	413.7	302.3	263.5	266.0	235.8	267.2	236.3	234.3	251.6	299.0	292.7	271.7	294.5	287.1	259.7
	457.6	393.9	317.7	277.2	273.7	277.8	261.6	285.2	273.6	285.9	310.9	306.0	294.8	313.1	297.8	292.0

LDLR(811-834) (nM)

	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	0.4883	0.2441	-
MYLIPA(1-298 C-His6) (10 nM) LDLR(811-859)-K-biotin (30 nM)	363.3	318.2	296.1	277.8	281.4	277.0	275.5	289.5	284.4	294.7	301.1	307.8	293.1	308.9	301.9	301.5
	359.2	320.3	301.4	303.5	290.9	277.4	275.4	295.1	292.3	298.8	309.2	304.1	310.3	300.0	307.9	326.3
	367.7	347.2	357.0	306.3	316.4	308.9	317.6	312.0	299.9	302.2	324.9	320.2	326.4	331.7	323.4	333.8

Figure 41: AlphaScreen control experiment. Titration of non-biotinylated LDLR(811-834) against set concentrations (as indicated) of MYLIPA(1-276 C-His6) or MYLIPA(1-298 C-His6) and LDLR(811-859)-K-biotin. Signal is in thousands.

3.3.1.5 HTRF – MYLIPA(1-298, T279D, S282D, C-His6) vs biotin-VLDLR(820-842)

biotin-VLDLR(820-842) (nM)

2h

MYLIPA(1-298, T279D, S282D, C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000		304	261	287	275	289	259	272	272	264	265	263	261	278	276	254
2000	270	278	268	297	309	280	269	262	277	269	262	260	270	270	274	274
1000	273	298	282	286	276	298	283	269	282	275	278	289	267	283	270	273
500	308	304	317	310	297	308	289	271	280	275	261	271	279	294	262	266
250	352	362	347	384	350	365	312	282	300	277	268	308	281	290	288	291
125	545	511	490	490	435	448	372	324	328	303	309	304	289	316	287	304
62.5	638	595	602	639	583	565	456	397	375	336	321	327	312	339	313	318
31.25	688	733	717	767	719	726	528	411	372	347	359	367	329	382	333	325
15.63	1685	838	788	815	770	761	761	429	403	359	358	363	351	373	368	350
7.81	701	715	715	719	700	703	521	399	377	352	354	346	361	364	352	332
3.91	615	644	565	599	536	527	421	353	340	327	303	345	316	342	323	330
1.95	461	459	422	450	395	386	366	315	331	308	304	317	299	323	300	293
0.98	383	383	346	356	357	326	326	292	308	279	280	305	282	294	291	302
0.49	341	322	317	347	315	302	297	284	307	286	293	293	278	303	289	271
0.24	308	298	290	315	306	307	309	284	287	285	292	286	290	288	292	267
-	300	290	280	303	272	295	295	281	297	285	287	284	280	294	274	268

Figure 42: HTRF titration of phospho-mimetic MYLIPA(1-298, T279D, S282D, C-His6) against biotinylated VLDLR(820-842) peptide. Wells where air bubbles led to a seeming increase in FRET ratio are greyed out.

3.3.2 Cooperativity of MYLIP and LDLR adapter proteins

3.3.2.1 Protein-peptide titrations of DAB1 and DAB2 PTBs against receptor tail peptides in AlphaScreen and HTRF

These experiments were performed by Charline Giroud during my parental leave.

biotin-VLDLR(820-842) (nM)

DAB1A-p001 (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	105.4	114.9	103.3	71.3	183.4	68.6	44.2	27.5	15.4	4.7	1.7	1.0
75	200.7	226.6	207.7	212.1	214.6	135.7	84.6	43.8	17.2	4.3	1.3	0.9
37.5	305.8	361.2	344.1	375.5	361.4	239.6	171.0	85.5	34.3	4.7	1.0	1.0
18.75	327.2	404.6	422.0	457.6	442.2	319.6	216.6	110.2	50.1	7.8	1.6	1.0
9.38	285.7	367.1	385.4	445.5	452.3	321.6	197.6	93.9	36.2	4.3	1.7	1.1
4.69	167.6	259.7	323.8	360.0	360.3	277.6	164.2	76.7	21.4	4.5	1.2	1.1
2.34	103.5	147.2	200.2	221.1	238.8	173.9	90.0	33.7	8.5	1.7	0.6	0.6
-	1.0	0.4	0.4	0.5	0.6	0.6	0.9	0.8	0.6	0.7	0.8	0.5

DAB1A-p003 (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	317.7	354.2	349.8	356.3	349.0	390.0	333.5	231.9	126.5	39.3	5.0	0.3
75	383.5	413.8	409.0	425.5	430.7	443.5	371.1	240.0	125.0	39.6	6.7	0.7
37.5	405.9	445.4	429.8	475.0	462.9	474.6	376.3	241.3	142.0	45.5	6.1	0.7
18.75	478.0	540.2	554.8	536.5	581.8	503.0	387.0	245.8	125.5	39.0	6.9	0.7
9.38	521.8	622.7	624.0	703.0	673.0	561.6	412.8	235.0	129.2	43.8	5.1	0.8
4.69	473.6	160.3	685.0	696.3	752.1	604.2	395.3	228.7	103.5	27.0	5.1	0.9
2.34	485.6	637.0	706.3	771.7	770.5	577.8	341.4	192.9	83.9	23.3	4.6	1.0
-	1.2	0.5	0.8	0.7	0.7	0.7	0.7	0.3	0.8	0.8	0.5	0.6

DAB2A-p001 (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	219.0	255.9	238.6	257.1	240.2	191.4	128.0	64.9	19.7	1.9	0.5	0.8
75	294.3	346.2	333.6	372.9	336.3	251.9	151.8	66.6	18.0	2.2	0.5	0.8
37.5	285.0	355.9	289.8	395.8	380.3	288.7	174.7	77.6	21.1	2.0	1.1	0.6
18.75	249.3	332.8	470.1	396.0	403.6	297.8	159.2	74.6	19.9	2.5	0.8	0.9
9.38	191.5	257.5	310.4	333.8	330.8	247.4	126.2	53.1	10.8	1.4	0.6	0.5
4.69	101.7	138.9	187.4	205.9	216.6	161.3	74.1	32.7	6.6	1.1	0.7	0.6
2.34	50.5	73.8	97.7	101.8	124.8	80.4	28.8	10.3	2.0	0.4	0.8	1.0
-	0.6	0.6	0.8	0.4	1.1	0.7	0.5	0.6	0.4	0.7	0.3	1.6

DAB2A-p003 (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	385.2	445.2	438.9	470.3	456.3	455.2	379.1	224.0	99.9	20.3	1.9	0.7
75	493.1	582.0	554.2	599.1	561.8	515.0	388.8	215.9	97.6	19.3	2.9	0.7
37.5	466.4	588.6	575.8	637.7	606.9	563.6	394.4	217.4	99.0	24.7	2.7	0.4
18.75	549.6	682.3	732.7	764.5	736.9	607.6	361.5	203.8	81.2	15.6	2.6	0.7
9.38	544.0	703.8	723.3	801.9	776.6	643.6	388.0	185.5	75.4	23.5	1.7	0.8
4.69	454.7	607.8	686.0	736.0	768.8	609.2	331.3	163.8	51.3	4.8	1.6	0.8
2.34	383.2	525.1	598.7	631.4	662.7	515.3	258.9	120.0	57.8	8.6	1.4	0.9
-	0.7	1.5	0.8	0.8	3.4	0.6	0.7	1.0	0.8	0.8	1.0	1.3

Figure 43: AlphaScreen titration of biotin-VLDLR$_{(820-842)}$ against DAB1 and DAB2 PTBs (p001: N-terminally tagged, p003: C-terminally tagged in both cases). AlphaScreen signals are shown in thousands.

3.3.2.2 HTRF – MYLIPA$_{(1\text{-}276\ C\text{-}His6)}$ vs receptor peptides in presence of DAB1A-p004 (His$_6$ tag cleaved, C-terminally tagged)

These experiments were performed by Charline Giroud during my parental leave.

A

DAB1A-p004 (20 nM) / MYLIPA(1-276 C-His6) (nM)	biotin-VLDLR(820-842) (nM)															
	3200	1600	800	400	200	100	50	25	12.50	6.25	3.13	1.56	0.78	0.39	0.20	-
3200	100	97	94	96	95	94	96	96	96	94	96	99	96	97	96	98
1600	105	98	98	98	97	99	100	96	100	98	98	99	97	99	97	97
800	99	98	96	95	95	96	99	97	97	97	98	98	96	97	97	101
400	97	100	96	97	98	99	97	98	100	97	98	101	98	97	97	99
200	100	100	98	103	99	104	102	101	101	101	101	101	100	101	101	100
100	100	102	100	101	100	99	101	100	105	99	98	101	97	100	99	98
50	98	101	100	103	101	102	104	115	100	99	98	99	98	99	99	99
25	99	102	97	101	100	103	104	102	103	101	103	100	99	100	99	101
12.50	99	100	101	102	104	101	107	97	100	99	120	97	97	99	97	97
6.25	101	104	104	107	106	106	107	105	106	102	138	104	101	101	99	101
3.13	105	101	101	103	103	104	103	104	103	102	138	102	102	100	98	99
1.56	117	104	106	109	109	114	111	108	108	107	132	104	101	98	98	97
0.78	100	102	100	100	99	103	103	99	101	99	121	101	99	99	99	100
0.39	99	99	97	99	100	145	100	97	99	99	114	99	98	99	97	97
0.20	96	100	99	102	101	169	102	99	101	99	109	101	100	99	100	101
-	114	102	100	101	101	138	101	100	102	98	99	99	100	99	102	105

B

DAB1A-p004 (20 nM) / MYLIPA(1-276 C-His6) (nM)	biotin-ApoER2(849-871) (nM)															
	3200	1600	800	400	200	100	50	25	12.50	6.25	3.13	1.56	0.78	0.39	0.20	-
3200	103	104	98	99	99	98	100	96	102	98	101	101	98	101	100	99
1600	103	100	97	100	99	99	101	98	102	100	100	104	98	98	101	103
800	102	103	98	101	98	100	103	98	100	101	100	102	101	102	102	104
400	97	100	99	102	100	100	103	100	102	99	100	99	100	98	99	98
200	102	104	102	102	103	105	107	103	103	103	103	103	102	102	101	103
100	100	102	102	104	104	104	108	107	108	101	101	102	101	100	98	102
50	103	104	102	105	102	107	111	108	107	103	102	103	100	99	101	99
25	102	103	100	105	104	105	109	108	106	104	101	103	98	102	101	99
12.50	103	104	100	102	101	105	107	106	108	102	102	101	98	97	99	97
6.25	105	104	103	106	106	107	113	108	108	106	104	103	99	100	99	100
3.13	100	103	103	103	100	102	105	104	102	102	101	103	100	100	99	99
1.56	98	98	98	99	98	101	102	100	102	100	99	101	98	96	97	95
0.78	99	100	98	101	100	102	104	102	103	98	100	101	98	101	98	99
0.39	99	100	96	100	97	99	103	100	100	96	97	97	98	98	96	98
0.20	101	100	100	100	99	103	105	102	99	100	101	103	98	99	98	99
-	102	101	98	102	100	103	103	101	101	101	99	102	100	99	100	99

C

DAB1A-p004 (20 nM) / MYLIPA(1-276 C-His6) (nM)	biotin-APP(745-767)-biotin (nM)															
	3200	1600	800	400	200	100	50	25	12.50	6.25	3.13	1.56	0.78	0.39	0.20	-
3200	103	102	100	99	98	95	100	95	99	95	100	100	102	99	100	97
1600	101	102	98	97	97	97	99	98	101	97	100	103	99	96	102	102
800	99	101	100	99	98	99	101	97	99	100	103	100	98	100	98	102
400	96	98	96	97	97	98	100	99	99	98	101	100	98	106	95	97
200	99	99	99	101	99	100	102	100	102	102	101	102	100	100	99	100
100	115	90	96	100	100	101	103	102	98	100	102	104	99	101	100	98
50	99	101	99	102	102	101	105	99	99	99	100	99	97	98	99	98
25	101	103	98	101	102	102	105	104	107	101	101	101	100	101	100	101
12.50	99	101	101	104	108	111	103	104	102	99	97	99	97	97	97	99
6.25	100	103	105	107	107	106	108	104	104	105	102	105	100	103	97	100
3.13	98	102	100	104	102	102	104	104	104	103	102	105	98	99	99	99
1.56	97	99	100	101	99	102	103	101	100	102	101	100	95	97	94	97
0.78	98	100	100	101	100	100	102	102	101	101	110	102	97	101	99	100
0.39	92	98	97	100	99	98	101	97	99	98	99	102	96	99	96	97
0.20	113	101	101	102	102	95	105	101	102	100	103	101	101	102	101	100
-	99	100	99	101	99	112	101	99	101	99	97	101	101	101	98	99

Figure 44: HTRF protein-peptide titration of MYLIPA$_{(1\text{-}276\ C\text{-}His6)}$ in presence of 20 nM DAB1A-p004 against receptor tail peptides. **A:** biotin-VLDLR$_{(820\text{-}842)}$. **B:** biotin-ApoER2$_{(849\text{-}871)}$. **C:** biotin-APP$_{(745\text{-}767)}$.

3.3.2.3 HTRF – Phospho-mimetic MYLIP against receptor peptides in presence of PTBs

A

O/N DAB1A-p002 | LDLR(811-834)-K-biotin (nM)

MYLIPA(1-298 T279D S282D C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	393	1008	665	584	623	596	563	349	227	213	196	204	193	191	206	193
2000	836	924	927	905	950	784	820	363	296	232	201	219	200	201	202	196
1000	1226	1277	1137	1253	1275	762	1188	634	346	283	232	236	208	216	211	201
500	1332	1252	1345	1402	1322	1278	1094	778	407	244	268	241	231	217	215	202
250	769	999	1201	1142	1049	1126	962	704	425	337	288	262	233	221	223	209
125	782	813	812	777	697	726	604	498	355	292	270	252	238	226	226	208
62.5	709	703	681	617	597	579	492	371	280	245	251	260	240	230	230	215
31.25	654	649	618	582	566	502	453	360	280	272	245	275	268	254	272	216
15.63	657	664	644	620	603	489	461	346	363	279	257	288	284	283	286	224
7.81	629	628	634	609	569	512	460	333	282	282	271	300	307	301	312	223
3.91	645	634	647	626	575	710	470	363	296	297	283	297	273	301	322	233
1.95	648	622	624	616	584	560	467	379	302	286	289	301	307	303	319	231
0.98	628	593	592	609	573	717	482	350	290	321	286	301	300	298	301	238
0.49	598	586	581	579	535	544	521	326	284	279	276	293	300	301	299	236
0.24	587	550	575	534	497	493	412	302	282	271	285	290	310	312	297	244
-	501	457	454	400	399	384	334	264	227	231	221	230	237	245	267	229

B

O/N DAB1A-p002 | biotin-VLDLR(820-842) (nM)

MYLIPA(1-298 T279D S282D C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	221	221	236	245	231	237	222	206	208	196	194	202	193	195	195	199
2000	273	271	295	294	276	304	261	220	216	212	202	206	210	196	200	198
1000	382	387	414	428	417	462	410	270	240	229	218	215	207	211	205	200
500	446	476	528	508	522	538	456	309	249	227	218	220	206	208	202	201
250	464	506	520	479	535	545	440	336	267	248	251	212	212	209	220	211
125	535	438	412	400	437	432	401	316	288	263	279	223	212	210	208	212
62.5	376	382	384	383	399	392	337	273	280	281	256	226	223	220	214	210
31.25	448	441	431	442	447	464	401	350	322	328	293	242	225	216	209	204
15.63	565	629	738	676	641	649	534	423	403	434	392	283	244	219	227	213
7.81	521	517	508	528	520	532	418	332	373	435	398	279	246	240	213	220
3.91	554	515	504	545	533	537	504	352	414	593	413	310	267	239	226	210
1.95	565	536	566	544	580	544	458	361	442	598	424	319	273	246	242	230
0.98	565	551	507	520	588	545	538	372	402	524	400	315	284	249	234	229
0.49	491	523	481	474	493	481	450	335	394	436	375	302	267	246	231	217
0.24	464	453	461	451	423	466	362	306	364	384	345	294	274	249	224	231
-	252	248	269	280	280	290	280	245	228	234	242	255	244	240	241	218

C

O/N DAB1A-p002 | biotin-ApoER2(894-871) (nM)

MYLIPA(1-298 T279D S282D C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	229	225	296	219	213	214	198	209	203	194	203	203	189	213	196	195
2000	291	297	299	252	264	275	237	217	213	203	215	220	202	202	205	202
1000	385	382	378	260	270	217	241	219	220	217	213	219	216	214	222	213
500	248	274	295	215	226	266	237	224	218	222	212	213	214	209	215	213
250	276	264	280	210	202	229	235	232	219	216	205	208	209	204	215	213
125	222	261	223	226	220	217	225	217	218	217	216	207	213	206	212	208
62.5	226	227	214	212	262	216	218	223	221	218	215	214	207	216	215	215
31.25	224	226	231	235	231	248	233	235	239	238	229	219	220	214	215	223
15.63	281	283	265	263	250	306	260	283	305	288	265	243	216	219	222	217
7.81	253	260	258	262	251	266	262	276	294	300	258	218	227	232	217	226
3.91	264	268	269	263	261	279	275	294	315	313	265	231	230	223	223	221
1.95	302	262	264	268	255	278	285	294	317	309	273	223	222	220	218	229
0.98	259	266	267	265	272	274	264	288	314	303	260	229	213	212	223	227
0.49	268	251	240	264	259	272	268	284	396	281	253	224	215	220	221	224
0.24	259	259	255	246	256	256	270	268	289	276	256	237	222	222	220	241
-	222	221	217	235	212	224	215	232	234	252	241	234	232	220	221	227

D

O/N DAB1A-p002 | biotin-APP(745-767) (nM)

MYLIPA(1-298 T279D S282D C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	344	362	391	430	428	477	552	481	243	213	203	200	198	199	195	195
2000	486	605	663	761	762	850	912	936	362	242	222	211	200	199	207	205
1000	785	960	1112	1172	1341	1480	1786	1120	508	277	249	216	213	204	205	209
500	969	1233	1370	1377	1467	1760	2047	1555	685	333	269	240	215	212	209	205
250	843	1131	1359	1318	1417	1621	1936	1546	679	363	286	257	226	225	219	217
125	330	483	1081	951	1018	1179	1628	987	482	318	254	252	223	220	206	215
62.5	598	662	865	827	917	995	1009	1075	538	326	247	252	228	221	218	218
31.25	817	1101	1490	1087	1115	1206	1363	940	571	379	355	316	294	257	237	221
15.63	680	806	1001	1036	1078	1167	1444	1029	590	413	307	303	252	244	228	220
7.81	696	767	1011	1008	1088	1135	1238	1273	659	439	334	339	279	266	239	224
3.91	870	1179	1054	1060	1105	1201	1364	1032	607	459	337	339	286	267	267	215
1.95	1218	1405	1356	1301	1359	1370	1378	1265	1004	774	588	547	515	391	400	241
0.98	802	1020	1099	1039	1077	1157	1200	1057	648	497	397	391	328	307	289	228
0.49	745	882	1034	923	914	982	1066	884	527	409	332	325	305	282	278	225
0.24	578	798	851	775	909	770	991	1073	537	388	313	318	309	295	264	230
-	261	295	338	335	427	253	632	520	276	243	229	250	252	247	259	227

Figure 45: HTRF titration of phospho-mimetic MYLIPA$_{(1\text{-}298\ T279D\ S282D\ C\text{-}His6)}$ against receptor tail peptides in presence of 20 nM DAB1A-p002. **A:** LDLR$_{(811\text{-}834)}$-K-biotin. **B:** biotin-VLDLR$_{(820\text{-}842)}$. **C:** biotin-ApoER2$_{(849\text{-}871)}$. **D:** biotin-APP$_{(745\text{-}767)}$.

3.3.2.4 HTRF – Effect of DAB PTBs on assay and displacement with non-tagged phosphomimetic MYLIP

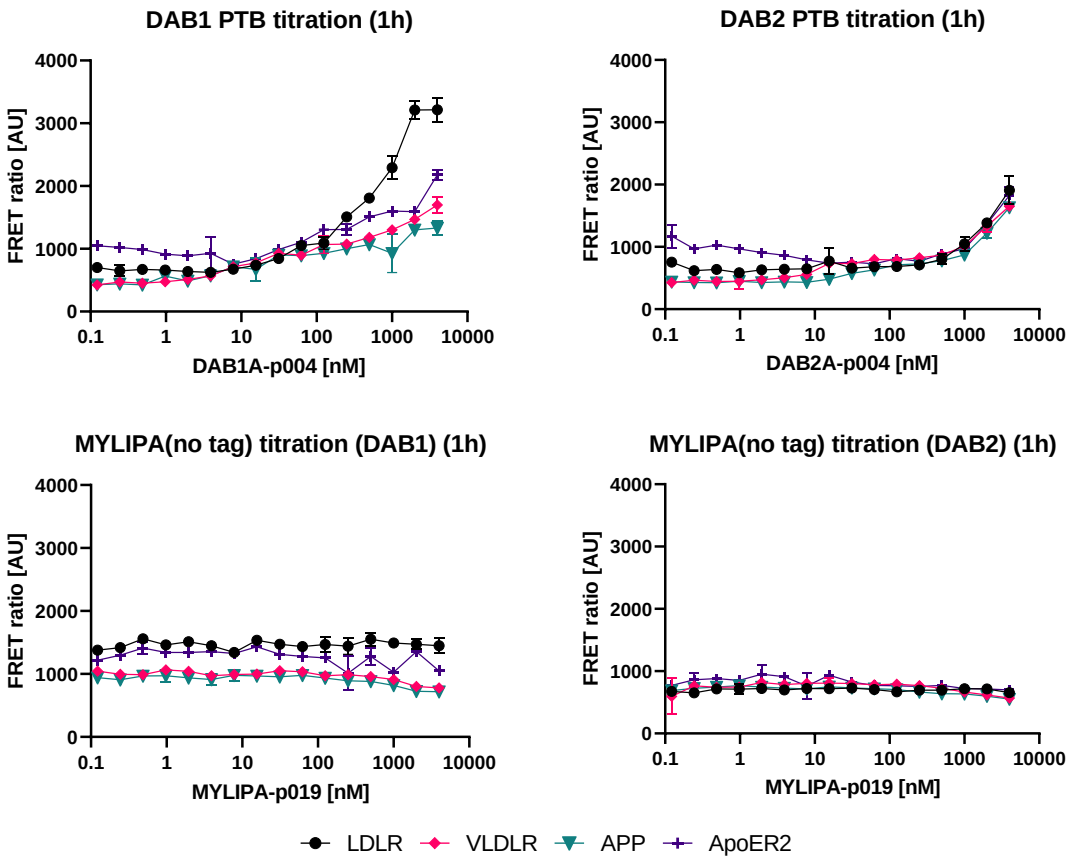


Figure 46: MYLIPA-p018 (1-298, T279D, S282D, C-His_6) (250 nM) and receptor peptides (62.5 nM). Top row: Titration of DAB1A-p004 (left) and DAB2-p004 (right) PTB. Bottom row: Titration of MYLIPA-p019 (1-298, T279D, S282D, C-His_6, His_6-tag cleaved) against MYLIP, receptor tail and DAB1 PTB (left) and DAB2 PTB (right) at 200 nM.

3.3.3 Other biophysical assays

3.3.3.1 Isothermal Titration Calorimetry (ITC)

Three ITC experiments were carried out with two different MYLIP constructs: MYLIPA-p002(1-276,pNIC-CTH0) and MYLIPA-p006(1-312,pNIC-ZB) and untagged $LDLR_{(811-834)}$. No precipitation was observed after dialysis. Post run, precipitate was observed in recovered cell contents. Buffer salt content was increased from 300 mM NaCl to 500 mM and buffer supplemented with 2 mM DTT and the cell temperature lowered to 15°C. Before the third run, the peptide precipitated in cold buffer and solubility could be recovered by pre-warming the buffer. Cell temperature was then increased to 25°C. After sudden change in heat rate, not

indicative of a gradual saturation of bindings sites, was observed, the run was aborted, and cell and syringe contents recovered. Both showed precipitate.

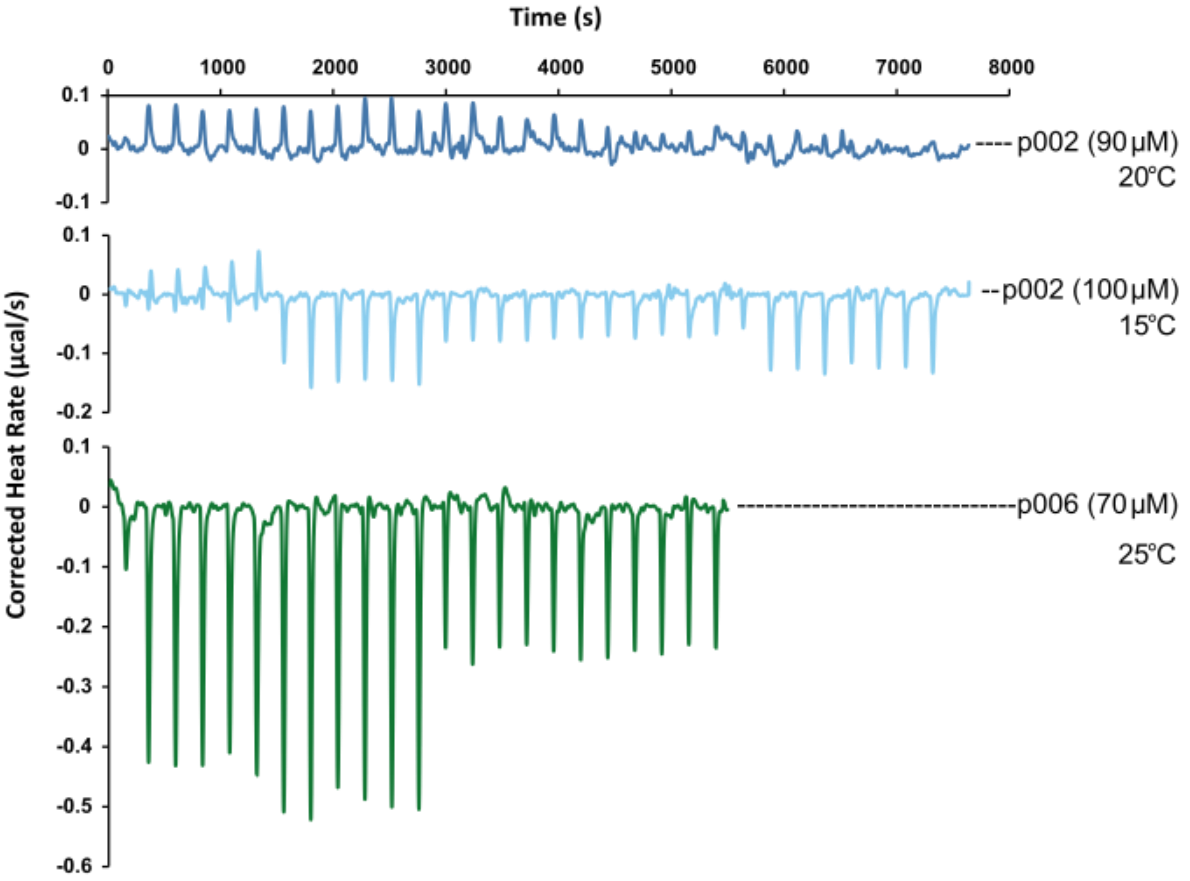


Figure 47: Isothermal titration calorimetry of $LDLR_{(811-833)}$ at 1 mM (against p002, 1-276, C-His_6) or 0.9 mM (against p006) into MYLIPA preps. Third run (p006, 1-250, N-His_6 ZB) was stopped early. Recovery of cell contents showed precipitation. Inspection of injection syringe showed precipitation of peptide.

3.3.3.2 Isothermal Fluorimetry

Stability of $LDLR_{(811-834)}$ was assessed via fluorimetry and light scattering at 20°C, with and without use of detergent (TWEEN-20) (**Figure 48**). Presence of detergent resulted in a forward shift of the onset of increase in light scattering and the decrease in fluorescence. The fluorescence intensity at the end of measurement was also significantly reduced by detergent for the highest peptide concentration.

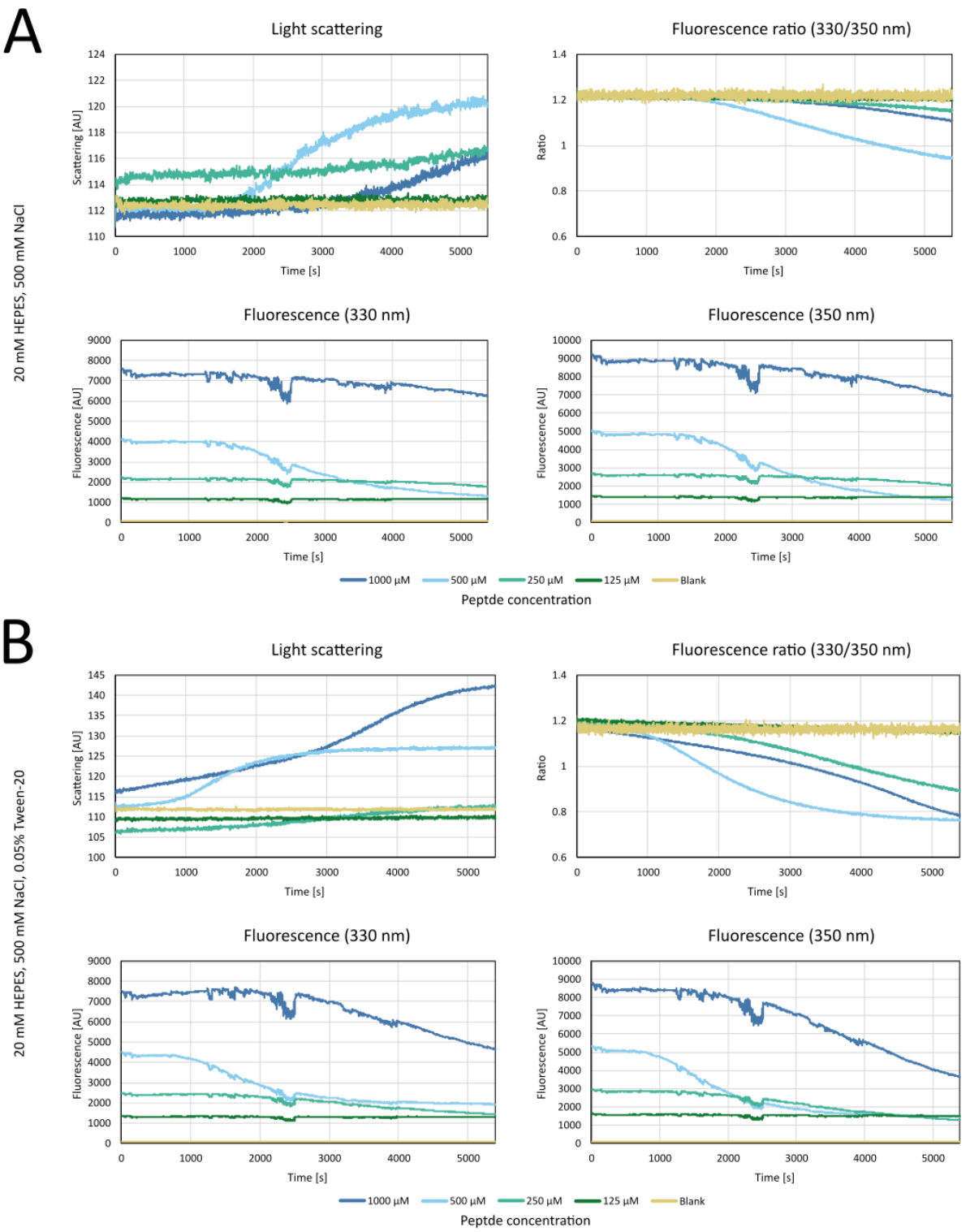


Figure 48: Light scattering and fluorescence of $LDLR_{(811-833)}$ at 20°C, in absence (**A**) and presence (**B**) of 0.05% TWEEN-20 in the assay buffer.

3.3.3.3 Microscale Thermophoresis

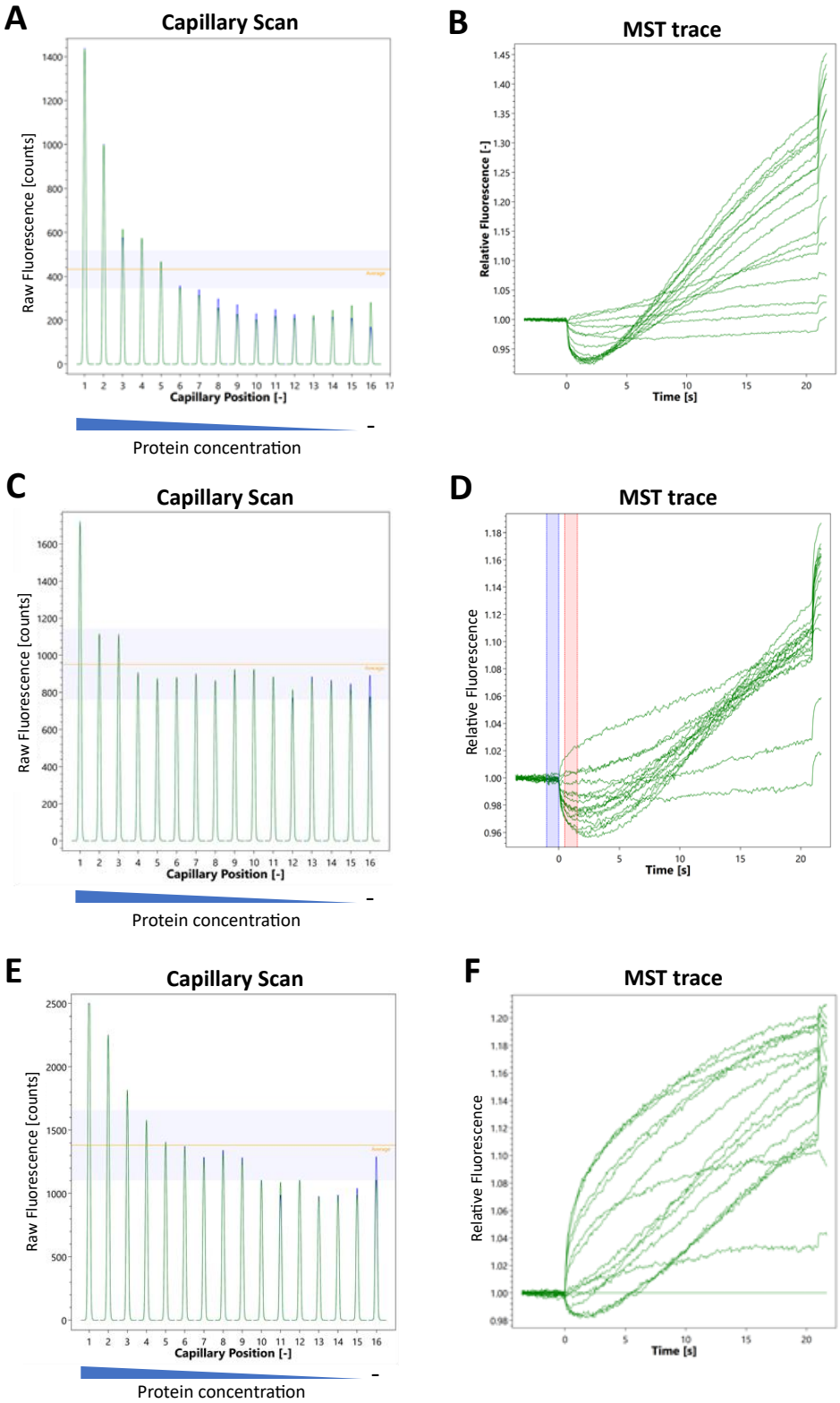
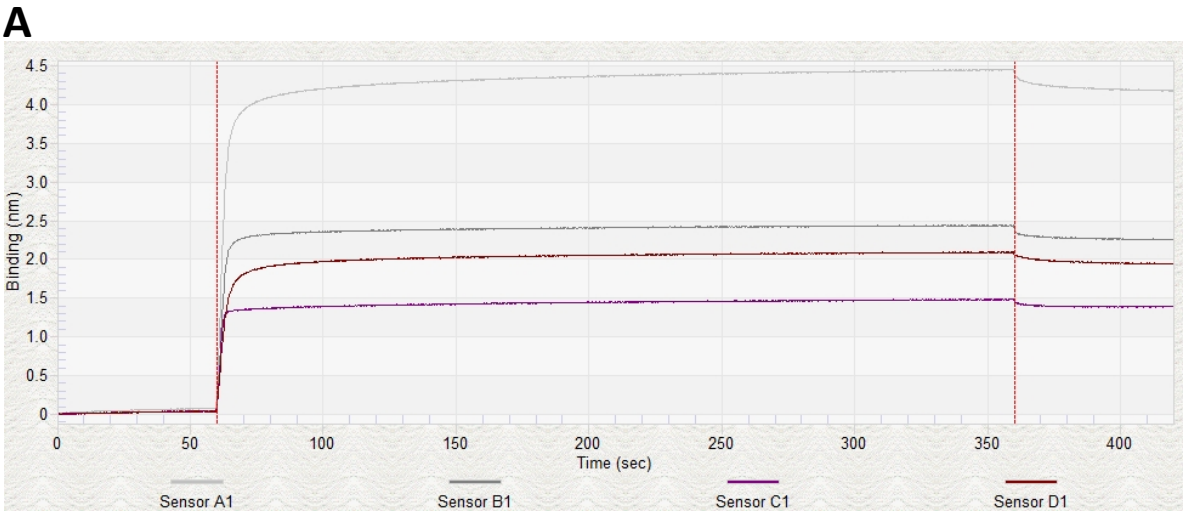


Figure 49: Microscale Thermophoresis serial dilutions (1:1) with MYLIPA-p003 (1-276, C-His_6) and Cy5.5-C-$LDLR_{(811-833)}$ at 10 µM in different buffers: **A**,**B**: MYLIP at 40 µM top concentration in 20 mM sodium phosphate, 150 mM NaCl, 1 mM TCEP, pH 7.5. **C**,**D**: MYLIPA at 40 µM top concentration in 50 mM Tris-HCl, 500 mM NaCl, 2 mM DTT, 10% glycerol, pH 7.5. **E**,**F**: MYLIPA at 100 µM top concentration in same buffer as **C**,**D**.

3.3.3.4 Biolayer interferometry

A binding check was performed with four MYLIP preps (p003: 1-296 C-His_6; p011: 1-276 C-His_6; p018: 1-298 T279D S282D C-His_6, p020: 1-445 N-His_6; **Figure 50**) at 40 µM and 20 µM. Super-streptavidin sensors were saturated with peptide ($LDLR_{(811-834)}$-K-biotin; biotin-$VLDLR_{(820-842)}$, biotin-$ApoER2_{(849-871)}$, biotin-$APP_{(745-767)}$). Binding was observed for all four constructs and $LDLR_{(811-834)}$-K-biotin and biotin-$APP_{(745-767)}$.

It was then attempted to determine K_D for MYLIPA-p003, p011 and the two binding peptides. Sensors were dipped into a 10 point 1:1 serial dilution, starting at 40 µM. Reliable fitting of the traces for $APP_{(745-767)}$ could not be achieved due to heterogeneous binding. K_D values for APP were not determined due to the heterogeneous binding behaviour observed. For biotin-$LDLR_{(811-834)}$, K_D values were obtained: 2.2 µM for MYLIPA-p003 (residues M1-S298) and 3.2 µM for MYLIPA-p011 (residues M1-R276).



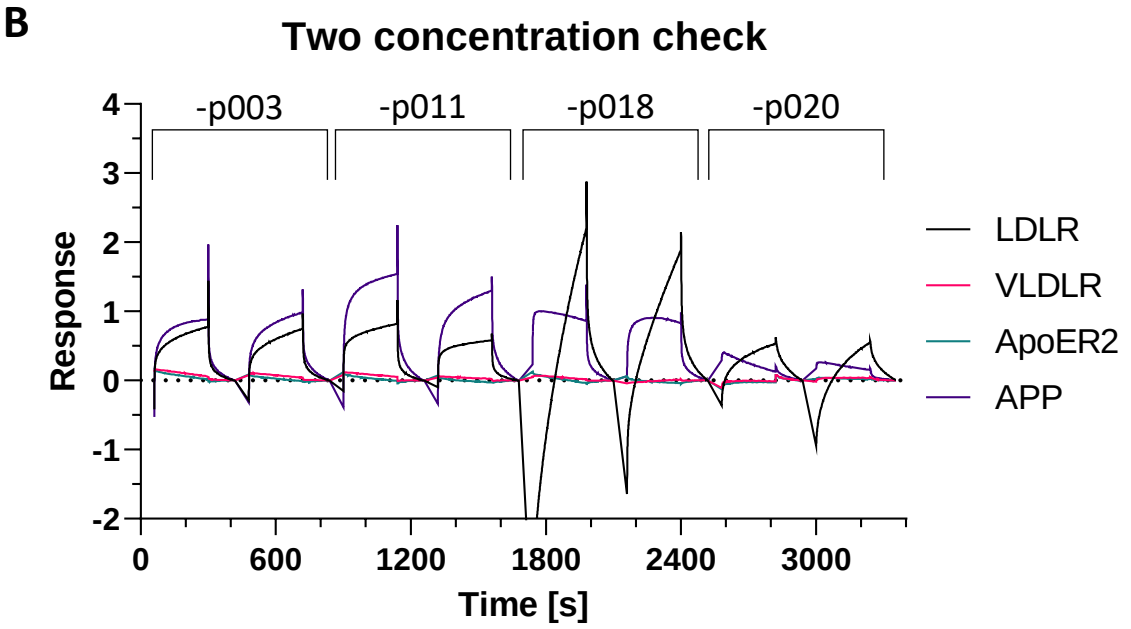


Figure 50: A: Sensogram of sensor loading with peptides (A1: LDLR$_{(811-834)}$-K-biotin; B1: biotin-VLDLR$_{(820-842)}$; C1: biotin-ApoER2$_{(849-871)}$; D1: biotin-APP$_{(745-767)}$). **B** Baseline corrected and y-axis aligned sensograms of two concentration check. MYLIP preps (p003: 1-296 C-His$_6$; p011: 1-276 C-His$_6$; p018: 1-298 T279D S282D C-His$_6$, p020: 1-445 N-His$_6$) were tested at 40 µM and 20 µM (high concentration first).

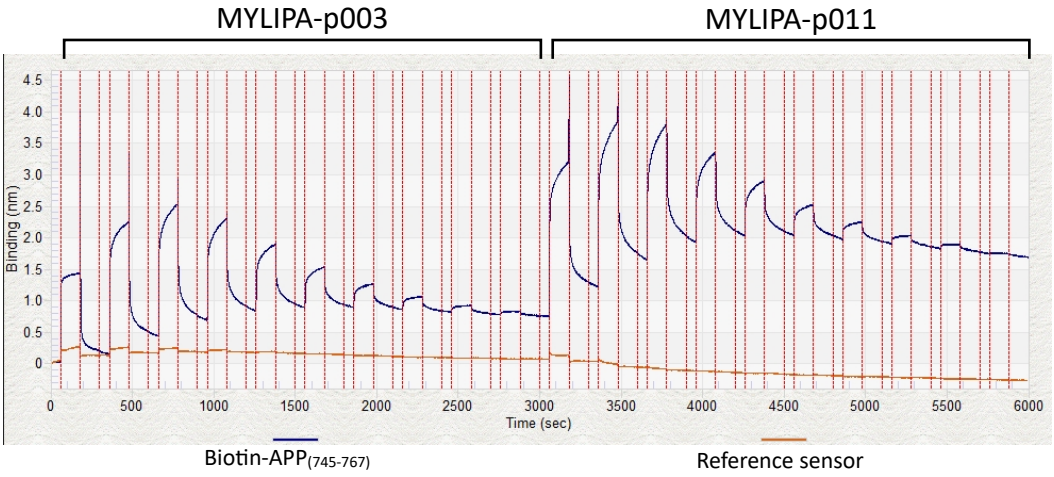
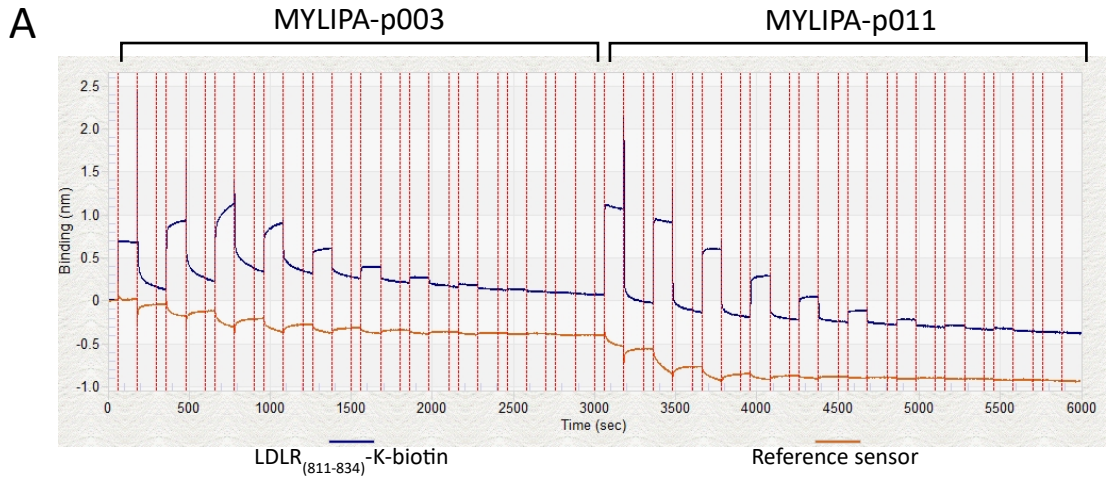


Figure 51: Sensogram for K_D determination experiment with MYLIPA-p003, MLIPA-p011 and biotin-APP$_{(745-767)}$



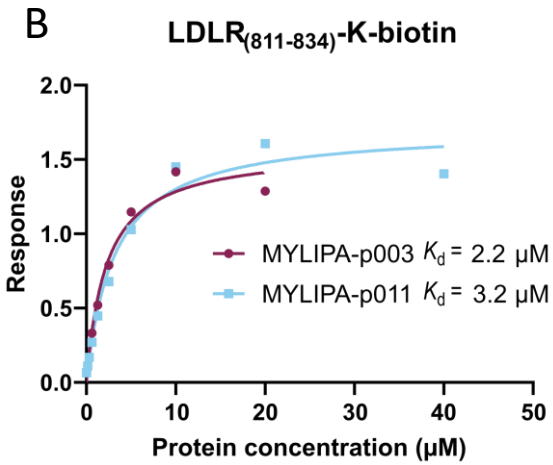


Figure 52: A: Sensogram for K_D determination experiment with MYLIPA-p003, MLIPA-p011 and LDLR$_{(811-834)}$-K-biotin. **B:** Plateau levels of observed binding response (nm) for different concentrations of MYLIP FERM domain fitted to a one-site association model using GraphPad PRISM.

3.3.4 Summary and Discussion

Successfully purified and stable MYLIP FERM domains were tested in protein-peptide titrations in AlphaScreen®, HTRF® and fluorescence polarisation assays, but no definitive binding could be detected. While two constructs gave positive results in AlphaScreen (M1-R276-His6 and M1-S298-His6), these results were singular to one peptide tested ($LDLR_{(811-859)}$-K-biotin) (**Table 22**) and could not be reproduced with the other substrate peptides, which should have a higher affinity for MYLIP FERM domain, or HTRF. Given the greater avidity of AlphaScreen (using micro-beads coated with multiple binding sites) in comparison to HTRF (using a bivalent antibody), it is not surprising that the two different assays would yield quantitatively different results. However, it is striking that no binding could be observed in HTRF at all. When it was attempted to out-compete the biotinylated receptor tail peptide with non-biotinylated peptide, the generated signal seemed to be not significantly affected by the presence of the non-biotinylated peptide.

Possible cooperative effects with the phospho-tyrosine binding domains, which are known to bind LDLR and associate on the cytosolic side of the plasma membrane, were then explored. Of these, Dab-1 and Dab-2 were successfully expressed and again tested against the receptor tail peptides, where binding could be observed (**Figure 43**, **Figure A39** to **Figure A43**). However, a specific co-operative effect of Dab-1 or Dab-2 PTB domains (after His_6-tag cleavage to eliminate α-His_6 antibody or AlphaScreen bead binding) with MYLIP FERM domain and the receptor tail peptides could not be observed (**Figure 44**). While an interaction between these three components *in vivo* cannot be excluded, it was not possible to obtain it *in vitro*.

3.4 Ubiquitination of receptor tail peptides by full length recombinant MYLIP

After the protein-peptide interaction assays did not yield results amenable to the development of a high-throughput assay, it was decided to investigate whether the in-house produced recombinant full length MYLIP ($MYLIP_{FL}$) could effectively ubiquitinate the receptor tail peptides in vitro and whether this reaction i. could be scaled down to 384 well microtiter plate compatible reaction volumes, ii. could be detected using available assay technologies, and ii. be validated by adding non-detectable receptor tail peptide to compete for the binding to $MYLIP_{FL}$, leading to a reduction in readout of the experiment.

3.4.1 Detection of LDLR ubiquitination via Western Blot

The first E1 purification was active despite being exposed to air on the column. Detection of LDLR-peptide ubiquitination ($Cy5.5\text{-}LDLR_{(811\text{-}833)}$) via Western blot showed a moderately greater activity of the in-house, *E. coli* produced $MYLIP_{FL}$ prep compared to the commercially available $MYLIP_{FL}$ procured from AMSBIO, which had been produced in eukaryotic cells (insect cell line Sf9). After overnight incubation, there is a marked reduction in ubiquitin in the sample, which can be explained by the precipitation observed in the sample.

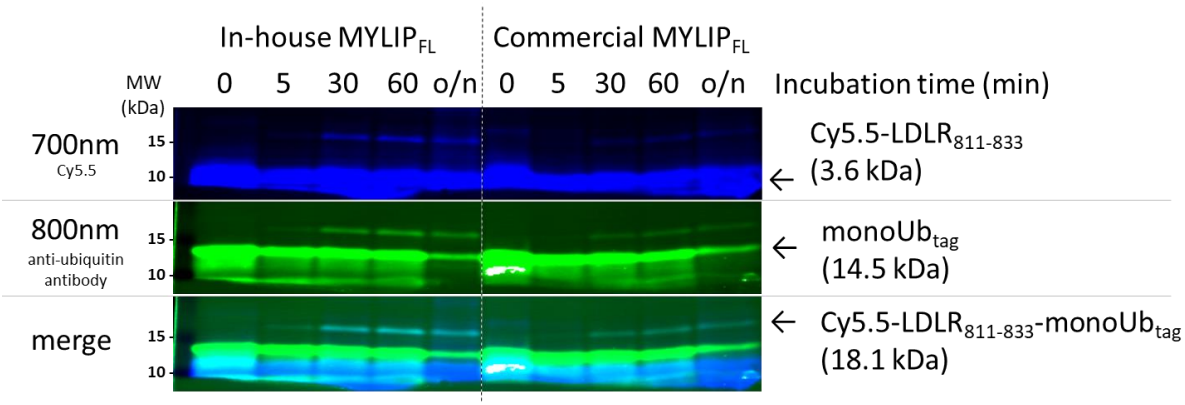


Figure 53: Detection of LDLR ubiquitination by in-house produced and commercially sourced recombinant $MYLIP_{FL}$ via Western Blot. Samples taken before addition of ATP (0), after 5, 30, and 60 minutes of incubation, as well as after over-night (o/n) incubation were blotted. Top Panel: 700 nm channel. Middle Panel: 800 nm channel. Bottom Panel: Merge.

3.4.2 HTRF based ubiquitination assay with discrete timepoints

The FRET signal for the $LDLR_{(811\text{-}833)}$-K-biotin reaction developed in line with the literature reference for the assay. For the other three peptides, no significant signal over baseline was

observed (**Figure 54**). For illustrative purposes, a one phase association model was fitted to the data using scipy.optimize.curve_fit():

$$Y = Y_{initial} + \left(Y_{plateau} - Y_{initial}\right) * (1 - e^{-K*x})$$

Equation 5

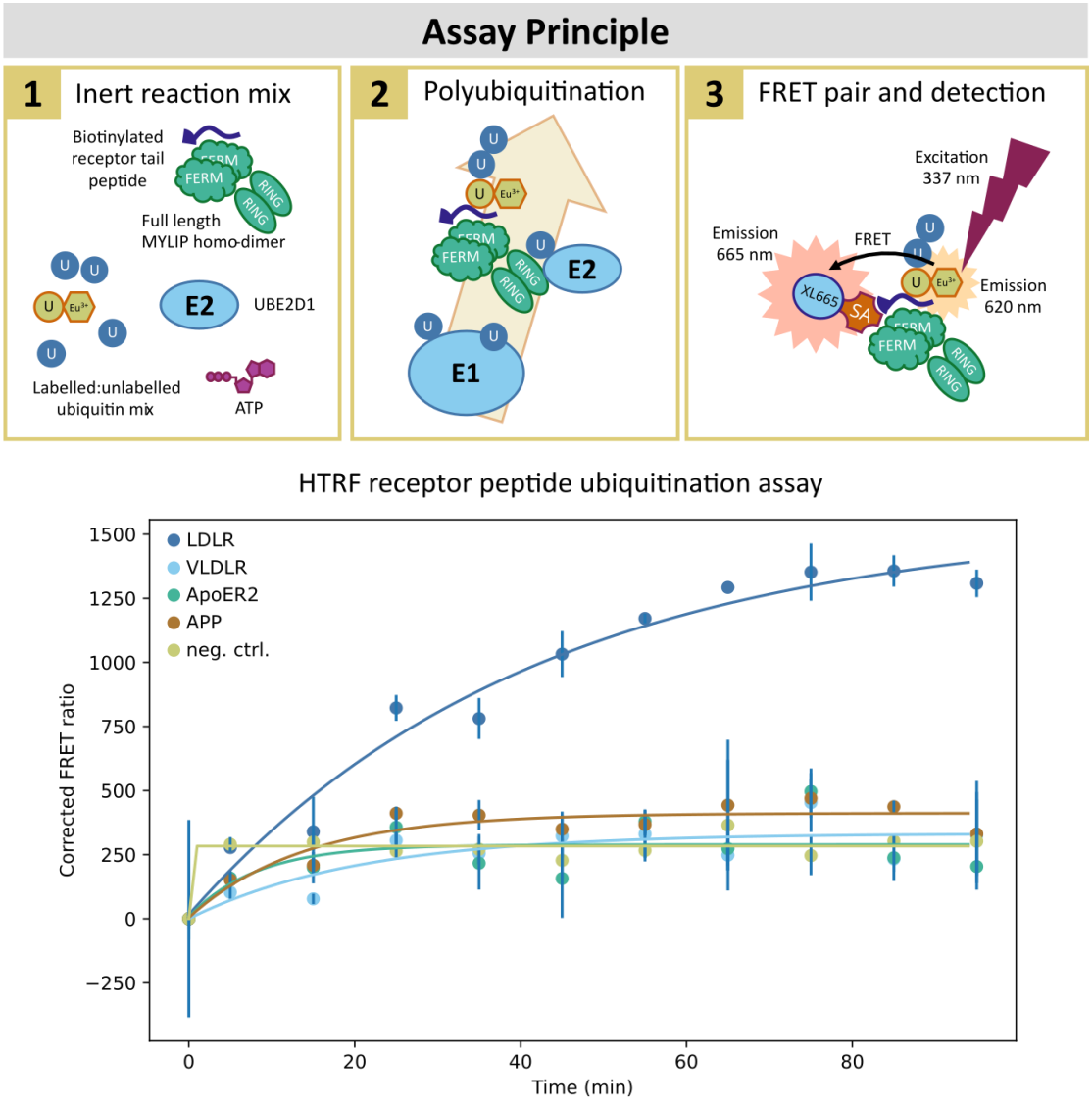


Figure 54: HTRF receptor peptide ubiquitination assay. **Top panel:** Recapitulation of Assay principle. **1:** Reaction mix is inert prior to addition of E1. **2:** Polyubiquitination of LDLR tail peptide commences upon addition of E1. **3:** After the reaction is stopped, streptavidin-XL665 (FRET acceptor) is added and it and Eu3+-cryptate labelled ubiquitin form FRET pairs. **Bottom panel:** Corrected FRET ratios, initial data point was subtracted as baseline. Peptides are **LDLR**: $LDLR_{(811-834)}$-K-biotin, **VLDLR:** biotin-$VLDLR_{(820-842)}$, **ApoER2:** biotin-$ApoER2_{(849-871)}$, **APP:** biotin-$APP_{(745-767)}$. **neg. ctrl:** $LDLR_{(811-834)}$-K-biotin without addition of ATP.

3.4.3 Continuous HTRF based ubiquitination assay

The first attempt of setting up a continuous assay showed good progress curves for $LDLR_{(811-834)}$-K-biotin and btn-$VLDLR_{(820-842)}$ substrate peptides that appeared to reach a plateau after 60 minutes. Product formation seemed to plateau for all concentrations of E1 around the same time.

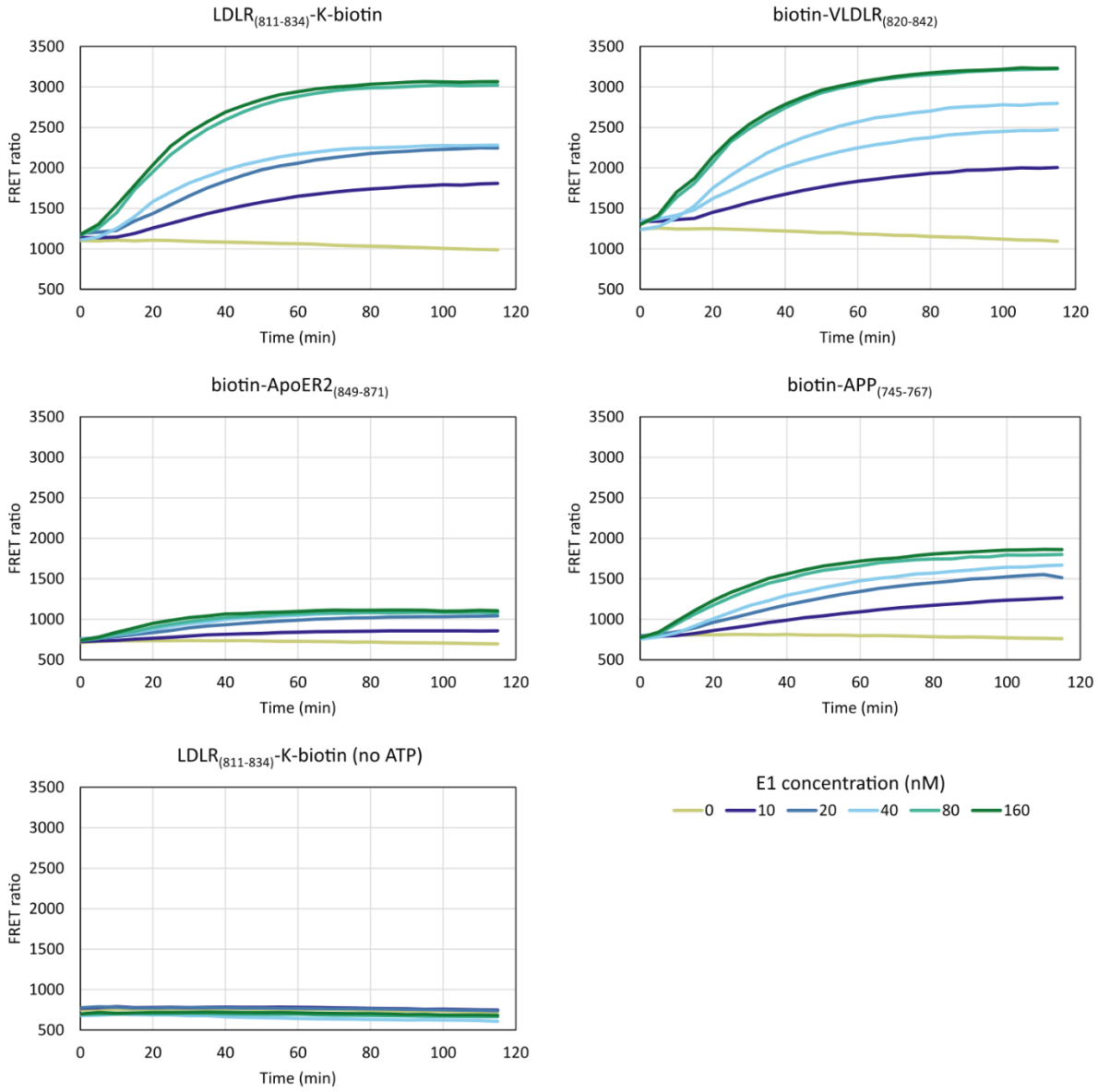


Figure 55: First continuous HTRF receptor peptide ubiquitination assay. Ubiquitin chain formation appears to reach a plateau at around 80 minutes.

3.4.3.1 Reagent optimisation for continuous assay

As before, product formation seemed to cease between 60 - 80 minutes (**Figure 55**). This is consistent with the literature pointing towards an inefficient reaction and previous experiments, but it needs to be investigated.

Overall signals are much higher than in previous experiment (3.4.3). Peptides were taken from 1 mM DMSO stock this time whereas previously they were taken from 1:100 pre-diluted aliquots that have undergone several freeze-thaw cycles. Pipetting inaccuracies are also likely a contributing factor at such small volumes.

Increasing the concentration of labelled ubiquitin correlates with a decrease in baseline FRET ratio (**Figure 56** to **Figure 59**). (both initial signal at start of the reaction and baseline signal with 0 nM mUBE1, see **Figure 60**, **Figure 61** and **Figure 62**).

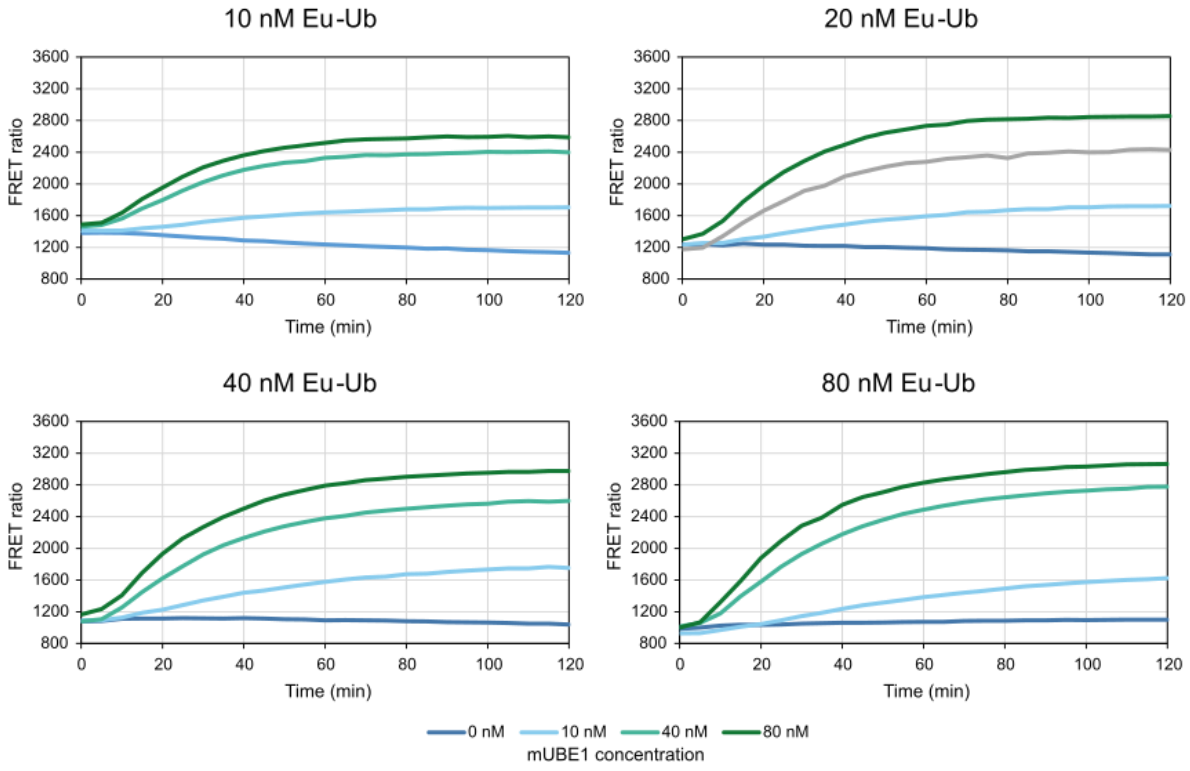


Figure 56: Titration of E1 and Eu^{3+}-cryptate labelled ubiquitin in continuous HTRF assay with biotin-$LDLR_{(811-834)}$ substrate peptide.

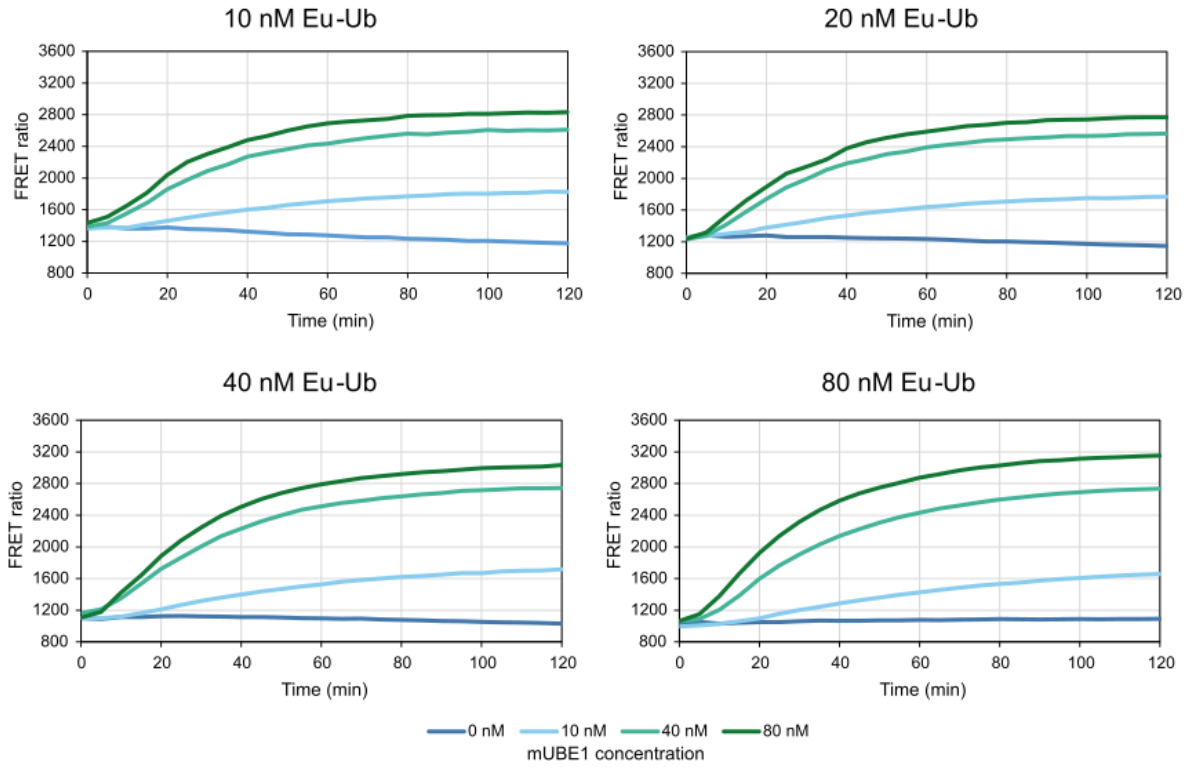


Figure 57: Titration of E1 and Eu^{3+}-cryptate labelled ubiquitin in continuous HTRF assay with $VLDLR_{(820-842)}$-K-biotin substrate peptide.

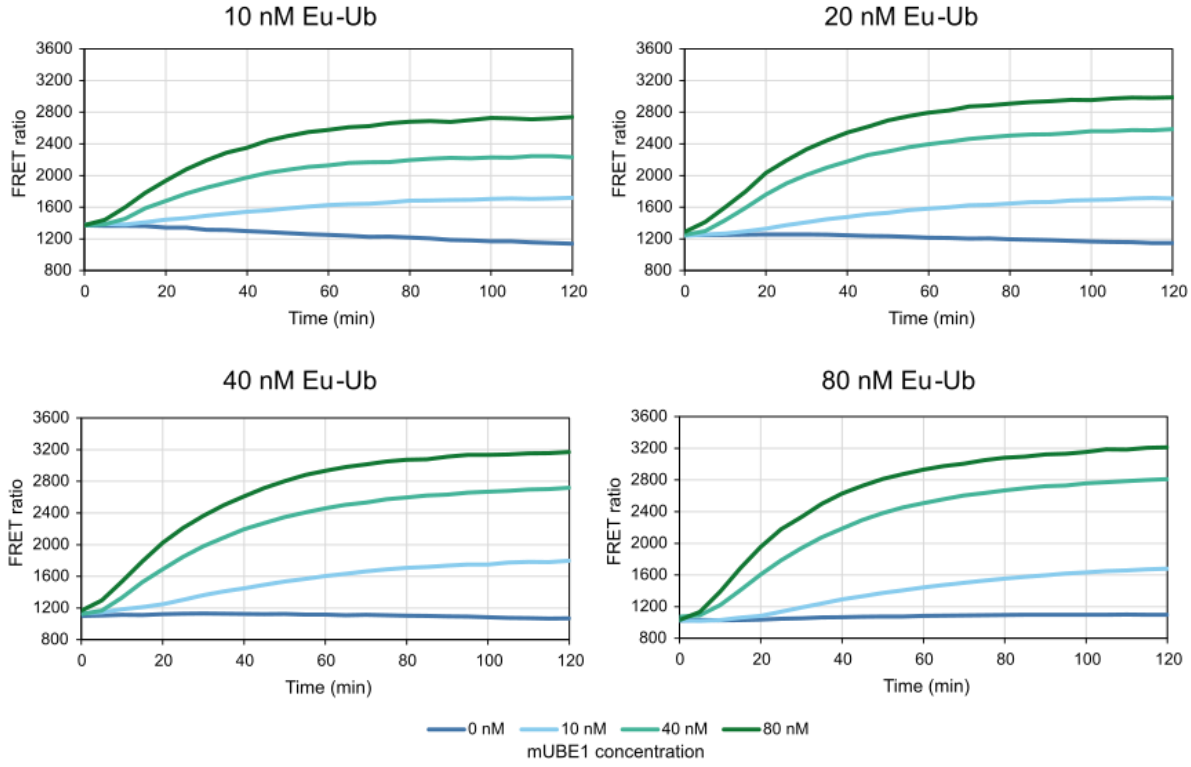


Figure 58: Titration of E1 and Eu^{3+}-cryptate labelled ubiquitin in continuous HTRF assay with $ApoER2_{(489-871)}$-K-biotin substrate peptide.

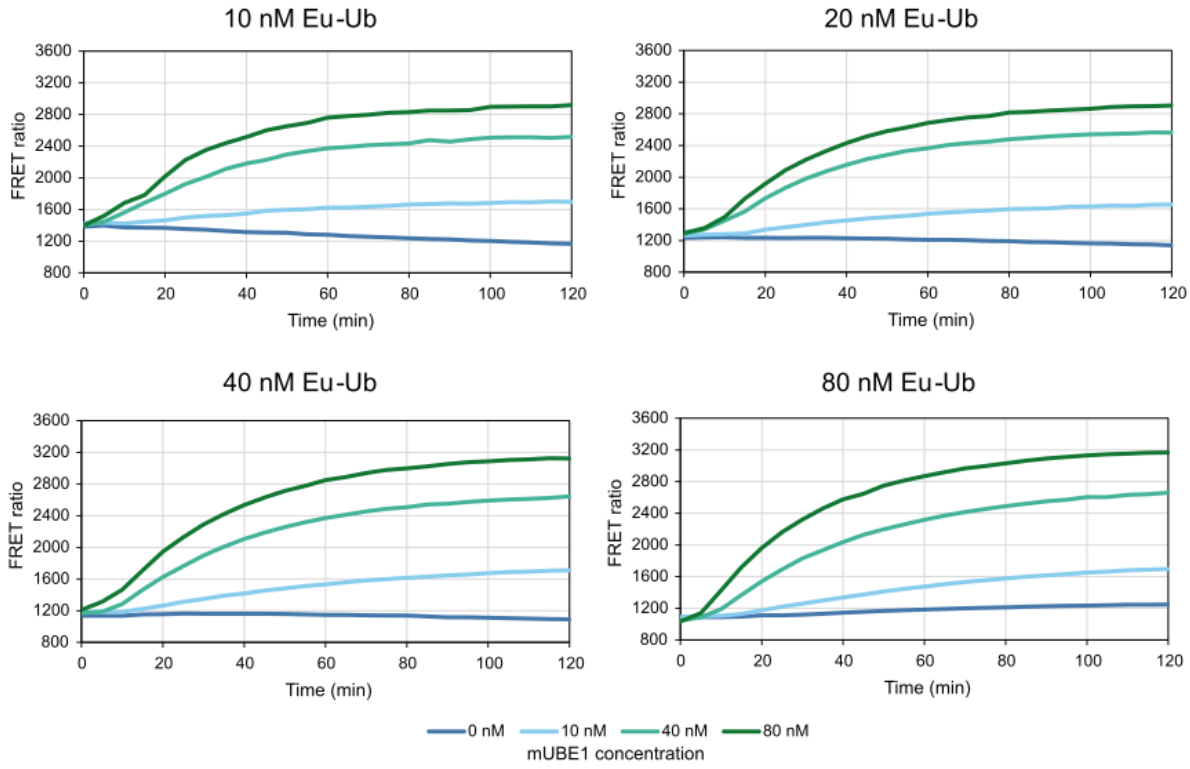


Figure 59: Titration of E1 and Eu^{3+}-cryptate labelled ubiquitin in continuous HTRF assay with $APP_{(745-767)}$-K-biotin substrate peptide.

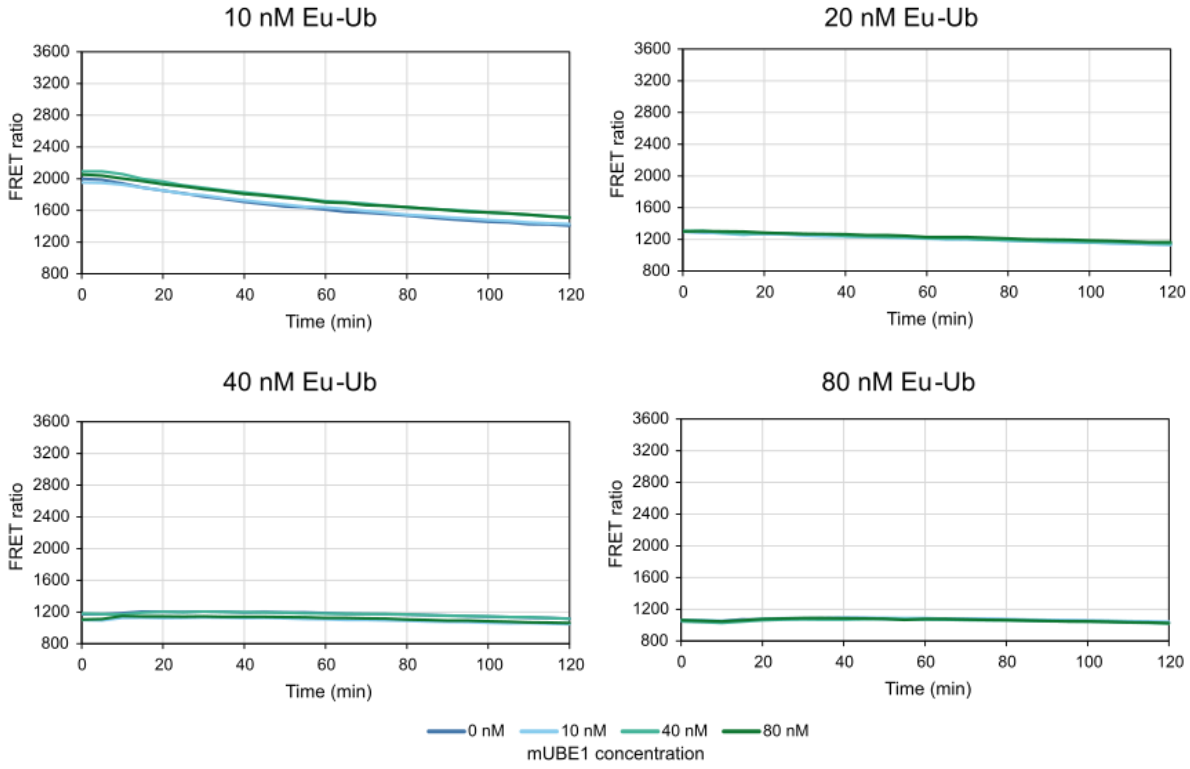


Figure 60: Titration of E1 and Eu^{3+}-cryptate labelled ubiquitin in continuous HTRF assay with biotin-$LDLR_{(811-834)}$ substrate peptide in absence of ATP (negative control)

Ratio (Final:Initial)

Eu-Ubi	10				20				40				80			
mUBE1	0	10	40	80	0	10	40	80	0	10	40	80	0	10	40	80
LDLR	0.9	1.15	1.57	1.67	0.97	1.27	1.92	2.07	1.02	1.41	2.16	2.35	1.09	1.45	2.4	2.78
VLDLR	0.86	1.33	1.88	1.98	0.91	1.41	2.09	2.24	0.94	1.56	2.36	2.73	1.05	1.67	2.58	2.96
ApoER2	0.83	1.25	1.63	1.99	0.92	1.38	2.07	2.32	0.97	1.59	2.42	2.72	1.07	1.64	2.62	3.12
APP	0.84	1.19	1.8	2.09	0.93	1.31	2	2.24	0.96	1.44	2.26	2.58	1.15	1.56	2.55	3.06
LDLR(no ATP)	1.02	1	1.08	1.08	0.92	0.9	0.89	0.9	0.98	0.88	0.95	0.88	0.95	0.95	0.98	0.99

Ratio (X nM mUBE1 : 0 nM mUBE1)

Eu-Ubi	10				20				40				80			
mUBE1	0	10	40	80	0	10	40	80	0	10	40	80	0	10	40	80
LDLR	1	1.3	1.83	1.99	1	1.31	1.89	2.25	1	1.4	2.11	2.48	1	1.26	2.27	2.59
VLDLR	1	1.55	2.22	2.41	1	1.54	2.24	2.42	1	1.66	2.66	2.94	1	1.52	2.5	2.89
ApoER2	1	1.51	1.96	2.4	1	1.49	2.26	2.61	1	1.68	2.54	2.96	1	1.53	2.56	2.93
APP	1	1.46	2.16	2.5	1	1.45	2.25	2.55	1	1.57	2.42	2.86	1	1.36	2.13	2.54
LDLR(no ATP)	1	1.02	1.07	1.07	1	1	1.01	1.02	1	0.94	1	0.95	1	1.01	0.99	0.99

Figure 61: Signal to background ratios for reagent optimisation of continuous HTRF ubiquitination assay. Top panel: Ratio of final to initial signal. Bottom panel: Final ratios at X nM E1 to final ratios at 0 nM E1)

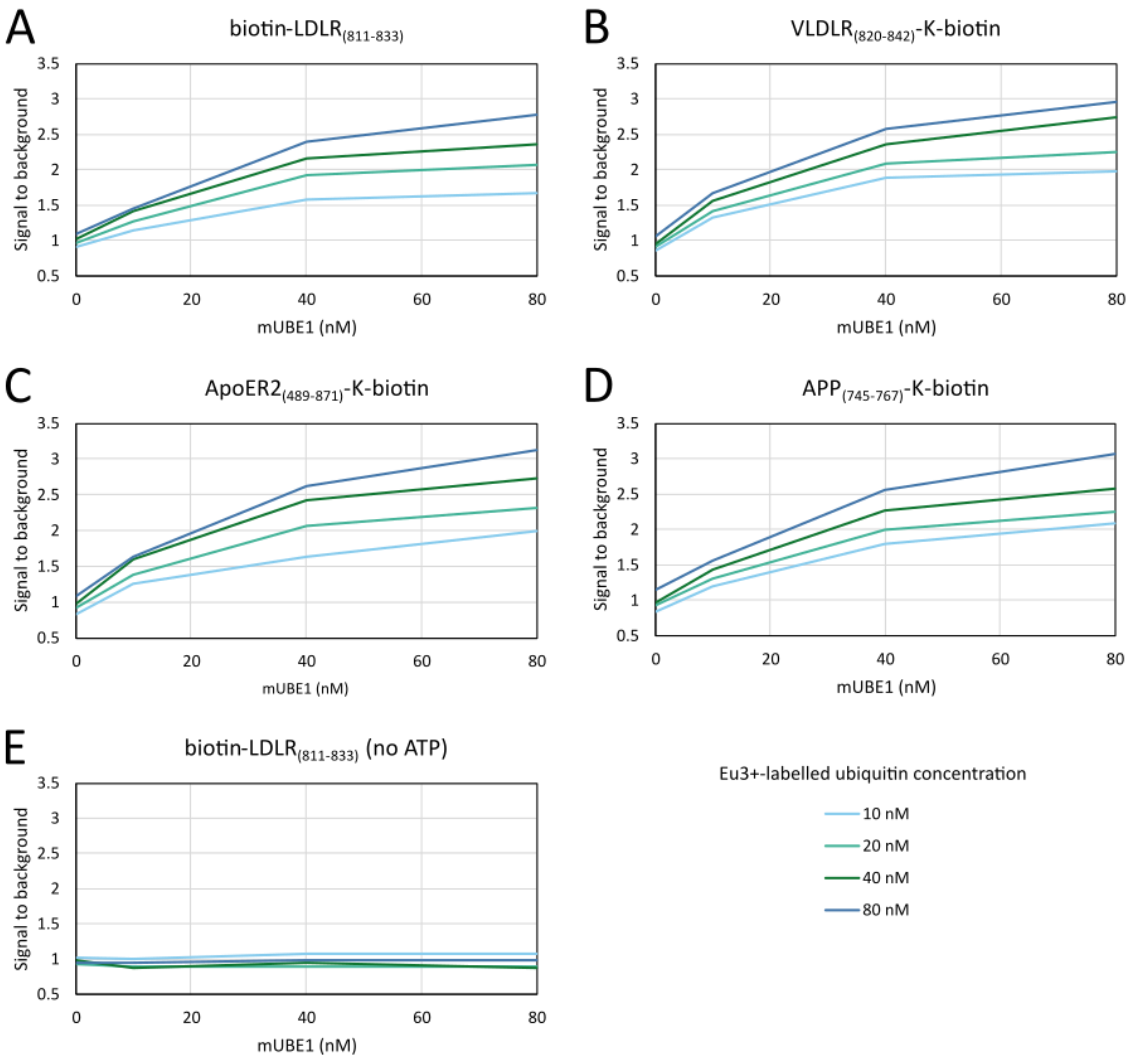


Figure 62: Change in signal to background at endpoint of continuous HTRF ubiquitination assay dependent on concentration of E1 and labelled ubiquitin.

3.4.3.2 Validation of new E1 batch

In the first instance, the gel fluorescence assays showed substrate peptide ubiquitination with MUBE1A-p002 and all three UBE2D isoforms. MUBE1A-p002 was tested at 7 concentrations plus a no-enzyme negative control. For the highest concentrations, the FRET ratio reached a plateau at around 30 min (**Figure 63**). For up to 15 minutes, the FRET ratio appeared to increase in a linear fashion. There was no significant increase in signal between 40 nM and 80 nM MUBE1 in the assay, so 40 nM MUBE1 were selected to move forward with. For potential endpoint assays (i.e., where the reaction is stopped after a set time and then only read once), 15 min was considered as a stopping point for the reaction.

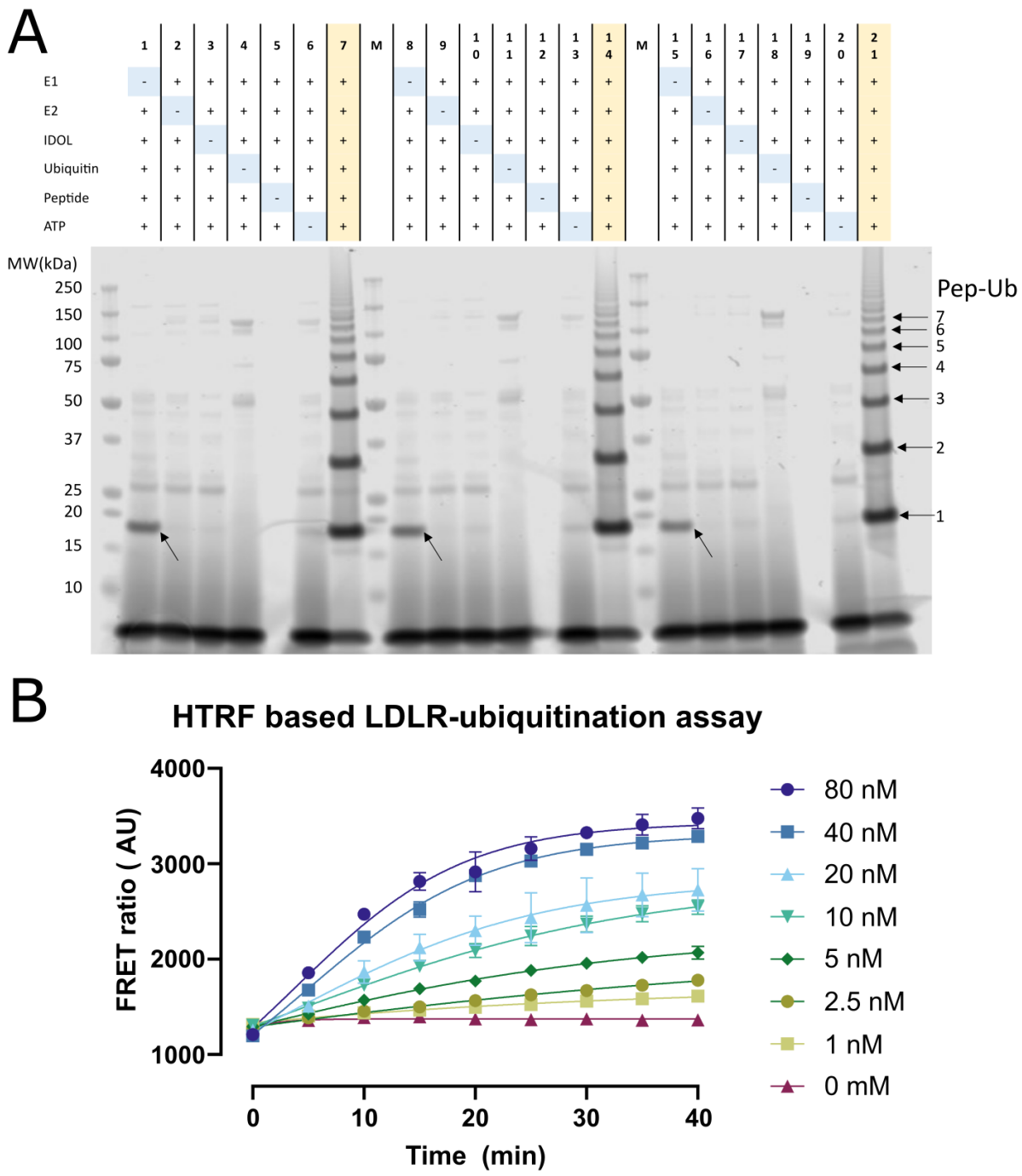


Figure 63: A: Gel fluorescence-based evaluation of MUBE1A-p002 with all three UBE2D isoforms tested in this work. Polyubiquitination states of the CY5.5-LDLR(811-834) peptide are clearly discernible for up to seven ubiquitin units. Theoretical MWs of poly-ubiquitinated peptides: Pep-Ub_1:18.1 kDa, then a further 14.5 kDa for each ubiquitin unit. A band corresponding to mono-ubiquitination can be observed in the no-E1 lane. **B:** Titration of MUBE1A-p002 in continuous, HTRF based IDOL-LDLR ubiquitination assay.

3.4.3.3 Competition assay with non-biotinylated peptide

A non-biotinylated peptide concentration dependent reduction in FRET signal could be observed when $LDLR_{(811-834)}$ was titrated (400 µM to 1.6 µM) into the reaction of a continuous assay (**Figure 64**). All reactions appeared to reach a plateau at roughly the same time point, albeit earlier than in previous experiments (**Figure 63**).

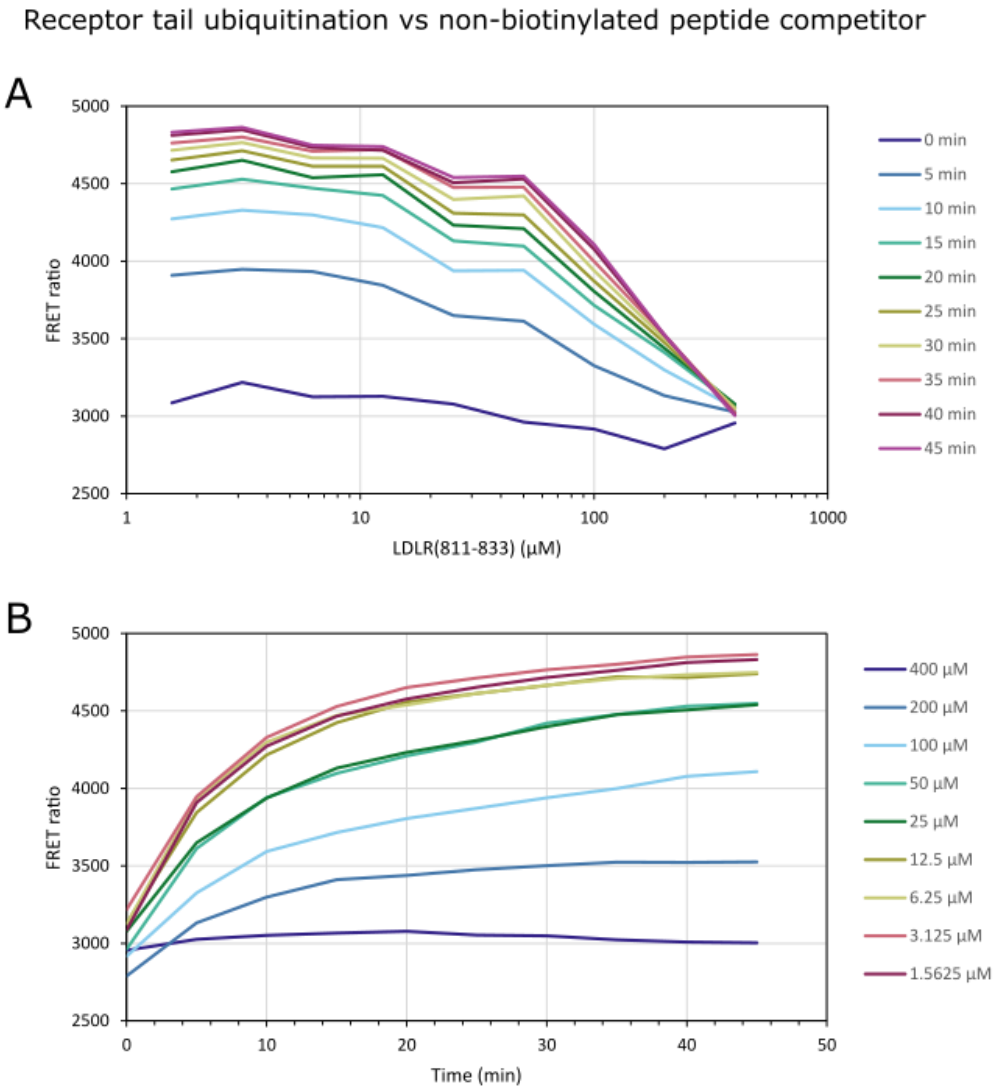


Figure 64: Competition of biotinylated and non-biotinylated LDLR tail peptide in continuous ubiquitination assay. **A**: FRET ratios plotted against concentration of non-biotinylated peptide. **B**: FRET ratio plotted against time.

3.4.4 Summary and Discussion

After the unsuccessful attempts to establish a receptor-tail-MYLIP binding assay and under the pressures of the COVID-19 pandemic, an enzymatic approach was pursued and murine E1, human UBE2D (isoforms 1, 2 and 3) and human mono-ubiquitin were produced in *E. coli* cells (3.2.3). Western Blot and in-gel fluorescence assays with Cy5.5 labelled $LDLR_{(811-834)}$ peptide showed clearly discernible poly-ubiquitination states of the receptor tail and the full length MYLIP produced in E. coli cells showed comparable activity to commercially sourced MYLIP produced in eukaryotic SF9 cells(**Figure 53**). Using $Europium^{3+}$-cryptate conjugated human ubiquitin, it was then attempted to establish a plate-based HTRF assay. Initial experiments were promising, and the results resembled the results originally published by the developers of the method (**Figure 54**). Attempts to optimize the signal-to-baseline window of the assay were undertaken and an acceptable ratio of 3:1 was reached (**Figure 61**) and titration of non-biotinylated peptide as a surrogate inhibitor showed a dose dependent reduction in FRET ratio (**Figure 64**). Attempts to identify possible unknown substrates of MYLIP within the LDLR super-family led to the discovery that even non-substrates produced signal in this assay. It was then discovered that part of the affinity-tag of the human ubiquitin construct used (TwinStrep) can bind streptavidin and thus streptavidin-XL665, leading to the creation of FRET fluorophore pairs along the length of the poly-ubiquitin chain on both the genuine substrate as well as in case of auto-ubiquitination of MYLIP.

3.5 Using the MYLIP autoubiquitination assay to screen for MYLIP inhibitors

After it was discovered that the supposed substrate ubiquitination assay in fact produced FRET fluorophore pairs along the polyubiquitin chain generated by MYLIP's auto-ubiquitination function, it was decided to optimise this assay as a primary assay to screen a subset of the ODDI 10k diversity library in a pilot screen, with the outlook of developing new assays to separate genuine MYLIP inhibitors from inhibitors of other steps of the ubiquitination cascade.

3.5.1 Time window for auto-ubiquitination assay

To determine the optimal time for the reaction in the pilot screen, MYLIP-auto-ubiquitination reactions were set up and samples taken at discrete time points. Detection reagent was added later and further incubated. Data are plotted in Figure and 30 min was chosen as the incubation time for the auto-ubiquitination reaction.

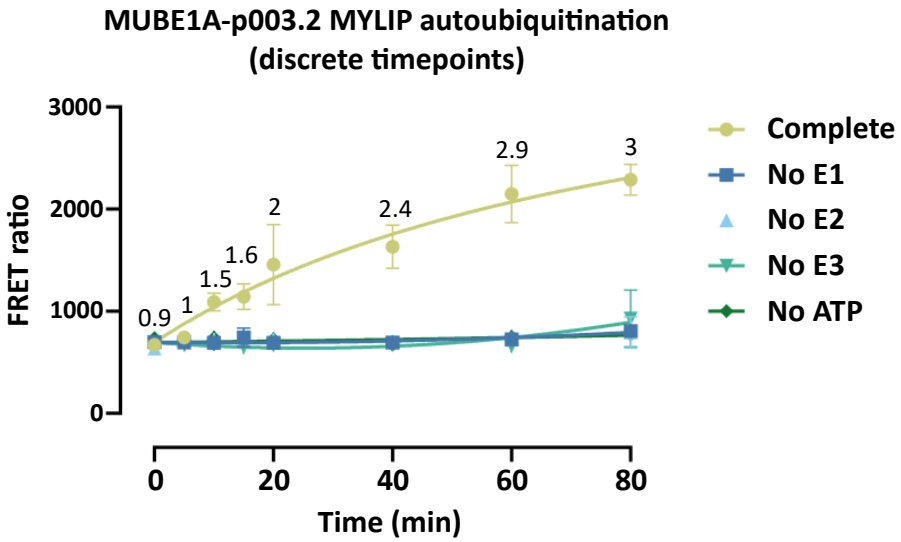


Figure 65: MYLIP auto-ubiquitination assay with discrete timepoints. Error bars represent standard deviation over three technical replicates. Numbers above datapoints represent ratio of complete reaction datapoint to no-ATP control.

3.5.2 MYLIP auto-ubiquitination assay pilot screen

Reagents were dispensed on eight plates as planned. After recovering the dead volume from the liquid handler, reagents were dispensed on two more plates using electronic multichannel pipettes. Due to air bubbles being aspirated, plate 10 was entirely excluded from analysis.

Due to a blockage in the liquid handler cassette, rows C and D of plates 1 – 8 had uneven liquid levels dispensed. After sighting the plates as heat maps in MS Excel and looking at the

manually calculated Z' values (**Table 23**), it was decided to remove certain wells and rows from analysis. Plate 1 was satisfactory with a Z' > 0.5. On Plate 2, well C1 was excluded as fluid levels and signal levels seemed not overly affected. On Plates 3 – 8, the entire rows C and D were excluded as the fluid levels were significantly affected. After this correction, the Z' was satisfactory for all plates (**Table 23)** and plates 1 – 9 were analysed with the "End Point Single Dose" workflow in BBQ. This workflow also calculated the Z' (robust), also shown in **Table 23.**

To identify any systemic errors, the median normalised FRET ratio of each well was calculated across all eight plates where reagents have been dispensed with the Multidrop Combi liquid dispenser (**Figure 67C**). The area affected by uneven dispensing due to the blocked nozzle is clearly visible.

Table 23: Comparison of Z' values with and without excluded wells (calculated in Excel) and Z' (robust) (*wells excluded,* calculated with BBQ workflow).

	Z' with wells…		
Plate	**included**	**excluded**	**Z' (robust)**
Plate 1	0.541	0.541	0.849
Plate 2	0.303	0.711	0.885
Plate 3	0.639	0.770	0.919
Plate 4	0.252	0.835	0.896
Plate 5	0.300	0.886	0.929
Plate 6	0.270	0.824	0.903
Plate 7	0.265	0.828	0.93
Plate 8	0.497	0.826	0.944
Plate 9	0.562	0.562	0.725
Plate 10	0.142	N/D	N/D

N/D = not determined

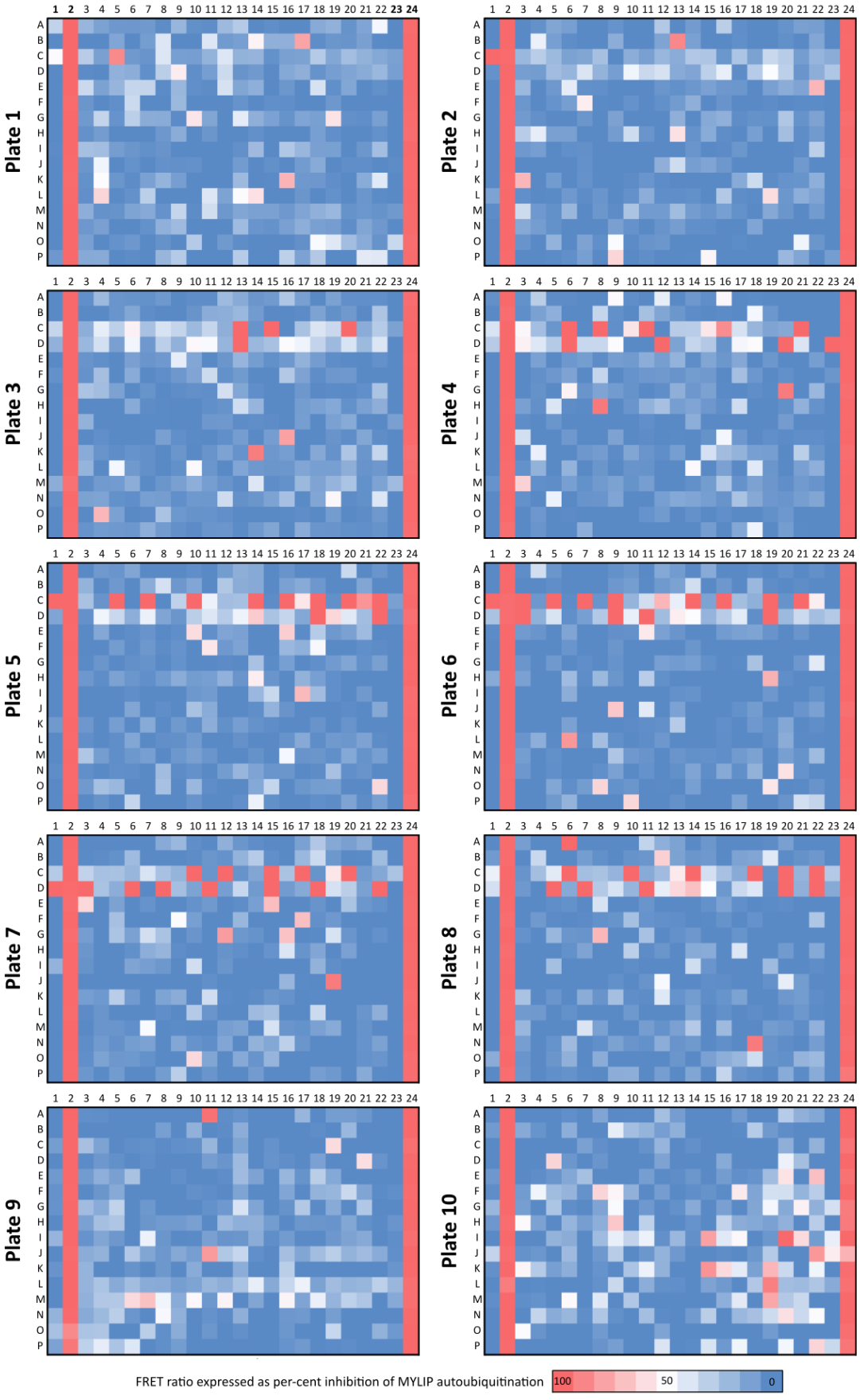


Figure 66: Heat maps of MYLIP auto-ubiquitination assay plates. Values shown are FRET ratio normalised as per-cent inhibition of enzymatic activity.

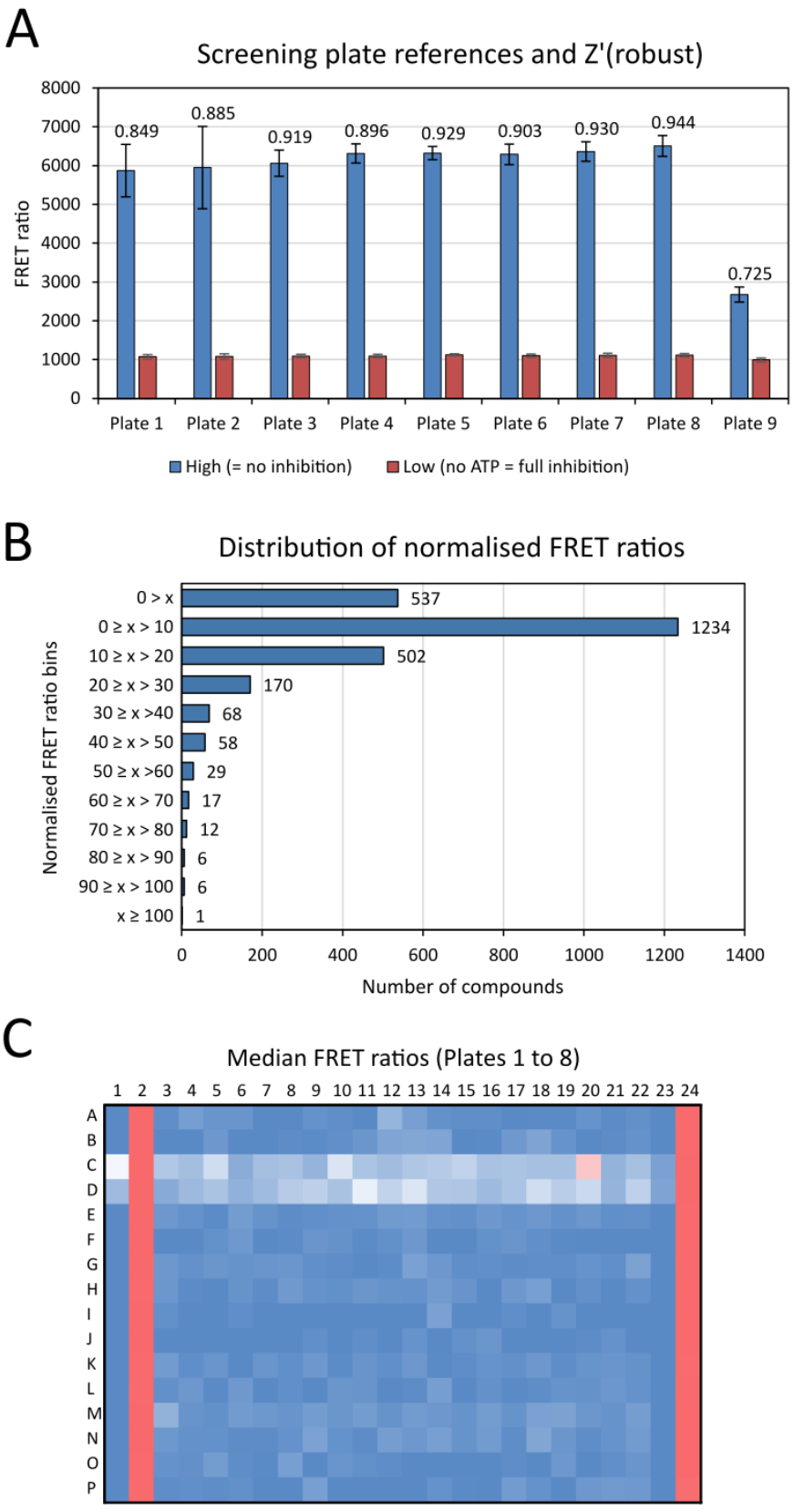


Figure 67: Data quality of the MYLP auto-ubiquitination pilot screen. **A** High (no inhibition) and low (full inhibition) signal references. Error bars are standard deviation. Values above bars are Z'(robust) values. **B:** Distribution of normalised FRET ratios in MYLIP auto-ubiquitination pilot screen. **C**: Median normalised FRET ratios for plates 1 to 8 (reagents dispensed with automated liquid dispenser).

3.5.3 Identified "Hits"

Compounds were then ranked by the calculated per-cent inhibition of auto-ubiquitination. The top 50 compounds and their inhibition values are listed in **Table 24** and their structures are shown in **Figure 69**, **Figure 70** and **Figure 71**. Structures were drawn using the Python library PIKAChU[163]. Of the 2640 compounds that were screened, only 25 (0.95%) showed a reduction in auto-ubiquitination of 70% or greater compared to reference. Performing multidimensional scaling, hierarchical clustering. and distance matrix analysis with ChemMine Tools [164], revealed no obvious similarities in these "hits" (**Figure 72**).

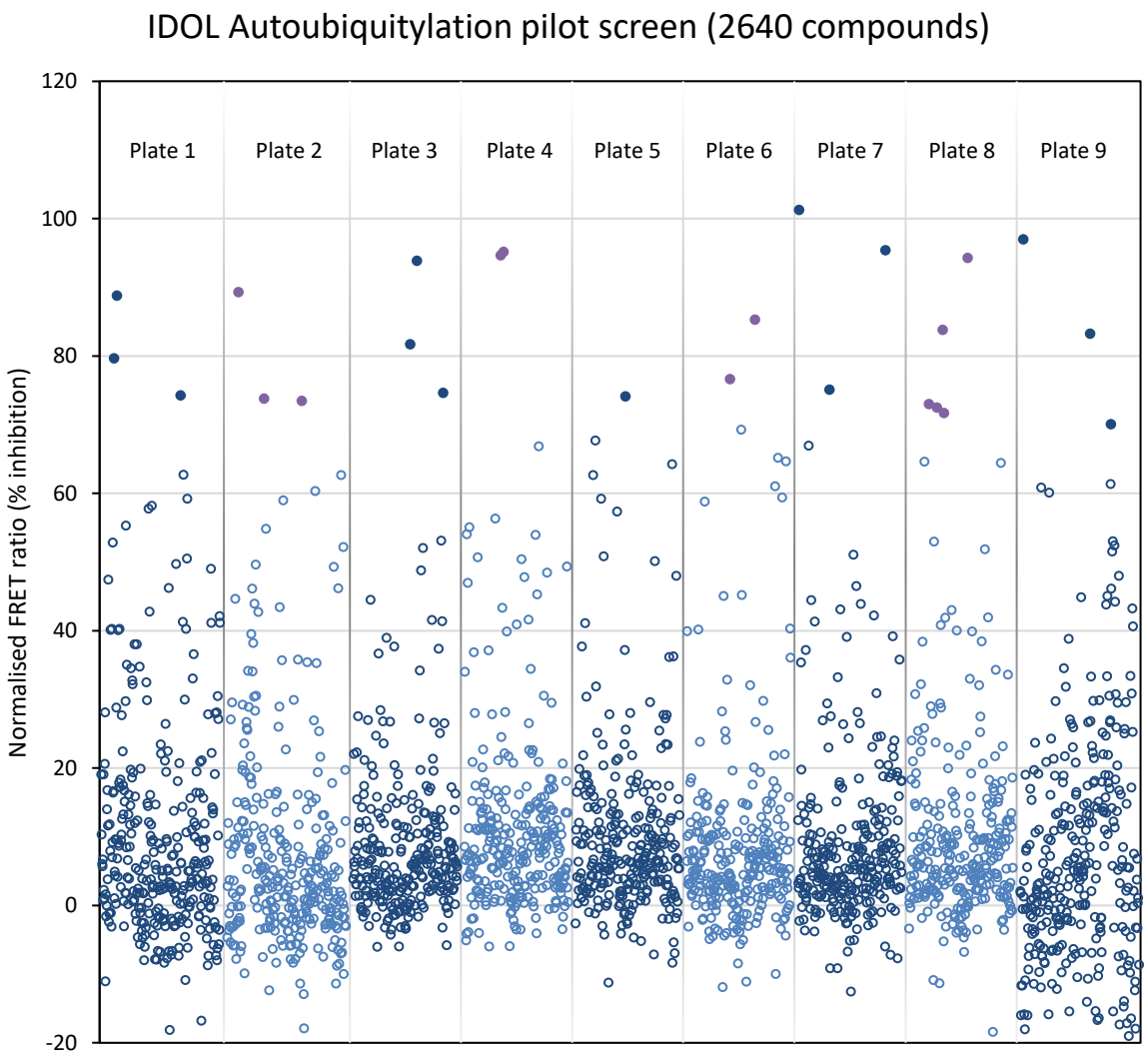


Figure 68: Results of the MYLIP auto-ubiquitylation screen (Plates 1- 9). Compounds with inhibition of enzymatic function > 70 % are shown as filled circles.

Table 24: Top 50 compounds of plates 1-9 of the MYLIP auto-ubiquitination screen.

Rank	Compound ID	Normalised FRET ratio (% inhibition)	Plate	Rank	Compound ID	Normalised FRET ratio (% inhibition)	Plate
1	ZZ003212a	101.2	7	26	ZZ003077a	69.21	6
2	ZZ003850a	96.91	9	27	ZZ002675a	67.66	5
3	ZZ003481a	95.33	7	28	ZZ003237a	66.91	7
4	ZZ002409a	95.11	4	29	ZZ002503a	66.81	4
5	XS338417b	94.59	4	30	ZZ003175a	65.13	6
6	ZZ003721a	94.22	8	31	ZZ003196a	64.62	6
7	ZZ002162a	93.8	3	32	ZZ003606a	64.58	8
8	ZZ001678a	89.27	2	33	ZZ003810a	64.37	8
9	ZZ001380a	88.74	1	34	ZZ002873a	64.18	5
10	ZZ003113a	85.24	6	35	ZZ001556a	62.65	1
11	ZZ003654a	83.77	8	36	ZZ002669a	62.64	5
12	ZZ004027a	83.2	9	37	ZZ001941a	62.61	2
13	ZZ002144a	81.67	3	38	ZZ004082a	61.29	9
14	ZZ001372a	79.61	1	39	ZZ003167a	60.98	6
15	ZZ003047a	76.57	6	40	ZZ003896a	60.81	9
16	ZZ003332a	75.03	7	41	XS332317b	60.31	2
17	ZZ002232a	74.57	3	42	ZZ003918a	60.04	9
18	ZZ001548a	74.22	1	43	ZZ003185a	59.38	6
19	ZZ002751a	74.06	5	44	ZZ002689a	59.16	5
20	ZZ001746a	73.76	2	45	ZZ001566a	59.15	1
21	ZZ001839a	73.4	2	46	ZZ001790a	58.96	2
22	ZZ003618a	72.96	8	47	ZZ002981a	58.73	6
23	ZZ003640a	72.42	8	48	ZZ001473a	58.17	1
24	ZZ003658a	71.66	8	49	ZZ001464a	57.71	1
25	ZZ004083a	70.02	9	50	ZZ002729a	57.31	5

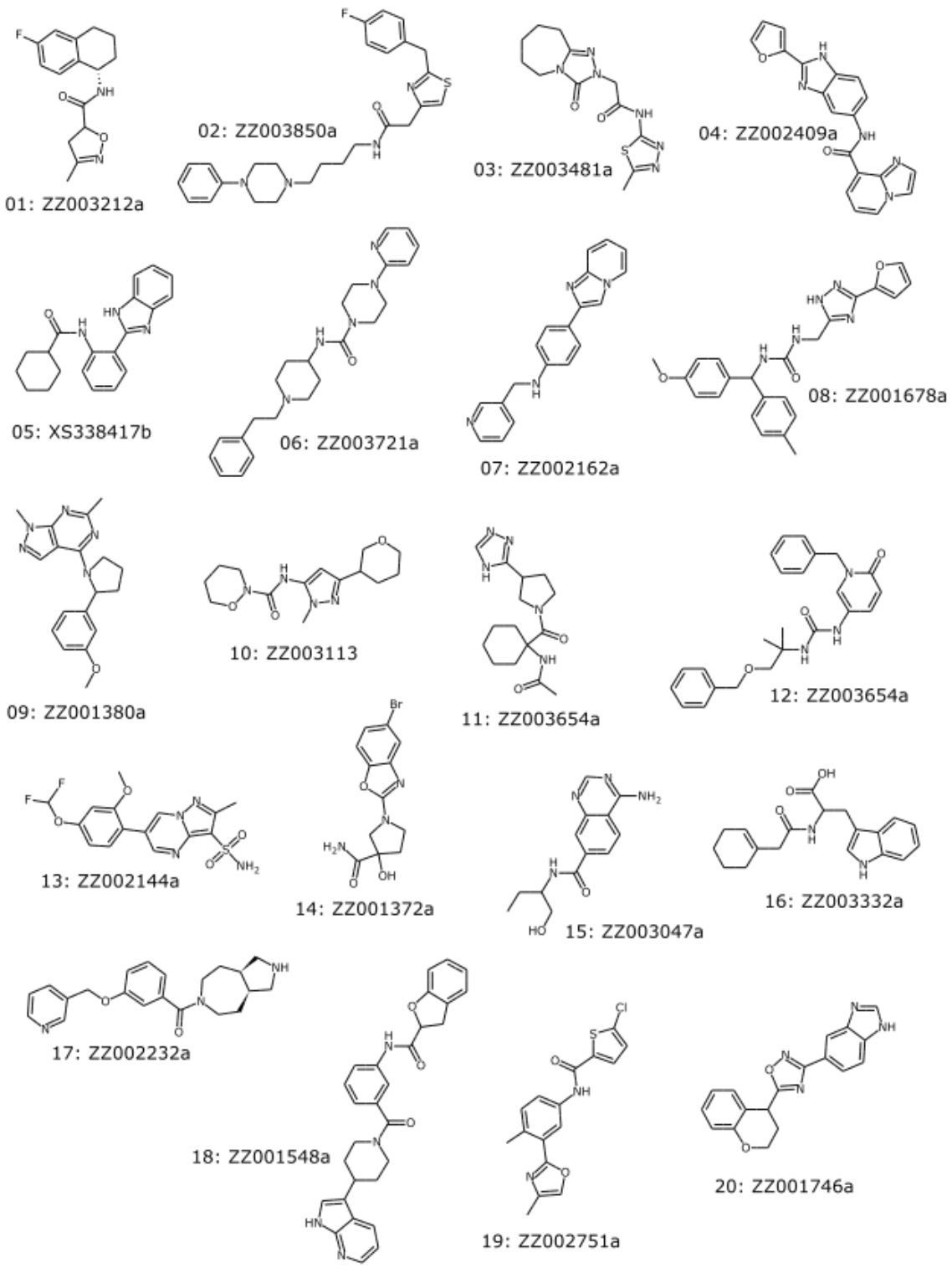


Figure 69: Structures of hit compounds 1 to 20, in order as per **Table 24**.

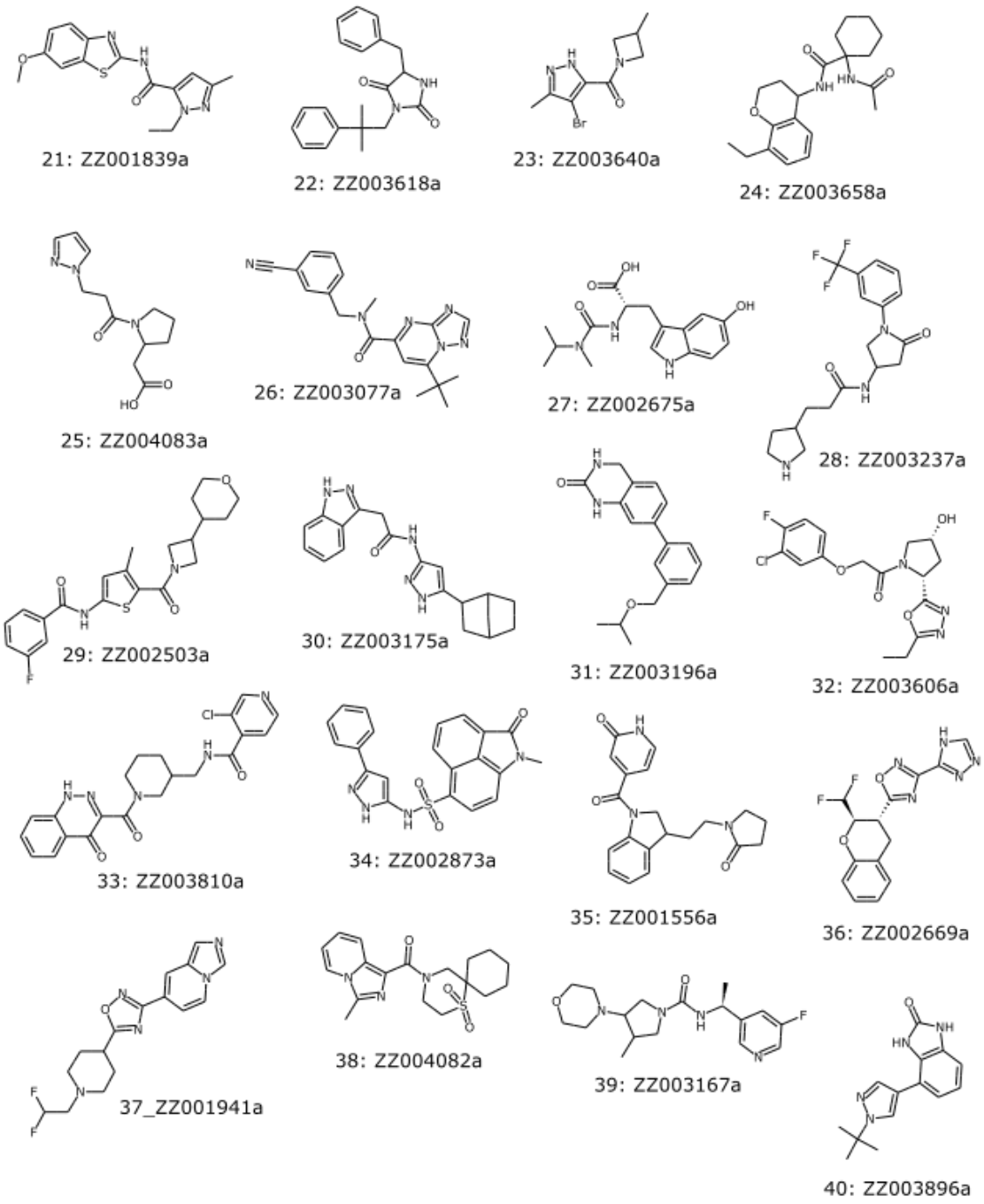


Figure 70: Structures of hit compounds 21 to 40, in order as per **Table 24**.

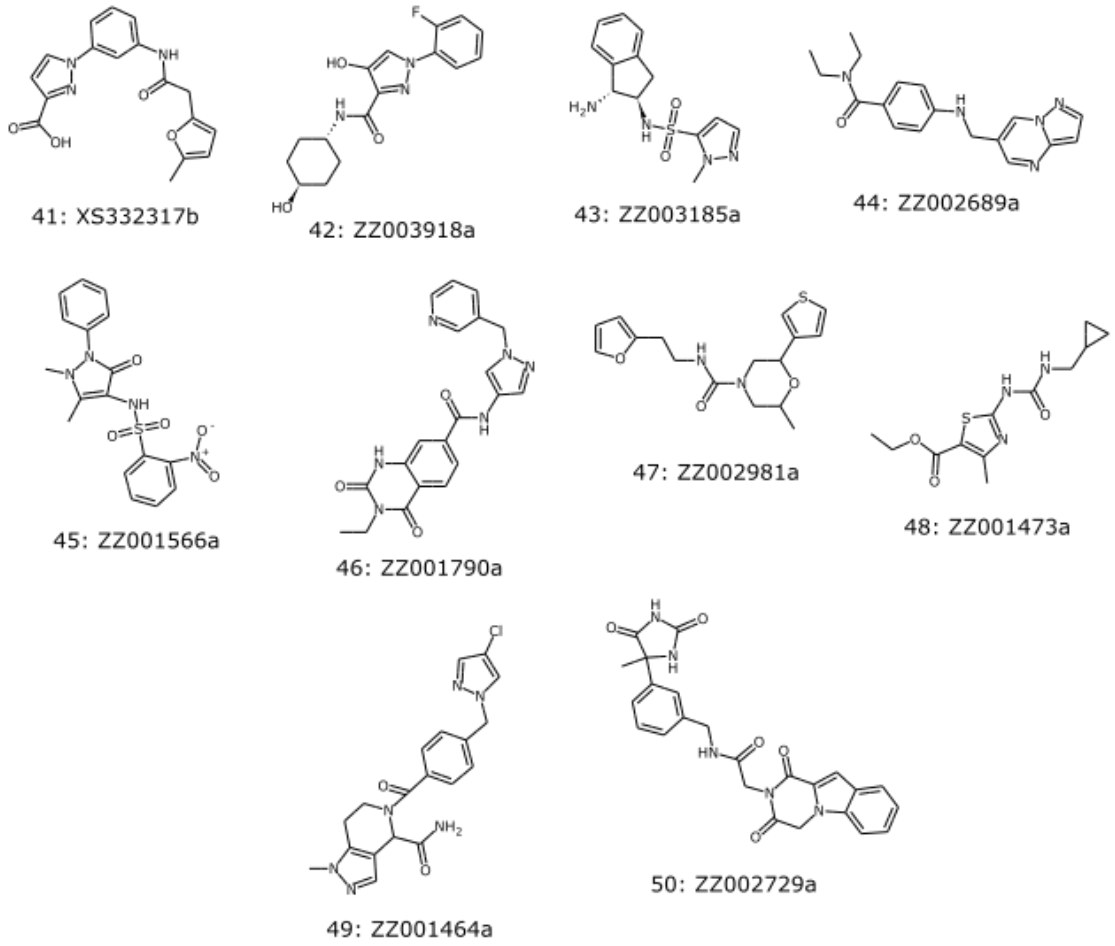


Figure 71: Structures of hit compounds 21 to 40, in order as per **Table 24**.

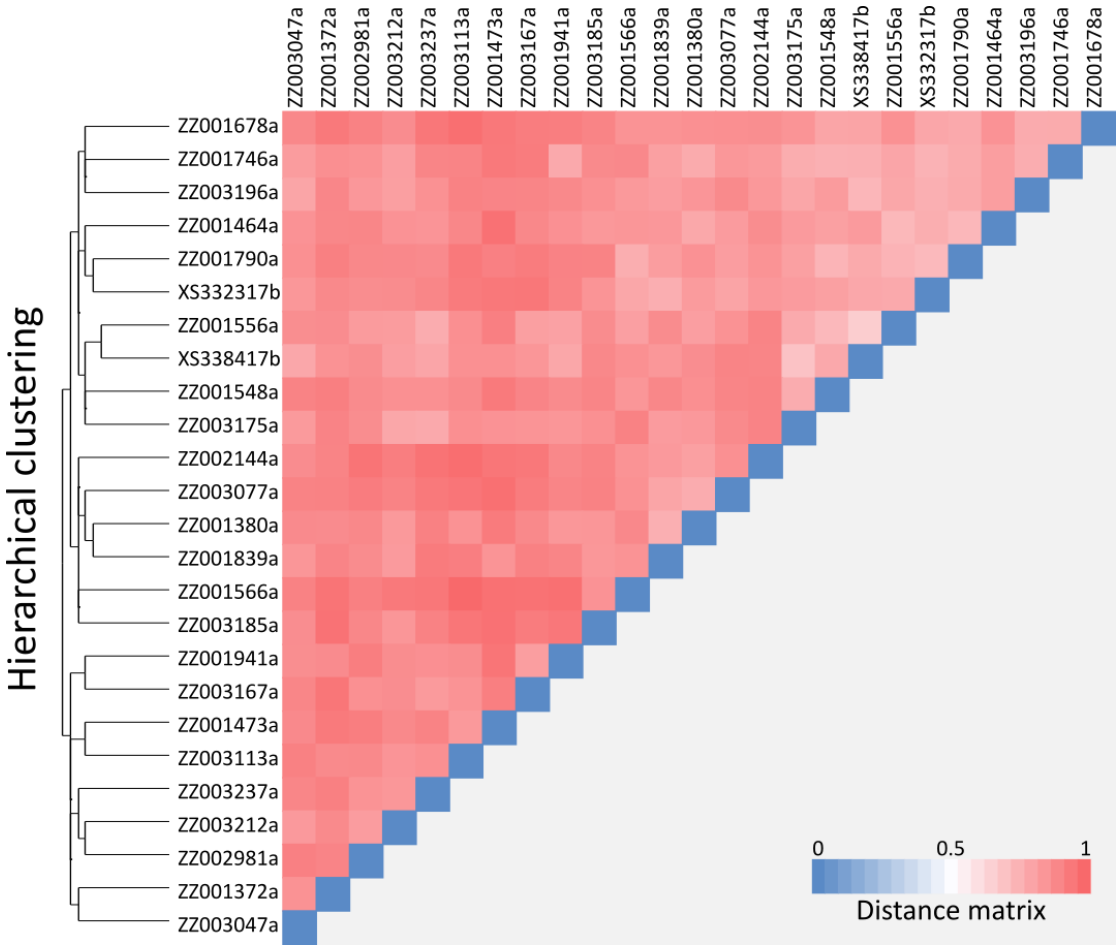


Figure 72: Hierarchical clustering and distance matrix for "hit" compounds with an inhibition ≥ 70%, as calculated with ChemMine Tools[164]. The branch length of the dendrogram is proportional to the similarity of the compounds.

3.5.4 Summary and Discussion

This serendipitously discovered MYLIP-auto-ubiquitination assay was then refined and used to conduct a pilot screen using a sub-set of the Oxford Drug Discovery Institute's 10k diversity library. Despite a blocked nozzle on a liquid dispenser, satisfying data quality (**Table 23**, **Figure 67**) was achieved across 9 assay plates after exclusion of affected wells (see **Figure 67C** for the affected area). A tenth assay plate was wholly discarded due to air bubbles affecting large areas of the plate. A range of compounds were identified as "hits" (70% or greater reduction of signal compared to no-compound addition reference). While no similarities could be identified using bioinformatical methods (**Figure 72**), validation of these hits could not be completed due to the absence of time. However, the lack of similarities and the low frequency of hits (25 out of 2640, or 0.95%) are encouraging.

Chapter 4 – Results of Software Development

4.1 Introduction

Biophysical and Biochemical Quantifier and Qualifier (BBQ) was developed in response to the first lockdown of the COVID-19 pandemic and becoming a father. Both these factors could pose major disruptions to laboratory work, and a developing a much-needed bespoke data analysis software for routine screening assays would i. insulate against such disruptions, ii. be of immediate benefit to the department, and iii. useful for the data generated in this work.

In the first instance, requirements were identified, and Python libraries selected to address these. Then, a user interface and a back end to process raw data and analyse it were created. Lastly, to future proof the software, development on a workflow editor has begun, which is still ongoing at the end of this work.

To disseminate the software in the future, Github and SourceForge repositories have been created: **https://github.com/tchristott/bbq_dissertation** and **https://sourceforge.net/projects/bbq-data-analysis/files/**, respectively. Due to file size limitations, the installation file for MS Windows is only available on the SourceForge repository.

4.2 Identification of requirements

4.2.1 All in one solution

The first and foremost requirement to be addressed is that all processing and analysis steps must be carried out in the same environment, without the need for different programs. The best way to achieve this for less technically minded end users is to build a front end/graphical user interface that takes the user through all the steps required to analyse the results of a particular experiment. Since most users are familiar with MS Office, the UI was designed with this in mind to make the user feel immediately feel familiar. In particular, the home screen of Office (2016) applications was used as a reference for the home screen of BBQ (**Figure 73**).

Out of the cross-platform user interface libraries assessed, wxPython was primarily chosen due to its licencing model, which allows for free use of the library while still allowing for the creation of commercial software based on it, without the need to pay licencing fees to the creators. Should the software developed here be turned into a commercial product or service,

this is invaluable. Another reason is the availability of a what-you-see-is-what-you-get user (WYSIWYG) interface editor (wxFormBuilder) and an extensive documentation.



Figure 73: Home screen of BBQ. Up to four assays can be pinned on the top half of home screen. To access the menu to pin assays, click on the pin symbol next to the “Pinned assays” heading. Up to ten recent projects are shown on the bottom half of the home screen. As in MS Office applications, this list is automatically updated as projects are saved and opened. SCARAB button to establish connection to the database. Help button opens the user manual. Version number, build date- and timestamp are displayed in the bottom left-hand corner.

4.2.2 Accessibility

While high contrast options and screen magnification are not easily implemented, colour blindness can easily be addressed by using a colour-blind safe colour scheme for all graphs. For this purpose, the “muted” colour scheme developed by Paul Tol[165] was chosen. After I discovered a slight red/green weakness myself, this colour scheme was also chosen for most figures in this work.

Table 25: Colour blind safe "muted" colour scheme by Paul Tol.

	Name	R	G	B	Colour
1	Indigo	51	34	135	
2	Blue	68	119	170	
3	Cyan	136	204	238	
4	Teal	68	180	153	
5	Green	17	119	51	
6	Olive	153	153	51	
7	Sand	201	204	119	
8	Rose	204	102	119	
9	Wine	136	34	85	
10	Purple	170	68	153	
11	Pale grey	221	221	221	

4.2.3 Automated and manual data processing

It must be possible to manipulate data, e.g., defining windows for fitting product formation rates in enzymatic assays or excluding data points from dose response curves (while keeping the data integrity intact by displaying the data points in a different colour or as empty markers).

4.2.3.1 Data organisation - pandas

pandas[166] (to be written in lower case) is used to write parsers for any and all liquid handler report and raw data files. The package has functions to read several file formats (read_csv, read_exel, read_json, and others). While most files are easily read CSV or excel files, in some instances it is necessary to try a few different readers and parameters before successfully accessing a file's contents.

Data imported from files is stored in so called data frames, which are essentially tables where each cell can hold any type of python object (strings, numbers (integers or floating-point numbers), lists, arrays, other data frames) and indexing can have multiple levels.

Pandas also allows for processing of the data held in data frames, e.g.

- filtering for certain values (e.g., to extract all entries in a transfer report associated with a sample ID)

- applying any function to the values in a column, row, or the entirety of the data frame, greatly simplifying data normalisation

4.2.3.2 NumPy

NumPy[167] is used for a broad range of mathematical operations, e.g. finding minimum and maximum values in a list that can contain NaN values that can trip up Python's built in minimum and maximum finding functions.

Also contains the gradient function to determine the gradient/first derivative of a continuously recorded dataset, e.g., from a DSF experiment, to determine the point with the highest gradient, i.e., the point of transition.

4.2.3.3 SciPy

SciPy[168] contains the "optimize" subpackage with the essential "curve_fit" function that is used to fit models to datapoints.

Also contains the "signal" sub-package with a Savitzky-Golay filter to smooth out continuously recorded raw data, e.g., from a DSF experiment.

4.2.3.4 PeakUtils

PeakUtils[169] is used to detect peaks and baseline in the first derivatives of DSF datasets.

4.2.3.5 Plotting data - matplotlib

To plot data, matplotlib[170] was used. This library allows for plotting data in many different plots and offers a vast range of customisation options. Additionally, it offers events that can be connected to functions in the main program, such as mouse clicks on the plot, hovering the mouse over datapoints, dragging the mouse over a plot, or picking datapoints. With these, interactive plots can be realised.

Implementations of interactive features are discussed in the corresponding workflow sections, but the two most common use cases are bringing up a context menu upon right clicking on the plot and including/excluding datapoints from a curve fit.

4.2.4 Database integration

The final requirement is the integration of the analysis software with the department's database to upload results with minimal user input and therefore minimal potential for user errors.

In our department, we use ICM-SCARAB[171], which is based on an Oracle database. To integrate BBQ with the database, Oracle's oracledb[172] Python library and instantclient (version 21.8) are used. The user is given the option to connect to the database on the main sidebar of BBQ or prompted to enter their credentials when clicking the "Upload to Database" button on the "Export to Database" tab. User credentials are currently not stored by BBQ.

In the ICM-SCARAB instance used by our department, some fields (e.g., sample and protein IDs) are not stored as such in assay results tables, but as by using the corresponding primary key from other tables where information about these reagents is stored as a reference. Prior to uploading the data, these primary keys are verified. If they cannot be found, the process is interrupted, and the user notified of which components have invalid IDs. These are also highlighted in yellow on the database table. Currently, the user will have to change the IDs in the liquid handler transfer file and/or the "Assay Details" tab of BBQ and repeat the analysis. Making the database table editable and having the edited information propagate through the data structure underlying the analysis workflow is part of future plans for BBQ. Once the data is verified, the data is uploaded via prepared SQL statements (**Code Excerpt 1**).

Code Excerpt 1: Database upload function for generic database upload table. The upload table is a property of the workflow calling the upload function (self.dfr_Upload). The SQL statement constructed with f strings is split into two lines for readability here. This is not the case in the actual code.

```
def db_upload(self, conn = None, db_table = None):
    """
    Inserts data in self.dfr_Upload.

    Arguments:
        conn -> database connection
        db_table -> name of database table

    Returns:
        shape[0] of pandas upload dataframe on success, 0 on failure
        or error code, if applicable.

    """

    if not conn is None:

        colnames = [str(col) for col in self.dfr_Upload.columns]
        str_colnames = ""
        str_valnames = ""
        for name in colnames:
            str_colnames += name + ", "
            str_valnames += ":" + name + ", "
```

```
        str_colnames = str_colnames[:-2] # removes last ", "
        str_valnames = str_valnames[:-2] # removes last ", "
        sql = f"INSERT INTO {db_table}({str_colnames})"
              + f"VALUES({str_valnames}) RETURNING PKEY INTO :pumpkin"

        pumpkin = conn.cursor().var(float,
                                    arraysize = self.dfr_Upload.shape[0]+1)
        colnames.append("pumpkin")
        sql_params = []
        for row in self.dfr_Upload.index:
            values = self.dfr_Upload.loc[row].tolist()
            values.append(pumpkin)
            dic = dict(zip(colnames, values))
            sql_params.append(dic)

        try:
            conn.cursor().executemany(sql, sql_params)
            conn.commit()
            return self.dfr_Upload.shape[0]
        except odb.Error as err:
            return str(err)
```

4.2.5 Future proofing

A major requirement is the need to address the creation of new workflows when equipment changes (e.g., replacement of a microplate reader with a newer model and therefore a change in raw data file format) or new assay technologies are added to the portfolio of a department. This would best be achieved by having a workflow editor with a graphical user interface and hints to allow even inexperienced users to quickly implement a new workflow.

For accessibility to other programmers, the code of BBQ follows the Python style guide PEP8 (**https://peps.python.org/pep-0008/**) with an exception made for line length.

4.3 Developed analysis workflows

There are several standard assays that are used in our department for a variety of targets (protein-protein interaction assays using AlphaScreen and HTRF, ADP-Glo, SYPRO-Orange based differential scanning fluorimetry). For these, specialised workflows were developed.

4.3.1 Common UI elements

To the top of the window, you will see the type of analysis being carried out. If the file has been previously saved, the file name will be shown below.

Below the file name (or placeholder text) is a row of buttons. You can either “Save” or “Save as” as with any other standard application. “Cancel” will cancel the analysis. “Analyse Data” will perform the analysis, provided all necessary information and data have been entered. “Previous” and “Next” will switch between the tabs of each analysis.

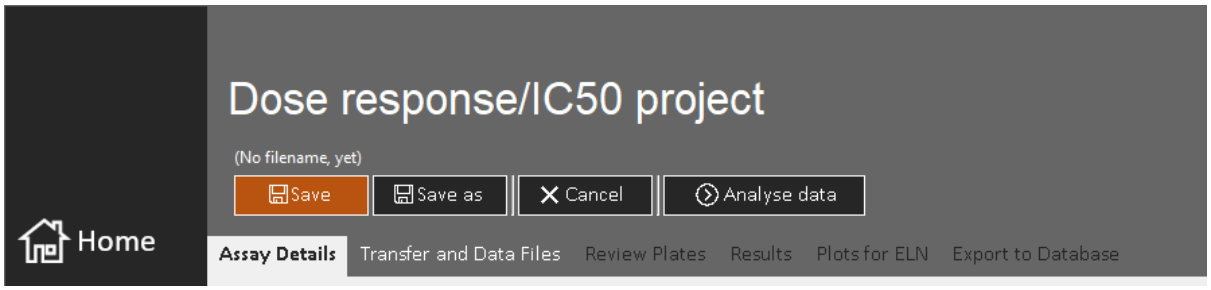


Figure 74: Example of the menu and tab structure of a workflow.

4.3.1.1 Assay Details

Here, the user can enter meta-data of the assay, such as names, IDs/batch numbers, and concentrations of assay reagents, buffer, solvent name, solvent concentration, ELN (electronic lab notebook) page reference, and date of Experiment.

4.3.1.2 Transfer and Data Files

This is applicable for microtiter plate-based assays only.

Here, the user can import a liquid handler transfer file, from which information about the assay plates is extracted. Raw data files can then be assigned to these entries for the actual analysis.

4.3.1.3 Review Plates

This is applicable for microtiter plate-based assays only.

To visually assess the quality of an assay plate, the raw readings are represented as a heatmap using matplotlib (**Figure 75A**). Mouse-over events were used to highlight connected areas on plates, i.e., all wells that belong to one sample and groups of control/reference wells. Tool tips are used to show the user information about the individual wells (e.g., well identifier, sample ID, well reading). The heatmap can be either copied to clipboard or exported a portable network graphics (PNG) file. Additionally, numerical measures to evaluate the assay quality are displayed (**Figure 75B**, section 4.3.2.2)

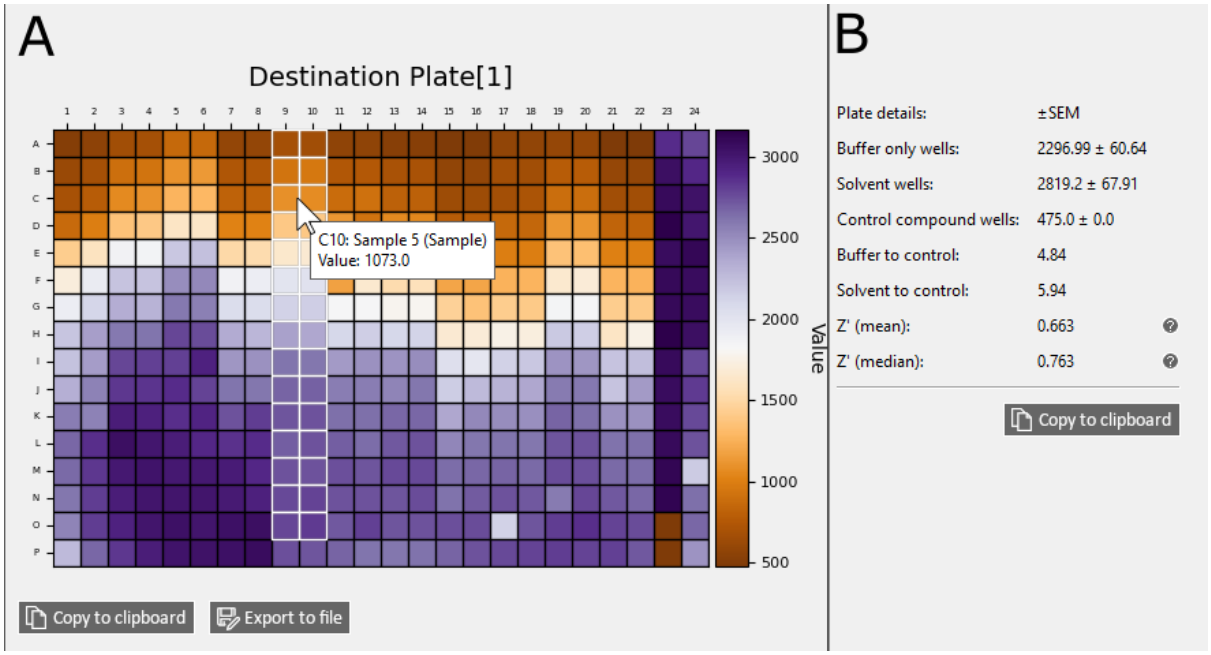


Figure 75: Assay plate review. **A:** Representation of an assay plate as a heat map. The heatmap can be copied to the clipboard or exported as a PNG file. **B:** Numerical assay quality measures. Values can be copied to clipboard. Not shown to the right: A list of all assay plates to select from for display.

4.3.1.4 Results

These tabs show a list of samples to the left, with relevant measures (e.g., IC_{50} value for does response assays, T_m and ΔT_m for DSF assays). Common with many applications, this list can be sorted by clicking on the column header and copied to the clipboard. To the right is a representation of the raw and analysed data in an interactive plot where the user can choose different display options (refer to individual assay workflow sections) as well as export the plots.

4.3.1.5 Plots for ELN

This is applicable for assays where there are individual datasets for each sample, e.g., dose response curves, fluorescence during thermal unfolding, etc.

This tab shows plots for each plate of an experiment to copy to the user's ELN page, if so desired. Samples are not plotted on the same axes but are shown as individual panels instead.

4.3.1.6 Export Results to Database

Results table for the SGC/CMD's SCARAB database, which can be either directly uploaded, exported as a file (MS Excel (.xlsx) or comma separated value (CSV)) or copied to clipboard.

4.3.2 Commonalities in behind-the-scenes data processing

In the following section, some ways in which data is handled in common ways across the different workflows are discussed.

4.3.2.1 Data fitting

Most assays have datasets that are either concentration, time, or temperature dependent (e.g. dose response curves). Equations that are used to describe these relationships can be fitted to the experimental data using the curve_fit function in the SciPy sub-package optimize.

The equations must be defined as Python functions and are passed as arguments to the curve_fit function alongside the raw data (as two separate lists/arrays, x_data and y_data). curve_fit then returns the calculated parameters of the function and their respective covariances, from which the 95% confidence interval can be calculated (see **Code Excerpt 2**, adapted from [173] under Creative Commons Attribution-ShareAlike 4.0 International License.)

Code Excerpt 2: Python function to calculate 95% confidence intervals for given parameters based on their covariance.

```
def calculate_confidence(n,pars,covar):
    """
    Calculates 95% confidence interval for parameters
    based on covariance values.

    Arguments:
        n -> integer. Number of datapoints
        pars -> list of floats. Values for parameters.
        covar -> list of floats. Covariances for each parameter
    """
    # http://kitchingroup.cheme.cmu.edu/blog/2013/02/12/Nonlinear-curve-
    # fitting-with-parameter-confidence-intervals/

    confidence = []
    p = len(pars)

    alpha = 0.05 # 95% confidence interval = 100*(1-alpha)

    # number of degrees of freedom
    dof = max(0, n - p)

    # student-t value for the dof and confidence level
    tval = t.ppf(1.0-alpha/2., dof)

    for i, p, flt_Variance in zip(range(n), pars, np.diag(covar)):
        if isinf(flt_Variance) == False:
            sigma = flt_Variance**0.5
            plusminus = sigma*tval
            confidence.append(plusminus)
        else:
```

```
            confidence.append(np.nan)
    return confidence
```

4.3.2.2 Data Normalisation for Endpoint Assays

In this work, any microtiter plate-based assay that produces one readout of a final state (e.g. equilibrium between bound and unbound state of a ligand to a protein, or the amount of product formed after stopping an enzymatic reaction) shall be referred to as an endpoint assay. These assays are used to screen potential inhibitors of the interaction/reaction and can be either single concentration screens of compound libraries or EC_{50} determinations for inhibitors being developed from initial hits. Such assays also contain references representing the highest and lowest possible signal, e.g., by dosing in a known inhibitor or leaving out a crucial component, or only adding solvent at the same concentration as it would be present in sample wells.

Data was normalised based on the assay type. In the case of AlphaScreen and HTRF assays, 100% inhibition would correspond to a complete *loss* in measured signal (or the mean signal level of wells with a condition known to fully prevent reaching of the endpoint, e.g., a known inhibitor or lack of a substrate) and this I would be calculated as follows:

$$I = 100 * \left(1 - \frac{M - \bar{C}}{\bar{R} - \bar{C}}\right)$$

Equation 6

Where $\bar{C}$ is the arithmetic mean of control measurements and $\bar{R}$ is the arithmetic mean of solvent reference measurements.

In the case of fluorescence polarisation assays, the readout is an *increase* in non-polarised fluorescence, and I would therefore be calculated as follows:

$$I = 100 * \left(\frac{M - \bar{C}}{\bar{R}}\right)$$

Equation 7

For all assays, the standard deviation SD was calculated as follows:

$$SD = \sqrt{\frac{\sum(I - \bar{I})}{n - 1}}$$

Equation 8

Where $\bar{I}$ is the mean inhibition and n is the number of replicates. From the standard deviation, the standard error of the mean was calculated:

$$SEM = \frac{SD}{\sqrt{n}}$$

Equation 9

4.3.2.3 Assay quality – Z' and Z'robust

In addition to the visual representation of the assay plate, a few numeric values are presented alongside the heatmap (**Figure 75B**).

i. Mean values and standard deviation of i. control compound (full inhibition) wells, ii. no sample/solvent reference wells (if used, to show effect of solvent on the assay), iii. solvent addition wells (no inhibition).

ii. Ratios of mean of buffer wells to mean of control wells and means of solvent addition wells to mean of control compound wells.

iii. To assess the quality of a particular assay plate, Z' and Z'_{robust} values were calculated using the control and reference values[148], [149]. The equations for these measures were implemented as functions to be called on in multiple assays.

$$Z' = 1 - \frac{3(\sigma_{pc} - \sigma_{nc})}{|\mu_{pc} - \mu_{nc}|}$$

Equation 10

Where σ_{pc}, σ_{nc} are the standard deviation of the positive and negative controls, respectively, and μ_{pc} and μ_{nc} are the mean of the positive and negative controls, respectively.

$$Z'_{robust} = 1 - \frac{3(M_{pc} - M_{nc})}{|m_{pc} - m_{nc}|}$$

Equation 11

Where M_{pc}, M_{nc} are the median absolute deviation (MAD) of the positive and negative controls, respectively, and m_{pc} and m_{nc} are the median of the positive and negative controls, respectively.

4.3.3 End Point Single Dose assays

For these assays, the raw data is simply normalised as described above (4.3.2.2) and the sample IDs are assigned to each data-point. Data was then displayed in scatter plots, both for individual plates and all results, and arranged in a results table suitable for upload to the department's database.

4.3.4 End Point Dose Response assays

Dose response curves were fitted to a sigmoidal dose response curve with four parameters (top, bottom, Hill slope (HS) and inflection point (IC_{50}). Data points with an SEM ≥ 20 were excluded from curve fitting. **Equation 12** was implemented as a simple function (**Code Excerpt 3**) to be passed to the scipy.optimize.curve_fit() function as an argument alongside the trimmed doses and response. Since this function takes the parameters of the equation as arguments, it was also used to calculate the idealised datapoints of the fitted curve for display on the results plots.

$$I(c) = I_{bottom} + \frac{I_{top} - I_{bottom}}{1 + \left(\frac{IC_{50}}{c}\right)^{HS}}$$

Equation 12

Code Excerpt 3: Function for sigmoidal dose response curve.

```
def eq_sigmoidal(x,ytop,ybot,h,i):
    """
    Function for sigmoidal dose response curve.

    Arguments:
        x -> float. concentration/dose
        ytop -> float. theoretical top of curve
        ybot: theoretical bottom of curve (e.g. 100% and 0% inhibition)
        h -> float. Hill slope
        i -> float. inflection point
    """
    return ybot + (ytop - ybot)/(1 + (i/x)**h)
```

The quality of the fit is then estimated via R^2, which was calculated as follows:

$$R^2 = \frac{\sum(I_{measured} - I_{calculated})^2}{\sum(I_{measured} - \bar{I})^2}$$

Equation 13

Where $I_{measured}$ is the per-cent inhibition at a given concentration calculated from the raw data using **Equation 6** or **Equation 7**, depending on assay type; $I_{calculated}$ is the calculated

inhibition at a given concentration calculated using the fitted parameters; and $\bar{I}$ is the mean (arithmetic) inhibition over all concentrations tested.

Dose response curves were plotted using the matplotlib library. Point picking events were used to include/exclude datapoints and to trigger re-fitting of the dose response curve. To ensure data integrity, any point removed from the curve is still displayed on the plot, but denoted by an empty circle, whereas included points are denoted by filled circles (**Figure 76**).

Code Excerpt 4: Function to calculate $R^{2.}$

```
def calculate_rsquare(data, fit):
    """
    Calculates R square value for given dataset and fit.

    Arguments:
        data -> list of numbers. Datapoints
        fit -> list of numbers. Fit to be assessed for quality.

    Returns
        Rsquare as float.

    R square is calculated as follows:

    Rsquare = 1 - ((residual sum of squares)/(total sum of squares))
            = 1 - ((Sum(Yi - Yi,fit)^2)/(Sum(Yi - Ymean)^2))

    Where Yi is the measured response at datapoint i, Yi,fit is the
    result of the fit at i and Ymean is the mean of all datapoints.
    """

    # remove any errant np.nan values from data
    data_trim = []
    fit_trim = []
    for i in range(len(fit)):
        if np.isnan(data[i]) == False:
            data_trim.append(data[i])
            fit_trim.append(fit[i])

    mean = sum(data_trim)/len(data_trim)
    RSS = 0 # residual sum of squares
    TSS = 0 # total sum of squares
    for i in range(len(data_trim)):
        RSS += ((data_trim[i]-fit_trim[i])**2)
        TSS += ((data_trim[i]-mean)**2)
    return round(float(1 - (RSS/TSS)),4)
```

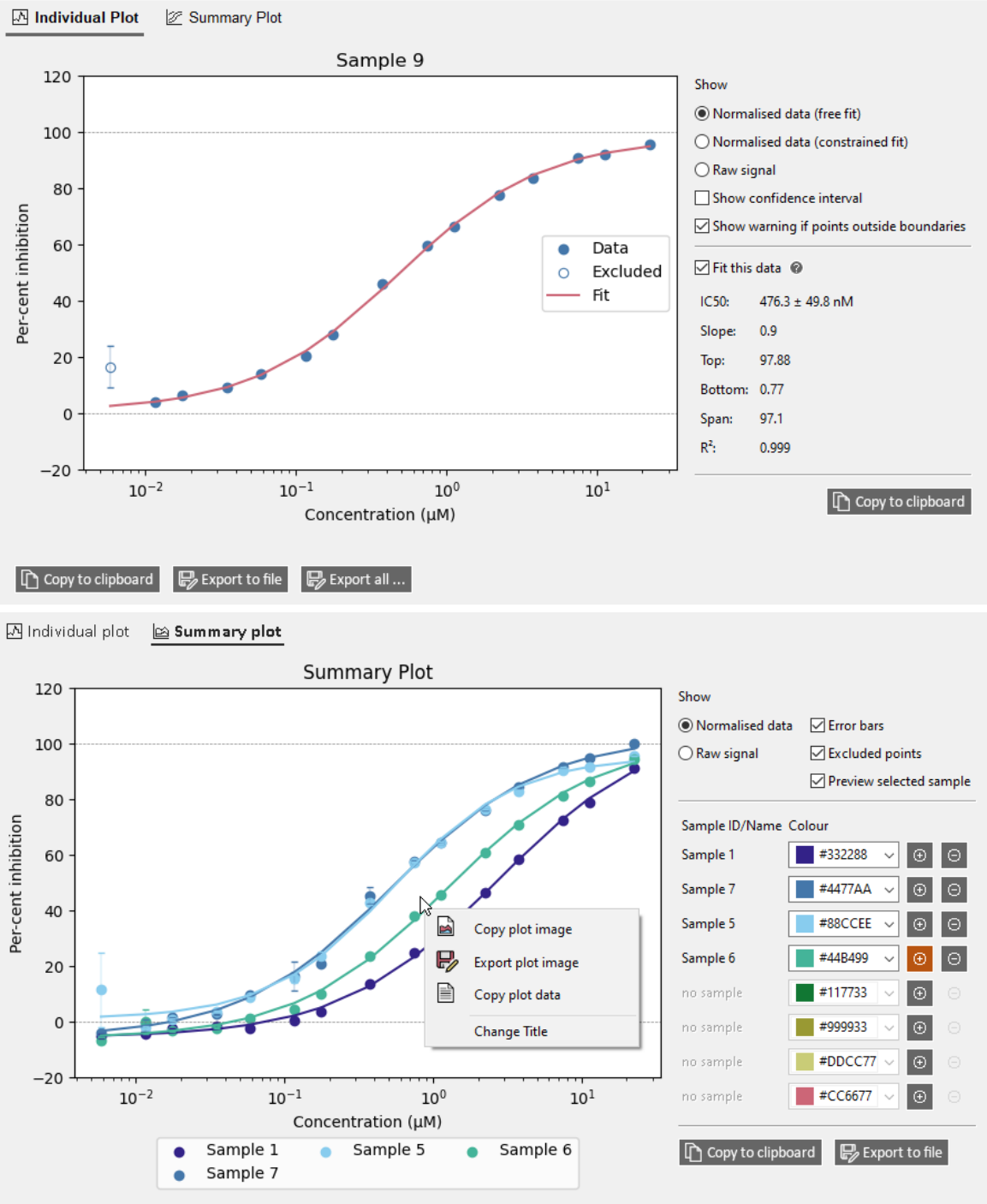


Figure 76: Screenshot of two results views of endpoint dose response analysis workflow. **Top panel:** Individual plot of a sample's dose response curve with one excluded datapoint. The user has the option to display the raw or normalised data as well as the confidence interval of the fit. **Bottom panel:** Summary graph of several dose response curves from one experiment. The grey curve is a preview of a sample selected from the list of all samples (not shown). The user has the options to display either the raw or normalised data, show error bars, and show excluded points.

4.3.5 Differential Scanning Fluorimetry

Differential scanning fluorimetry (DSF) is used to determine "melting point/temperature" (T_m) of proteins, that is the point at which the populations of folded and unfolded protein are equal. Many proteins can be stabilised by binding of a ligand, which shifts the transition point to higher temperatures. This shift is expressed as ΔT_m (**Equation 14**). Using dyes, such as SYPRO Orange, that become fluorescent upon binding to exposed hydrophobic amino acids and real-time polymerase chain reaction (RT-PCR) machines, the unfolding process can be measured. We have two models of machines in our department, a Roche LightCycler, which can read 96 well and 384 well plates, as well as two Agilent Mx3005P Real-Time PCR systems, which can read 96 well plates.

$$\Delta T_m = T_{m,ligand} - T_{m,reference}$$

Equation 14

Parsers were written to extract data from the raw data files from both device types and for all plate formats.

The user is given three options for sample IDs: Extract from liquid dispenser transfer report or entered manually (copy + paste from MS Excel). Alternatively, well designations (A01, etc.) can be used. In addition, multiple proteins can be processed on one plate. The layout of the plate can be defined using a graphical representation of the plate (**Figure 77**), and either be applied globally for all plates in the experiment/analysis or individually for each plate. At least one reference well per protein is required to calculate ΔT_m values. If no reference wells are defined, only T_m values are calculated.

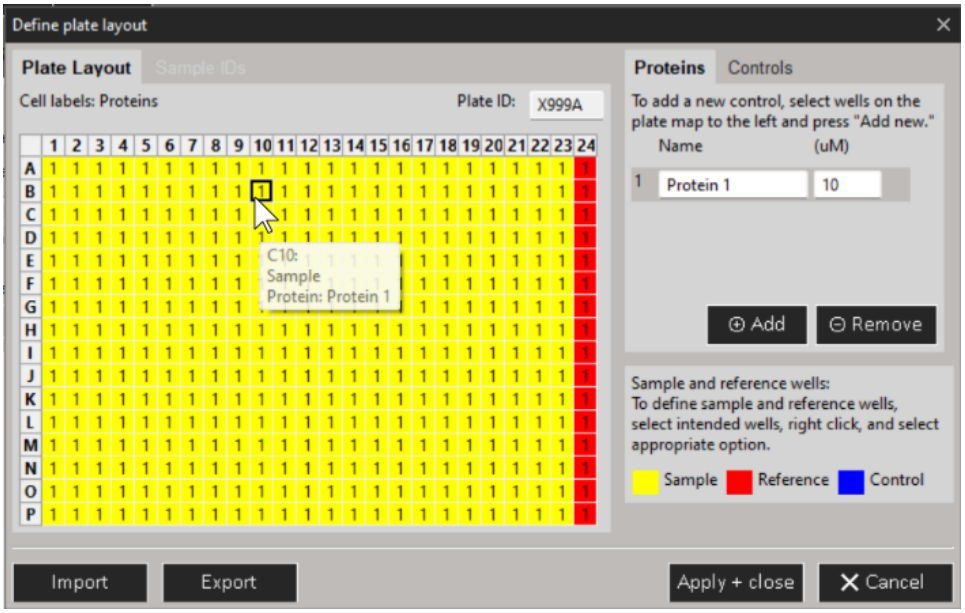


Figure 77: Plate layout definition menu. Wells can be modified by right clicking on them and bringing up the context menu.

In the first instance, the inflection points of the entire curve are calcuated. After applying a Savitzky-Golay filter from SciPy's signal sub-package, the gradient, or first derivative, of the curve was determined using NumPy's gradient function. All peaks that were at least five degrees C wide and greater than or equal to 5% of the maximum of the derivative were extracted using the find_peaks function from the signal sub-package. A simplified version of the code is shown in **Code Excerpt 5**. In the actual code, data is stored in pandas data frames.

Code Excerpt 5: Simplified excerpt from the data processing function for differential scanning fluorimetry for one sample. Full code in the lib_datafunctions.py file. y_data is the fluorescence intensity, x_data the temperature. np is numpy, scsi is scipy.signal, pd is pandas

```
derivative = np.gradient(scsi.savgol_filter(y_data,20,2))
deriv_baseline = pu.baseline(np.asarray(derivative), deg = 3)
max = np.nanmax(derivative)
peaks, props = scsi.find_peaks(derivative, width = 5, height = max*0.05)
inflections = [int(p) for p in peaks]
slopes = [derivative[x] for x in peaks]
peakstarts = x_data[int(x)] for x in props["left_ips"]]
# find inflection point at which derivative is greatest:
if len(inflections) > 1:
    deriv_arr = np.asarray(derivative)
    inf_deriv = deriv_arr[inflections]
    dic_inf = dict(zip(inf_deriv,inflections))
    inf_max = dic_inf[np.nanmax(list(dic_inf.keys()))]
```

```
    dic_slopes = dict(zip(inflections,slopes))
    slope = dic_slopes[inf_max]
elif len(inflections) > 0:
    inf = inflections[0]
    slope = slopes[0]
else:
    inf = np.nan
    slope = np.nan
inf_max = inf
slope_max = slope
# Set default Tm to inflection point with maximum heght
Tm = inf_max

transition = None
if not pd.isna(inf_max):
    for dp in range(inf_max,0,-1):
       if derivative[dp] < deriv_baseline[dp]:
           transition = processed.loc[k,fit][dp]
           break
    tmguess = processed.loc[k,"Temp"][inf_max]
else:
    tmguess = middle_of_list(x_data)
pars, confidence, stderr, success = fit_tm_boltzmann(x_data,
                                                     y_data,
                                                     tmguess,
                                                     transition)
if pars[0] > np.nanmax(x_data):
     pars = [np.nan] * 4
     drawn = [np.nan] * len(processed.loc[k,"Temp"])
else:
     drawn = draw_tm_boltzmann(x_data,pars)
```

Then, the data was fitted as described by Niesen et al[145] by fitting the Boltzmann equation (**Equation 15**) to the region around the transition between the entire protein population being folded and unfolded. Here, FI is the fluorescence intensity at a given temperature T, FI_{peak} and $FI_{minimum}$ are the fluorescence intensities at the peak and minimum of the curve, respectively, corresponding to fully unfolded and folded protein, respectively. T_m is the temperature at the point of transition between the two states and a is the slope of the curve at the transition point.

$$FI(T) = FI_{min} + \frac{\left(FI_{peak} - FI_{min}\right)}{\left(1 + e^{\left(\frac{T_m - T}{a}\right)}\right)}$$

Equation 15

To extract the transitional region over which the fit is applied, the maximum was determined, and the dataset truncated behind that. Next, the minimum point left of the maximum was determined and the dataset truncated before that. The Boltzmann equation (**Equation 15**) is

again implemented as a function (**Code Excerpt *6***). then fitted to the truncated data set. The inflection point calculated in **Code Excerpt *5***, alongside the raw data (temp, fluo) is passed to the fitting function (**Code Excerpt *7***) as an argument. Using a lambda function, the temperature at which the fluorescence intensity reaches maximum is fixed. Since the point of minimal fluorescence intensity could be at the beginning of the dataset, it is left variable.

Code Excerpt 6: Boltzmann equation implemented as a Python function.

```
def eq_boltzmann(T, Tm, LL, UL, a):
    """
    Describes transition between fully folded and fully unfolded protein
    in differential scanning fluorimetry.

    From Niesen, Berglund, Vedadi
    doi:10.1038/nprot.2007.321

    Arguments:
        T -> float. Temperature
        Tm -> float. Melting temperature, point of inflection.
        LL -> float. Lower limit
        UL -> float. Upper limit
        a -> slope
    """
    return LL + (UL - LL)/(1+np.exp((Tm - T)/a))
```

Code Excerpt 7: Fitting function for differential scanning fluorimetry.

```
def fit_tm_boltzmann(temp, fluo, tmguess, transition = None,
                     parsonly = False):
    """
    Calculates the Tm by fitting a Boltzmann equation to fluorescence
    curve.

    Arguments:

        temp -> list of floats. Temperature in degrees C
        fluorescence -> list of floats. Fluorescence intensities
        transition -> float. Deprecated. Was used to set boundaries
                      for fit.
        parsonly -> boolean. Whether to return parameters only.

    Returns if parsonly == true:
        pars -> list of calculated parameters of equation.
        success -> boolean. Whether curve was successfully fitted.

    Returns if parsonly == false:
        pars -> list of calculated parameters of equation.
        confidence -> list of tuples. 95% confidence intervals for
                      paramters
        stderr -> standard error of parameters
        success -> boolean. Whether curve was successfully fitted.
    """
    # Adjusting to Kelvins for fit
    tempK = [t + 237.15 for t in temp]
```

```
    # cut off values after maximum first IF maximum is not in first quarter
    # of datapoints
    max = np.nanmax(fluo[10:-10])
    fluo_trim = []
    temp_trim = []
    tempK_trim = []
    for dp in range(len(fluo)):
        if not fluo[dp] == max:
            fluo_trim.append(fluo[dp])
            temp_trim.append(temp[dp])
            tempK_trim.append(tempK[dp])
        else:
            fluo_trim.append(fluo[dp])
            temp_trim.append(temp[dp])
            tempK_trim.append(tempK[dp])
            break

    # Reverse lists to make cutting off before minimum easier
    fluo = fluo_trim.copy()
    fluo.reverse()
    temp = temp_trim.copy()
    temp.reverse()
    tempK = tempK_trim.copy()
    tempK.reverse()
    fluo_trim = []
    temp_trim = []
    tempK_trim = []

    # append values to _trim lists if they are not the minimum value
    # once the minimum has been reached, append it, then stop
    min = np.nanmin(fluo)
    for dp in range(len(fluo)):
        if not fluo[dp] == min:
            fluo_trim.append(fluo[dp])
            temp_trim.append(temp[dp])
            tempK_trim.append(tempK[dp])
        else:
            fluo_trim.append(fluo[dp])
            temp_trim.append(temp[dp])
            tempK_trim.append(tempK[dp])
            break

    # reverse lists back to correct orientation
    fluo_trim.reverse()
    temp_trim.reverse()
    tempK_trim.reverse()

    #
    fluo_coarse = []
    temp_coarse = []
    tempK_coarse = []

    for dp in range(0,len(fluo_trim),3):
        fluo_coarse.append(fluo_trim[dp])
        temp_coarse.append(temp_trim[dp])
        tempK_coarse.append(tempK_trim[dp])

    # More than 15 datapoints are necessary for good fit
    if len(fluo_coarse) < 15:
        pars = [np.nan] * 4
```

```
        confidence = [np.nan] * 4
        stderr = [np.nan] * 4
        success = False
        if parsonly == True:
            return pars, success
        else:
            return pars, confidence, stderr, success

    if transition is None:
        transition = min

    # prepare guesses for curve fitting
    guess = np.array([tmguess, min, 1])

    try:
        # Use lambda to define custom function with fixed upper limit:
        fix_boltzmann = lambda T, Tm, LL, a: eq_boltzmann(T, Tm, LL, max,
                                                          a)
        pars, covar = curve_fit(fix_boltzmann,
                                xdata = temp_coarse,
                                ydata = fluo_coarse,
                                p0 = guess)# ,bounds=bounds)
        pars = [pars[0],pars[1],max,pars[2]]
        confidence = calculate_confidence(len(temp_trim),pars,covar)
        stderr = np.sqrt(np.diagonal(covar))
        success = True
        if parsonly == True:
            return pars, success
        else:
            #return pars.tolist(), confidence, stderr.tolist(), success
            return pars, confidence, stderr, success
    except RuntimeError:
        pars = [np.nan] * 4
        confidence = [np.nan] * 4
        stderr = [np.nan] * 4
        success = False
        if parsonly == True:
            return pars, success
        else:
            return pars, confidence, stderr, success
```

In the results plot, the user has the option to i. display either raw or normalised fluorescence, ii. mark inflection points on the plot and if yes, iii. show all or only the main inflection point (i.e., with highest value for derivative), iv. plot the algorithmic fit (i.e., fitting **Equation 15** to the transition area) and v. mark the calculated T_m on the plot (**Figure 78**).

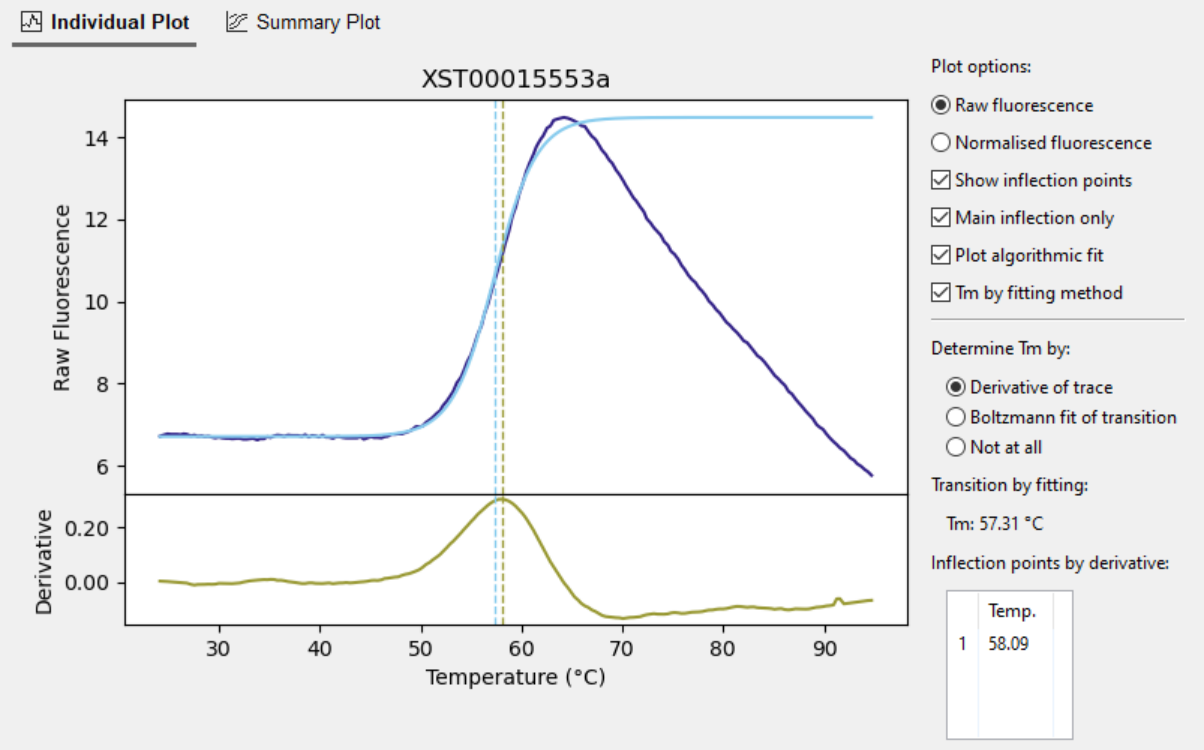


Figure 78: Results display for DSF analysis workflow. Not pictured to the right: List of samples to select from.

4.3.6 Cell Based Compound Screen

This analysis workflow was established in collaboration with Daniel Ebner's group at the Target Discovery Institute for cell based phenotypical screening of compound libraries. The experiments were designed and carried out by Dr Stephanie Hatch and Dr Timothy Johanssen. SH and TJ also defined the parameters of data analysis.

In a typical screen, a single cell line is used, and a compound library is screened against said cell line at four different concentrations (3 µM, 0.3 µM, 30 nM and 3 nM) in a combination of three conditions (Compound treatment, irradiation, and hypoxia) in two biological replicates, leading to a total of 64 plates per experiment. Each compound is evaluated against the reference condition (no compound treatment, no irradiation, normoxia).

4.3.6.1 Assay details

While the experimental parameters for the screens conducted by SH and TJ during this project are fixed, it was decided to add options for future flexibility.

In this tab, the user can define the name of the cell line, sample concentrations, the conditions and whether the conditions are independent or whether a combination of conditions is used, whether one condition is used as a reference condition, how many biological replicates are used, and what the plate layout looks like. The plate layout menu is essentially the same as for DSF assays, with the addition of control compounds. If more than one control compound is used, the user can select which one is used to calculate Z' and Z'_{robust} values.

4.3.6.2 Raw data

Raw data for these screens is coming from either of two image processing pipelines that extract the number of nuclei (stand in for living cell count) per well of 384 well cell-culture plates. Two different parsers had to be written to accommodate the Columbus Image Data Management and Analysis System (Perkin Elmer) and Harmony High Content Imaging and Analysis Software (also Perkin Elmer) pipelines. In Brief, there is a directory structure with directories for each condition (or cell line) analysed, with sub-directories for each plate of this cell line/condition. Result files are plain text files in these directories.

As with other assays, Z' and Z'_{robust} are calculated (**Equation 10**, **Equation 11**).

Since the replicates are biological replicates on different plates, replicate correlation was also calculated. First, values for replicate 1 were plotted against values for replicate 2, both expressed as per-cent of the solvent references. Then, a linear function was fitted to the data and the R^2 value calculated. Second, Pearson's correlation coefficient was calculated. These operations were combined in one function (**Code Excerpt 8**).

Code Excerpt 8: Function to calculate R^2 and Pearson's correlation coefficient for replicate datasets. The dropna(), sort_values(), and reset_index() operations on dfr_Temp are broken up for readability and carried out in one line in the actual code.

```
def calculate_repcorr(rep1, rep2):
    """
    Calculates replicate correlation for two given data sets.

    Arguments:
        rep1 -> list of floats. Replicate dataset 1
        rep2 -> list of floats. Replicate dataset 2

    Returns:
        m -> float. Slope of linear fit
        c -> float. Offset of linear fit
        Rsquare -> float. R square value of linear fit to
                   datasets (rep1 xdata, rep2 ydata)
```

```
        Pearson -> float. Pearson's correlation coefficient.
    """
    # Sort along "x" axis
    dfr_Temp = pd.DataFrame(data={"rep1":rep1,"rep2":rep2})
    dfr_Temp = dfr_Temp.dropna(axis = 0)
    dfr_Temp = dfr_Temp.sort_values(by=["rep1"],ascending=[True])
    dfr_Temp = dfr_Temp.reset_index(drop=True)

    # Perform linear fit
    pars, covar = curve_fit(eq_linear, dfr_Temp["rep1"], dfr_Temp["rep2"])
    # Calculate R square and Pearson's correlation coefficient
    Rsquare = calculate_rsquare(data = dfr_Temp["rep2"],
                                fit = draw_linear(dfr_Temp["rep1"], pars))
    pearson = round(dfr_Temp.corr().loc["rep1","rep2"],4)

    return pars[0], pars[1], Rsquare, pearson
```

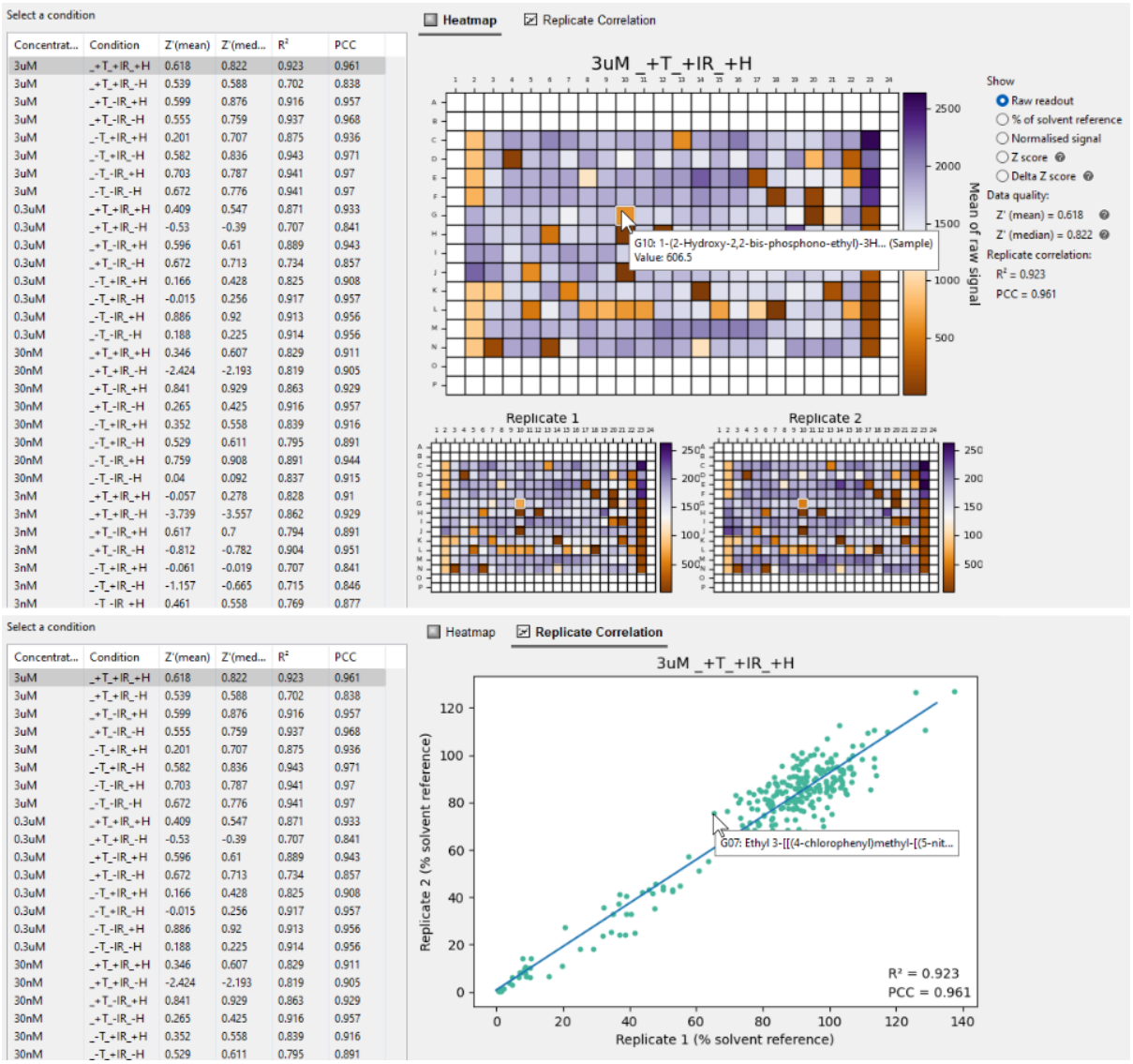


Figure 79: Review Plates tab of CBCS workflow. List of conditions on the left. **Top Panel:** Heat maps of mean values and replicates. With display options on the right. **Bottom Panel:** replicate

correlation plot on the right. Plots can be exported as a PNG file, copied to clipboard as a graphic, or copied to clipboard as table of datapoints via the context menu available on right clicking on it.

4.3.6.3 Data normalisation

Two normalisation techniques are applied per plate and the normalised data stored in the data structure.

A) Data is expressed as per-cent of the mean solvent control/reference reading.

B) Raw data is normalised by row and column. First, each well's reading in a row is divided by the median value of the row. Then, each well's median-normalised value is divided by the median of the column. Control and reference wells are excluded.

4.3.6.4 Data analysis and results

Data sets generated from B were then used to calculate Z score for each well (**Equation 16**):

$$Z = \frac{w - \mu}{\sigma}$$

Equation 16

Where w is the reading for the well, μ the mean of the sample wells on the plate and σ the standard deviation of sample wells on the plate.

For this project, data for each plate is compared to a reference condition. To this end, a ΔZ score was calculated:

$$\Delta Z = Z_{test} - Z_{reference}$$

Equation 17

The "Review Plates" tab for this assay has additional features: There are two smaller heatmaps under one big heatmap. The main map shows the mean value of either raw or processed data (there is a list of options to the left) and the two smaller heatmaps show the values for each of the two replicate plates.

To rank the compounds, the maximum absolute ΔZ across all conditions was determined. It is displayed in form of an interactive heat map and sample list. On the list, samples can be sorted by well, sample ID, or the maximum absolute ΔZ in ascending or descending order. Clicking on a well on the map or on the list will display the data on the plot. Hovering the mouse over

entries in the legend will increase the transparency of the other graphs on the plot. Leaving the legend will reset the transparency of all graphs. Again, right clicking on the plot brings up the context menu to either copy the plot to clipboard, export as PNG file, or copy the plot data to clipboard for plotting in another application. Additionally, the user can select multiple wells on the list and export all their corresponding graphs as PNG files.

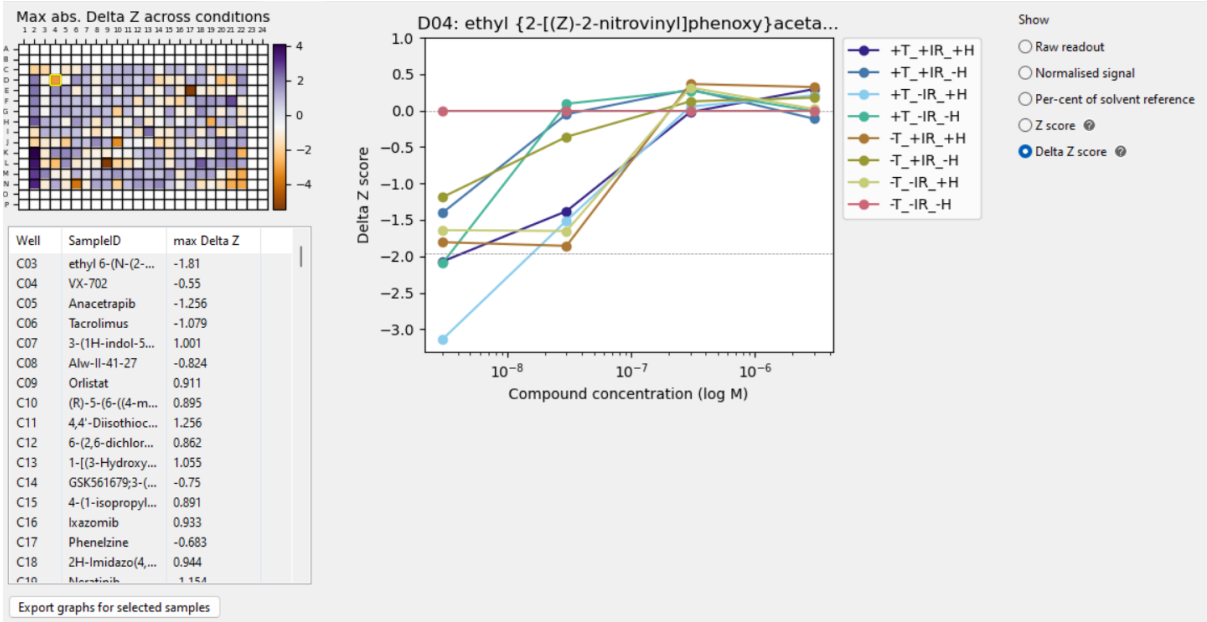


Figure 80: Results tab of Cell Based Compound Screen. Top left corner: Heat map of maximum ΔZ score. Bottom left: Sample List. Middle: Plot of all conditions for the selected well. Right: Plotting options.

4.3.7 Enzymatic Reaction Rate Dose Response assays

This workflow was developed for viral protease assays within the AI-driven Structure-enabled Antiviral Platform (ASAP) program (https://asapdiscovery.org/). The substrate for the protease is a short peptide that is flanked by a fluorophore (donor) and a quencher (acceptor). If the peptide is intact and the donor excited, energy transfer between the will quench the fluorescence coming from the donor. Here, a peptide flanked by the donor-acceptor pair Edans-Dabcyl is acting as the protease substrate[174]. The experiments were designed and conducted by Charline Giroud.

4.3.7.1 Raw data

As with the other assays, raw data files are assigned to entries from a liquid handler transfer report.

4.3.7.2 Data Processing and Analysis

The progression of product formation is approximated by a three parameter logarithmic function (**Equation 18**)[175] and the initial reaction rate v_i is then determined as the derivative at the origin, i.e., where $t = 0$ (**Equation 19**). Initial rates are then normalised to the maximum (no compound added, no inhibition) and minimum rate.

$$y = y_0 + b \ln\left(1 + \frac{t}{t_0}\right)$$

Equation 18

$$v_i = \frac{b}{t_0}$$

Equation 19

$$v_{i,norm} = 100 * \frac{(v_i - v_{i,min})}{(v_{i,max} - v_{i,min})}$$

Equation 20

After the data is processed and an initial fit has been performed, it is displayed on the "Results" tab with two plots next to the sample list (**Figure 81**). Empirically, a window from 0 s to 150 s has proven to produce workable results, and it is thus set as the default window. All reactions and the fitted model are shown on the major plot (**Figure 81**, left panel). The right-click context menu gives the user the option to zoom in and out of the plot, as well as to define the boundaries (Start and Stop) of the window over which **Equation 18** is fitted to the data. The left panel shows v_i plotted against the sample concentrations and a sigmoidal dose response curve fitted to the data (**Equation 12**). The plot is automatically updated when the user changes the fitting window. As with the End-point Dose Response workflow, points can be excluded from the fit, but remain on the plot. All plots can be copied/exported via the right click context menu.

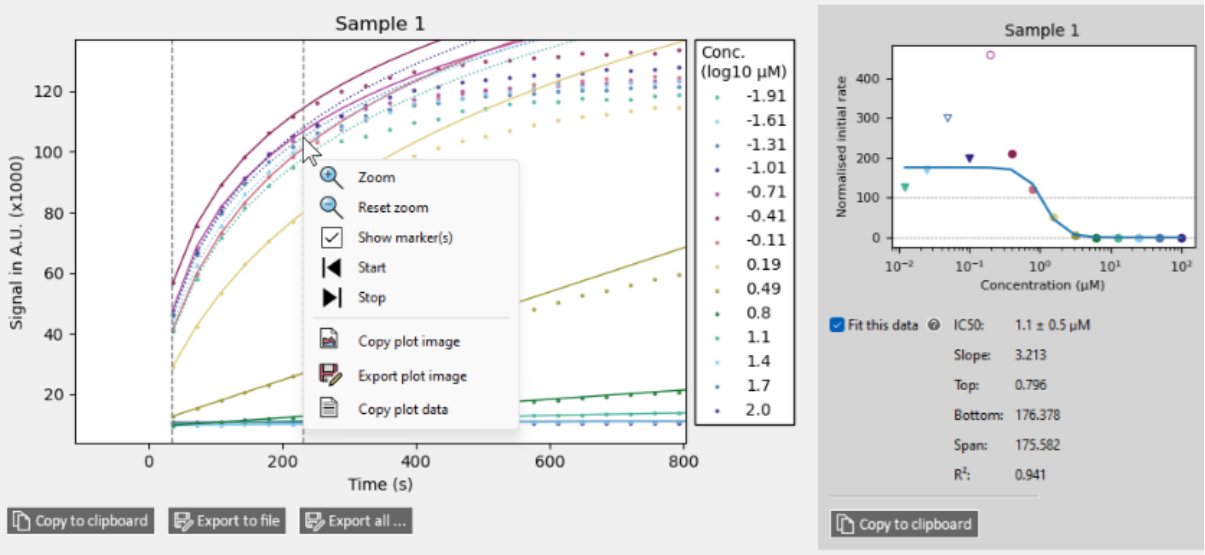


Figure 81: Results tab for the Enzymatic Reaction rate Dose Response workflow. **Left panel** shows the plot of the reactions as dots with the fitted model as continuous line. **Right panel** shows the dose response curve for the initial reaction rates.

4.4 Saving completed analyses for later use

Any data analysis software needs a file format to store completed analyses for later use (re-inspection, re-upload to the database, preparation of figures for meetings and publications).

For BBQ, analyses are stored as a series of CSV files inside a .zip archive with the custom .bbq file extension (**Figure 82**). From these files, the underlying data structure is rebuilt and, if they were populated at the time of saving, tabs will be repopulated.

The top level files in the archive hold general information about the assay (**Table 26**) while the sub-directories named after the individual assay plates' entries in the liquid handler transfer report hold the information pertaining to the contents of each plate (**Table 27** - **Table 29**).

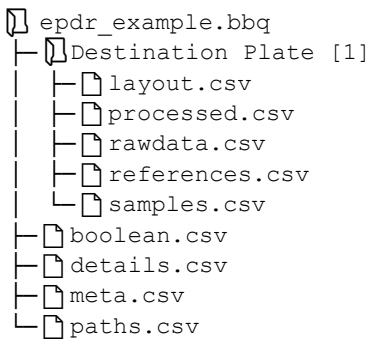


Figure 82: BBQ saved analysis file organisation.

Table 26: Top level files in .bbq saved analysis archive.

File	Purpose
boolean.csv	Variables that document the state of the analysis. E.g., have files been assigned, has the analysis been completed, have certain tabs been populated with results, have plots been draw, has the export to database table been populated.
details.csv	Assay details, such as reagent names and concentrations, assay type, devices, ELN page, date of experiment, etc.
meta.csv	Holds information about plates used: Plate name (e.g. “Destination Plate [1]”), how many wells, the assigned raw data file name
paths.csv	Stores the path of the liquid handler transfer file and the path where the raw data files are stored. Used for repopulating the

Table 27: Files in assay plate folder in .bbq archive

File	Purpose
layout.csv	Plate layout in table form. Wells are index column, properties of well are columns (see **Table 28**)
processed.csv	Holds the dataframe with the processed data, organised by sample
rawdata.csv	Contains the raw data

Table 27: Files in assay plate folder in .bbq archive

File	Purpose
references.csv	Holds values for mean, median, standard deviation and mean absolute deviation for all references and controls. Also holds calculated Z' and Z'_{robust} values.
samples.csv	Holds the locations, corresponding assay concentration, source concentration and transfer volumes (if applicable) of each sample on the plate.

Table 28: Contents of layout.csv in individual plate directory in .bbq archive.

Column	Purpose
WellType	r: Reference, s: Sample, c: Control, na (or empty cell): Not assigned
ProteinNumerical	Numerical identifier for protein, if used (0-n)
ProteinID	Batch/Lot ID/number
ProteinConcentration	Protein concentration, if used
ControlNumerical	Numerical identifier for control compound, if used (0-n)
ControlID	Batch/Lot ID/number
ControlConcentration	Control compound concentration, if used
ZPrime	If reference or control well, include this in Z' and Z'_{robust} calculations.
ReferenceNumerical	Numerical identifier for reference, if used (0-n)
ReferenceID	Batch/Lot ID/number or simple identifier (e.g., "DMSO", "Buffer"
ReferenceConcentration	Reference compound concentration, if used
SampleNumerical	Numerical identifier for sample, if used (0-n)
SampleID	Batch/Lot ID/number or unique identifier
SampleConcentration	Sample concentration, if used

Table 29: Columns in "samples.csv" file in individual plate directory in .bbq archive.

Column	Purpose
Destination	Name of destination plate
SampleID	Sample ID
Locations	List of locations on plate (indexed from 0-383). Lists in the following columns correspond to this list.
Concentrations	List of final assay concentrations as calculated from transfer volumes, source concentrations and assay volume.
TransferVolumes	List of transfer volumes.
Wells	Locations in plain text well identifiers (A01 – P384)
SourceConcentration	Concentration on source plate from which sample was dispensed.

4.5 Workflow editor

To enable faster development of workflows without the help of a Python developer, a flexible workflow editor is crucial. This editor needs to fulfil several requirements:

i. Easy to use with graphical user interface.
ii. Must be able to create parsing rules for liquid handlers reports and raw data files from plate readers.
iii. Offer simple data processing and normalisation options.
iv. Have an editor for mathematical equations for data fitting, ideally with pre-sets.
v. Options for data representation on the “Results” tab.
vi. Generate results tables and make database integration easy.
vii. Store the resulting rule set in an easily accessible file format.

4.5.1 Assay definition file format

Before the workflow editor was conceived, certain aspects of the pre-developed workflows were already harmonised to eliminate duplication of code. For example, a set of “widgets” was created for the assay details tab, so that the input fields for the assay details could be dynamically generated. Likewise, the column names for the results table and database table

were stored in the prototype of the assay definition files so that fewer functions were necessary to update the data. This concept was expanded to include rule sets to parse files, normalise data, fit models to the data and generate results tables.

To store these rule sets, the JSON (JavaScript Object Notification) file format was selected. The format is very similar to nested dictionaries in Python, and thus the standard Python module json (JSON encoder and decoder) can be used. An example of the syntax of JSON is given in **Code Excerpt 9**).

Code Excerpt 9: Example of JSON syntax. Shown is the first section of an assay definition file.

```
{
    "Meta": {
        "Shorthand": "EPDR",
        "FullName": "End Point Dose Response",
        "DisplayTitle": "Dose response/IC50 project",
        "MainCategory":"Plate Based",
        "SecondaryCategory":"Protein Protein Interaction",
        "OtherCategories":[
            "Dose Response"
        ]
    }
}
```

4.5.2 Workflow editor steps

For the user, the actual workflow editor is organised in a series of tabs, just like the actual workflow will be. Each tab corresponds to either a tab on the workflow's UI or a section on the back end (e.g., data processing, database integration). Each of these tabs is organised like a "wizard", i.e., the user is guided through the process step by step, and each step is verified before the user can progress to the next step.

4.5.2.1 Assay details and reagents

In the first section, the user can define: the name and other meta data about the workflow; the source of sample IDs (liquid handler transfer report, manual entry, or included in raw data file); how BBQ generates plate layouts internally; which reagents' details are recorded (**Figure 83**).

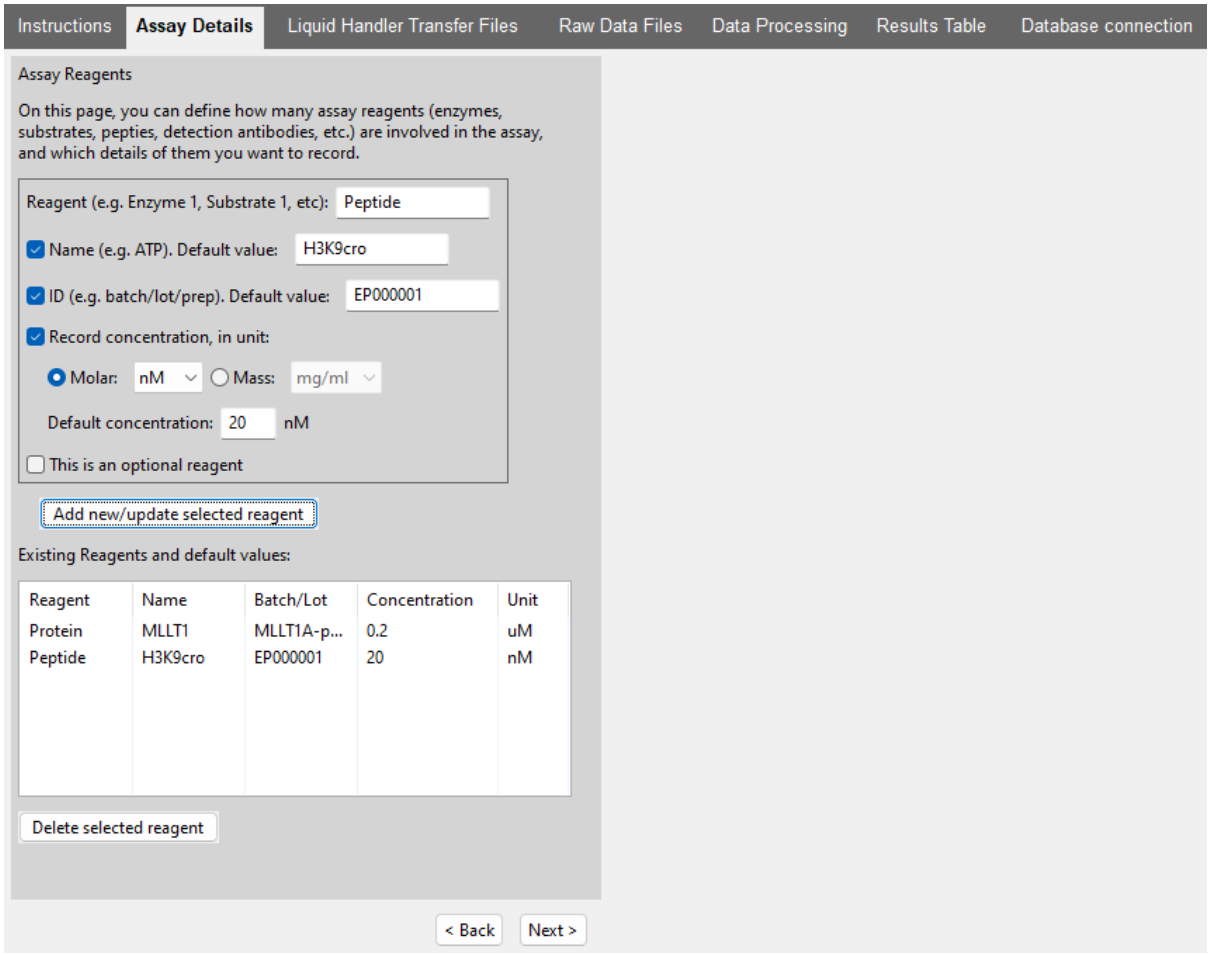


Figure 83: Reagent definition screen of the workflow editor.

4.5.2.2 Liquid handler transfer report parsing ruleset

The first step for the user is to select an example transfer file. If it is an excel file with multiple worksheets, the worksheet to extract information from can be selected. Once the file has been parsed, its contents are displayed, and the user can define the ruleset used to extract information from the file:

i. File verification: Choose a keyword in the file by right clicking on the cell on the preview field and selecting “File verification keyword” from the context menu. The user can also define whether the exact keyword is to be searched for or whether a partial match is allowed. In this case, the program will test each cell in the specified row or column for whether the defined keyword part of the text in the cell.

ii. Catch solvent-only transfers. Select this option if you want to automatically extract the location of wells that have had no compound dispensed but still backfilled with solvent.

iii. Find start of entries either by a keyword (exact or partial match, as above) or absolute coordinates. The defined keyword or coordinates must denote the row in which the column headers for the transfer entries are located.

iv. Find end of transfer entries either by keyword (as above), absolute coordinates, or an empty row.

v. Mapping of columns: The columns extracted from the example file must then be mapped to columns in the BBQ data structure necessary to process the data into a usable form (**Table 30**).

Table 30: BBQ data structure columns related to sample transfer.

BBQ Column Title	Comment
Source Plate Name	Either of these two is required for assigning raw data files to transfer file entries
Source Plate Barcode	
Source Concentration	
Source Concentration Unit	
Destination Plate Name	Either of these two is required for assigning raw data files to transfer file entries
Destination Plate Barcode	
Destination Plate Type	
Sample ID	Either of these two is required for proper sample identification
Sample Name	
Sample Transfer Volume	Optional. If given, can be used to calculate actual sample concentration in case of incomplete sample transfer.
Destination Well	
Destination Concentration	
Destination Concentr. Unit	
Solvent Transfer Volume	Optional

4.5.2.3 Raw data file parsing rule set

Like for the transfer file rule sets, the first step for the user is to select an example file. If it is an excel file with multiple worksheets, the worksheet to extract information from can be

selected. Once the file has been parsed, its contents are displayed, and the user can define the ruleset used to extract information from the file:

i. File verification: Choose a keyword in the file by right clicking on the cell on the preview field and selecting “File verification keyword” from the context menu. The user can also define whether the exact keyword is to be searched for or whether a partial match is allowed. In this case, the program will test each cell in the specified row or column for whether the defined keyword part of the text in the cell.

ii. What does the data represent: Datapoints corresponding to wells on a microtiter plate (and how many wells per plate) or a sample/collection of samples.

iii. How is the data organised: Either a in a grid pattern representing a microtiter plate or a table with two or more columns (first column being the well, following columns different readouts)

vi. Find start of data entries either by a keyword (exact or partial match, as above) or absolute coordinates. The defined keyword or coordinates must denote the row in which the column headers for the transfer entries are located. An offset must also be defined in case row and column labels are included in a grid representation.

iv. Are there multiple data sets per file? If yes,
   - define number of data sets (fixed number or determine dynamically)
   - whether they are organised vertically or horizontally
   - and how are datasets separated (same as main dataset, unique keyword and offset, fixed offset from preceding dataset, empty row).

v. Sub-datasets, e.g., multiple readings at different wavelengths for HTRF. If so, the same parameters and options as above are available.

vi. Find end of data entries either by keyword (as above), absolute coordinates, or an empty row.

4.5.2.4 Raw data processing

In this section, the operations applied to the raw data are defined.

First, data normalisation criteria are set out:

i. Subtract background? If yes, define what to use for the background signal: Control compound, solvent-only wells, or regents-only reference wells. Background-backup: If

no-addition/solvent is selected but not detected, use other as replacement, if it can be detected.

ii. Normalise data to reference, either as simple ratio (0 – 1) or as per-cent (0 – 100)

iii. Invert normalised data if your raw data corresponds to e.g., product formation but you want to express it as per-cent inhibition compared to a no-inhibitor reference (**Figure 84**).

iv. Where are replicates: On same plate or across plates.

v. Calculate errors: Either Standard Deviation (**Equation 8**) or Standard Error of the Mean (**Equation 9**).

vi. Data fitting - An editor to define equation to fit to the processed data. First, the independent variable must be defined. Second, the parameters of the function must be defined. Once all variables are defined, the user can enter the equation in a simple notation.

The interpreter behind the function is based on the exec function of python. Since this function can execute any text that is passed to it as an argument as if it were Python code, it is important to sanitise the input, that is make sure that no malicious code could be injected. To achieve this, the input in the text box is checked with every key stroke and certain criteria are enforced across this wizard page:

- Mathematical operators will always be separated from other characters.
- Periods, which could be used to call functions are only allowed between numbers.
- Parameters have a maximum character length of 10 alphanumeric characters.
- Paste events are also monitored to ensure that typing restrictions are not circumvented that way.

Additionally, a primitive linter has been constructed formats the equation dynamically. It enforces spacing, applies colour formatting to operators, parameters, numerical constants, prevents use of periods not between numbers, and underlines undefined parameters.

vii. Once the equation has been defined, it can be verified. That is, it will be checked for matched parentheses, that all parameters have been properly defined and that there are no operators that don't have the correct number of arguments, e.g., a plus sign at the beginning or end of the equation.

viii. Once the equation has been verified, a test data set can be entered, and the equation can be fit to it. The results are displayed to allow the user to verify the accuracy of the fit. On the display, the user can change from a linear x axis to a logarithmic x axis.

ix.

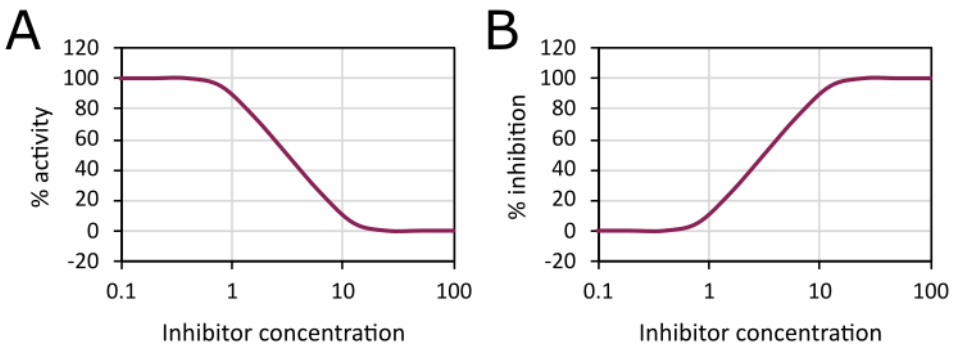


Figure 84: Example of normalised dose response assay data. **A:** Expressed as per-cent activity **B:** Expressed as per-cent inhibition.

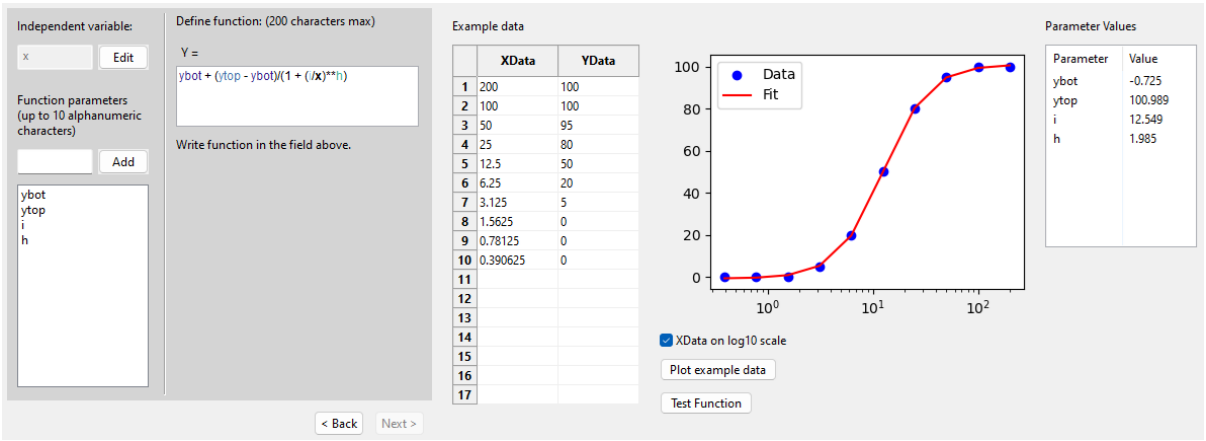


Figure 85: Equation editor page of the workflow editor. Right: Equation definition area. Middle two: Sample data and sample data plot Right: Calculated values of parameters.

4.5.2.5 Results Tab

Here the user can choose how to display the data. In general, a three-panel structure is presented:

i. Sample table with selected parameters from the curve fit.

ii. A plot of the raw and fitted data. If the dataset is made up of discrete datapoints, the user can choose to exclude datapoints from the fit.

iii. Panel with the parameters of the fit.

4.5.2.6 Results Table

In this tab, the user can define the results table that can later be uploaded to the database: All assay details, fitting parameters, sample concentrations, means of datapoints, errors, and plate quality measures can be assigned to columns.

4.5.2.7 Database integration

The user can define the name of the database and database table and assign columns of the results table to columns from the database table, e.g., “Protein Concentration [µM]” to “PROTCONC”. BBQ is currently limited to Oracle Databases, integration with other databases using structured query language (SQL) is planned.

4.5.3 Workflow interpreter

While the custom-built workflows have a lot of bespoke code, the workflows built with the editor need a common interpreter that assembles user interface elements and performs different tasks based on the parameters within the assay definition files. Following simple if-then-else branching, many functions of the existing code base can easily be copied and folded into this interpreter. At the time of submission, only the functions for parsing transfer and raw data files have been fully implemented.

4.6 Dissemination

Initially, the setup file was deposited into a directory on our department’s share drive. For ease of use, it was then decided to start a Microsoft Teams channel to alert current users of updates, provide change logs and direct download links via the institutional OneDrive. For public dissemination, a GitHub repository was created (**https://github.com/tchristott/bbq_dissertation**). However, GitHub imposes file size limitations, so it was not possible to deposit the windows installer. Now, the source code and a windows installer for the stand-alone version are deposited at **https://sourceforge.net/projects/bbq-data-analysis/files/.** Furthermore, a manuscript is in preparation for submission to an appropriate journal.

4.7 Summary and Discussion

Having identified the need for a simplified data analysis workflow for our department, I set out to identify which requirements need to be fulfilled to effectively address this need and how to best implement them. The Python programming language was chosen due to its rich

ecosystem of open-source libraries, cross-platform useability, and the fact that memory management is handled by the back end of Python, which simplifies software development for beginners.

The most important requirement was a user interface (UI) that guides the user through all steps of the analysis. While software packages like PRISM offer the user great flexibility for normalisation of data and fitting of models to the data, there is still manual preparation required before the data can be entered, such as transposing raw data into a usable format and associating datasets with samples. On the other hand, software packages like R allow automated readout of raw data files and plate layouts but require the user to run scripts from the command line or an IDE. These same restrictions often apply to python. Using the wxPython library, a user interface was constructed to circumvent these limitations. While each of the workflows developed for BBQ offer less flexibility, they provide productivity gains by being specific to each assay and guiding the user through the analysis process in a visual manner.

The part of the analysis process invisible to the user requires functions to read the files specified by the user and perform operations on the data (normalisation, fitting of models to data, association of sample IDs to datasets). The pandas library handles the parsing of data files and general organisation of data, while NumPy, SciPy and PeakUtils were used to process experimental data and perform curve fitting operations. To connect BBQ to the department's Oracle database and upload results, the built-in SQL functionalities of pandas as well as libraries supplied by Oracle were used.

Several workflows for routine assays were realised for different assay technologies (plate-based protein-protein-interaction assays, differential scanning fluorimetry, cell viability screens, enzymatic assays) to support both our group and other groups in the department. With the HTRF dose response workflow (**4.3.4**), the total processing time was reduced from tens of minutes per plate to a few minutes; with the Cell Based Compound Screen (**4.3.6**), analysis time was reduced from hours for a 64 plate data set to mere minutes.

During the development of the software, an additional requirement was identified: To future-proof BBQ, a way to create new analysis workflows without the need for in depth knowledge

of Python (or other programming languages) and databases is highly desirable. While work on this workflow editor has begun, it could not be completed due to the absence of time.

Chapter 5 – Discussion

5.1 MYLIP

5.1.1 Introduction

The function of an E3 ligase is to transfer ubiquitin from the E2 enzyme to the target protein. Conceptually, one straightforward approach to find an inhibitor of this function is to establish an assay aimed at disrupting the binding of the E3 ligase to its target. Proximity based assays such as AlphaScreen or HTRF are ideally suited for this purpose by generating a readout that can be reduced in a manner dependent of the dose of the putative inhibitors in a high-throughput format. Having previously established such an assay for the YEATS domain protein family [136], [176], [177], I chose the same approach for MYLIP.

In 2011, Calkin et al.[78] had published a putative mechanism by which MYLIP recognises the cytoplasmic tail of the LDLR via its FERM domain. According to the model, the non-canonical extension of the domain was essential. The titration experiments of FERM domain constructs (residues 1-273 and 1-344) into a set concentration of fluorescently labelled $LDLR_{(811-833)}$, $VLDLR_{(820-842)}$ and $ApoEr2_{(849-871)}$ peptides determined K_D values in the single and low double digit micromolar range (their Figure 7A). At the same time, Sorrentino et al.[88] were also working on the LDLR-MYLIP interaction and published a putative binding mode based on the homology of MYLIP FERM and DAB1 PTB domain (their Figure 8B). While they could not provide an actual crystal structure, they were able to demonstrate that mutations in the identified residues led to a partial (H272A) and near complete (Y265A) rescue of LDLR abundance. These two studies lent a high degree of confidence in the proposed binding mode and the project was pursued.

5.1.2 Isolated MYLIP FERM domain expresses well in *E. coli*

MYLIP FERM domain constructs spanning a range of boundaries and using different affinity and solubility tags were expressed in *E. coli* BL21(DE3) (**Table 17**). While most of them were stable and expressed at sufficient yields, a number did not. Some of the least expressed constructs precipitated after attempted affinity-tag cleavage (MYLIPA-p009/c035, MYLIPA-p010/c010). Both carried the ZBasic[178] solubility tag. Another, also carrying this tag, precipitated after one freeze-thaw cycle (MYLIPA-p004,c028). A small number did not express at all (MYLIPA-p001/c003, MYLIPA-p016/c089, MYLIPA-p017/c090). c089 and c090 were

phospho-mimetic mutants (T279D-S282D, **Figure A31**). CG used different buffers (basis buffer was 50 mM HEPES, 500 mM NaCl, 2% (v/v) glycerol, 0.5 mM TCEP, pH 7.0., supplemented with imidazole for IMAC washes and elution). It is conceivable that these mutants had an impact on solubility under these conditions. Indeed, the theoretical isoelectric points (pI) of c089 and c090 shifted closer to the pH of the buffer, while the pI of c088 shifted away from the buffer pH (**Table 31**).

Table 31: Isoelectric points of phospho-mimetic mutants and parent constructs.

Parent construct	Daughter Construct	Parent pI	Daughter pI
c016	c088	6.39	6.16
c029	c089	8.55	8.09
c035	c090	8.88	8.09

5.1.3 Isolated MYLIP FERM domain does not bind to receptor tail peptides in AlphaScreen or HTRF assays

While I was able to determine K_D values for two MYLIP constructs (c016, M1-MS98 and c014, M1-R276) and $LDLR_{(811-834)}$ using biolayer interferometry (2.2 µM and 3.3 µM, respectively; **Figure 52**), it was not possible to reproduce the observed binding in isothermal titration calorimetry (**Figure 47**), or protein-peptide titrations in AlphaScreen®, HTRF® or fluorescence polarisation (Results section **3.3.1**, **Figure 37** to **Figure 41** and **Figure A33** to **Figure A38**) assays. Another scientist working at the Centre for Medicines Discovery, Charline Giroud, also could not reproduce the results with a phospho-mimetic mutant of MYLIP (**Figure 42**).

Some experiments produced apparent positive results, such as a titration of $MYLIPA_{(M1-R276;pNIC-CTH0)}$ and $MYLIPA_{(M1-S298;pNIC-CTH0)}$ against $LDLR_{(811-859)}$-K-biotin in AlphaScreen® (**Figure 40A**). However, there was no activity at all in the corresponding HTRF assay (**Figure A35**). To validate this result, two ratios of MYLIP to receptor tail were picked (30 nM : 30 nM and 10 nM : 10 nM) and non-biotinylated $LDLR_{(811-834)}$ was titrated into MYLIPA-p011 with $LDLR_{(811-859)}$-K-biotin and MYLIPA-p012 with $LDLR_{(811-859)}$-K-biotin. The non-biotinylated peptide started at 16 µM and 4 µM and was titrated down in 14 steps in 1:1 dilutions. No dose-dependent decrease in signal could be observed (**Figure 41**), indicating that this was a non-specific interaction.

Martinelli et al. discuss the possibility that the membrane context could be essential to bringing the FERM domain into a conformation more amenable to receptor tail binding, something that was not explored in this work, although anchoring receptor tail peptides to lipid vesicles had been discussed briefly. Furthermore, Martinelli et al. showed that while the isolate F3ab region (terminating between residues 275 and 280) forms monomers in solution (at 187 µM), as indicated by small angle x-ray scattering (SAXS) analysis, the partially unravelled C-terminal helix of the region can occupy the peptide binding site of a neighbouring F3ab region when the protein is tightly packed in crystal form. While AlphaScreen beads have a high avidity, they are significantly larger than an antibody (250 nm vs 10 nm) and thus two bound molecules can be further apart than they would be on a bivalent antibody. Since the two MYLIP FERM constructs that showed positive results in AlphaScreen are also terminating at the end of the F3ab region, it is conceivable that two FERM domains are forced in an unfavourable conformation where they obstruct each other on an antibody, whereas they are distributed further apart on an AlphaScreen bead. **Figure 86** shows a comparison between a historical protein-peptide titration and the positive result obtained in this work. A clear drop-in signal with increasing MYLIP FERM domain concentration is observable after an initial peak at 62.5 nM, reminiscent of the high-dose hook effect in immunology.

Regrettably, the results of the fluorescence polarisation experiment published by Calkin et al. in 2011 could not be reproduced with any of the assay techniques tested here. However, there are several factors about the published work that are unclear and data that is missing. For instance, none of the parameters for curve fit are given, apart from K_D. Therefore, it is difficult to judge the quality of the data. Furthermore, it is not explained how the data were fit, i.e., whether the mean of the datapoints was taken and the curve fitted to them, or whether three independent curve fits were performed, and mean values of the parameters were shown. It is also not stated whether Figure 7 is representative data set or mean data set.

These uncertainties, in combination with the fact that there have been no follow-up investigations of FERM domain peptide binding should have been an indicator to not expend too much time and resources in attempting to reproduce the results. Instead, alternative assay techniques should have been explored sooner.

A

MLLT1A-p032(nM) \ biotin-H3K9ac (nM)	3200	1600	800	400	200	100	50	25	12.50	6.25	3.13	1.56	0.78	0.39	0.20	-
3200	959.4	966.7	989.6	976.8	716.8	972.4	741.4	905.0	584.9	224.7	49.0	22.1	23.9	17.9	25.0	24.6
1600	970.5	959.7	1002.6	945.8	969.9	1016.9	933.8	882.7	584.6	199.2	41.9	26.0	22.2	20.5	21.9	25.3
800	1032.4	1036.3	1064.6	986.4	996.1	1039.2	953.3	856.4	533.6	172.5	45.5	28.3	25.3	26.4	28.3	22.6
400	977.8	1013.1	1027.4	1008.9	959.4	1004.4	968.8	866.2	514.5	171.0	41.2	25.8	23.8	22.3	24.4	22.3
200	1065.5	1069.9	1089.1	1020.9	1027.6	888.7	948.6	722.4	328.6	88.2	28.3	15.5	12.6	12.9	13.3	12.5
100	1040.5	1012.4	1061.9	1010.2	995.7	11.5	1014.6	742.8	0.6	21.8	9.9	5.7	4.4	8.9	3.7	4.2
50	1030.8	1033.1	1096.5	1026.4	1036.0	1083.7	894.6	576.5	226.3	54.0	7.8	3.6	3.6	4.2	4.0	3.6
25	837.5	826.0	948.7	894.6	918.5	920.3	829.8	466.3	214.9	87.1	17.6	3.8	3.5	3.5	3.2	3.6
12.50	681.1	736.3	815.2	710.8	851.2	746.3	472.0	188.1	121.8	44.1	11.8	4.6	2.7	3.3	3.1	2.7
6.25	202.3	212.0	350.9	247.8	373.1	291.8	243.8	53.5	43.9	25.6	18.8	3.9	2.2	2.2	1.5	2.2
3.13	52.4	48.4	80.5	46.5	88.7	53.0	28.1	10.2	12.6	7.3	7.1	2.1	1.6	1.8	1.3	1.3
1.56	8.7	5.8	8.6	5.9	10.7	6.6	4.3	2.4	2.1	1.3	2.4	0.9	0.8	1.0	0.6	0.5
0.78	4.0	3.2	4.4	3.9	4.7	3.9	2.5	1.7	0.9	1.1	1.1	0.8	0.8	0.7	0.6	1.6
0.39	3.1	2.2	2.5	3.2	3.5	3.7	2.7	1.2	1.2	0.8	0.6	0.6	0.7	1.5	0.7	0.7
0.20	3.2	2.8	3.0	3.2	3.0	3.5	2.0	1.1	0.7	0.6	0.5	0.5	0.6	0.7	0.6	0.9
-	3.0	1.6	5.5	3.1	2.3	2.2	1.7	0.7	1.1	1.0	0.4	0.6	0.7	0.6	0.8	0.6

B

MYLIPA(1-276 C-His6) (nM) \ LDLR(811-859)-K-biotin (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
16000	443.4	303.9	446.2	406.1	422.3	406.3	417.4	367.5	346.2	309.3	234.1	117.4	60.8	26.6	12.2	5.9
8000	582.9	602.5	603.6	597.3	571.1	557.5	525.7	494.3	435.6	368.9	226.4	111.2	56.5	21	8.7	3.1
4000	592.4	611.4	604.8	565.1	539.7	511.3	474.1	444.7	409.6	318.4	198	103	53.3	23.8	9.7	3.9
2000	502.6	474	486.1	473.5	446.9	431.2	390.8	348	306.3	262.5	160.3	91.9	52.2	22.5	10.5	3.7
1000	226.8	278	315.6	328.2	321.8	311.8	279.7	255.3	216.1	205.1	125	73.3	41.7	18.9	12.2	3.7
500	95.8	113.3	130.1	137.7	134.2	214.5	240.1	213.7	311.7	249.5	156.4	103.8	59.7	33.8	19.7	4.6
250	313.1	400.8	523.8	625.1	622.1	627.5	641.8	591.1	574.7	461.2	289.1	197.4	131.3	74.4	43.8	7.4
125	339.7	460.4	576.8	748.1	806.1	836.8	861.6	874.7	810.5	681.1	441.3	285.3	196	107	61	9.3
62.5	356.5	478.5	673.1	780	759.1	928.4	983.1	978	917.2	818.5	559.2	384.7	266.8	158.9	86.4	7.3
31.25	316.8	377.9	504.4	591.1	692.2	815.6	832.6	852.6	840.3	761.4	518.4	333.6	262.4	159.4	95.1	7.1
15.625	7	925.4	444.8	512.4	620.5	687.1	703.4	691.7	662.3	605.5	390.7	275.7	189.7	116.5	64.9	6.4
7.8125	239	288.5	351	441.8	518.5	567.6	559.8	542.8	545.1	491	322.8	209.5	145.1	75.4	42.6	4.9
3.90625	223	259.9	335	402	456.7	444.5	435	414.6	370.2	336.8	218.5	140.8	85.3	49.1	25.2	3.6
1.953125	203.3	244.7	315.6	375	389	369.7	328.3	275.5	249.5	219.8	138.5	89.9	48.3	25.5	13.6	4.5
0.976563	212.5	233.5	296	363.7	365.3	310.8	224.3	181.5	135.3	112.6	72.8	38.3	27.6	12.1	6.4	2.9
-	183.2	210.8	278.6	359.5	335.4	212.4	86.6	30.1	9.3	5.3	3.2	3	4.1	3.1	3.2	3.1

Figure 86: Comparison of "typical" historical result of a protein-peptide titration and the positive result obtained in this work. Signal in thousands. **A:** Histone acylation reader MLLT1 and acetylated histone 3 peptide (H3K9ac) in AlphaScreen. Experiment performed 26 Mar 2018 by me. **B:** MYIPA$_{(1\text{-}276,\ C\text{-}His6)}$ versus LDLR$_{(811\text{-}859)}$-K-biotin.

5.1.4 PTB domains of DAB1 and DAB2 bind to receptor tails, but cannot facilitate binding to MYLIP FERM domain

As the direct one-to-one binding experiments remained unfruitful, possible adapter proteins were investigated. LDLR is known to associate with a range of proteins on the plasma membrane and it is conceivable that they form a tripartite complex. To this end, the phospho-tyrosine binding domains (PTB) of LDLR adapter protein 1 (LDLRAP1), Disabled-1 (DAB1) and Disabled-2 (DAB2) were expressed. It was possible to reproduce with high confidence previous findings and biotin-VLDLR$_{(820\text{-}842)}$, biotin-ApoER2$_{(849\text{-}871)}$ and biotin-APP$_{(745\text{-}767)}$ were all shown to bind to the PTB domains of DAB1 and DAB2, in both AlphaScreen® (**Figure 43**, **Figure A39**, **Figure A40**) and HTRF® (**Figure A41**, **Figure A42**, **Figure A43**) assays. It was then further investigated whether these PTB domains could facilitate binding between MYLIP FERM domain and LDLR receptor tails. Protein-Peptide titrations were repeated in presence of DAB1

(**Figure 44**) and DAB2 (**Error! Reference source not found.**). Unfortunately, no binding could be observed.

5.1.5 A serendipitously developed MYLIP auto-ubiquitination assay

After Martinelli et al. published a gel-based peptide ubiquitination experiment[106], efforts were shifted to produce an enzymatic assay. A Cy5.5-labelled $LDLR_{(811-33)}$ peptide had already been purchased for fluorescence polarisation assays. Indeed, Western blot (**Figure 53**) and pure gel-fluorescence based assays showed that all enzymes produced in *E. coli* were active and the substrate could be ubiquitinated. This was even though the first batch of E1 had been exposed to air during the ion exchange chromatography step. With HTRF reagents and biotin-labelled receptor tail peptides readily available in the lab, methods to detect poly-ubiquitin chains on high-throughput friendly microtiter plates were explored, and Eu^{3+}-cryptate labelled ubiquitin was purchased. In the first instance, the conditions listed on the PerkinElmer product website were used (see also[162]). The results looked promising: Product formation (**Figure 54**) closely resembled the data shown in the product literature for the reagent. However, efforts to improve the assay window (signal over baseline) were only moderately successful, reaching values between 2.59 and 2.78 for $LDLR_{(811-384)}$ (**Figure 61**, **Figure 62**) and I eventually returned to the standard protocol from reagent manufacturer.

To validate the results, non-biotinylated $LDLR_{(811-834)}$ titrated into the reaction with the expectation that it would compete with $LDLR_{(811-834)}$-K-biotin for binding to MYLIP. Lacking biotin, it would not bind SA-XL665 (FRET acceptor) and thus lead to a dose dependent reduction in FRET signal. This was indeed observed, albeit at very high concentrations (**Figure 64**).

Interestingly, when testing out other receptor tails with NPxY motif to screen for possible new substrates and some peptides without lysine (**Table 14**), signal was observed for all peptides as well as the no-peptide control. This result could not be explained at first. Investigating all components of the assay revealed that revealed that the ubiquitin construct used has TwinStrep tag, which can directly bind to streptavidin. It is feasible that this leads to the formation of FRET pairs along the poly-ubiquitin chain when MYLIP is auto-ubiquitinating (**Figure 87**) and essentially creates a MYLIP auto-ubiquitination assay. While I have not shown in this work that MYLIP auto-ubiquitinates, my colleague Karly Buchan has shown in liquid chromatography/mass spectrometry that the MYLIP construct produced in this work does

indeed auto-ubiquitinate. Interestingly, she could also not detect ubiquitinated receptor tail peptides.

Given the low turnover of substrate in the gel-fluorescence based assay even after overnight incubation, MYLIP does not have optimal activity in solution. Indeed, Martinelli et al. speculated that the F3c' helix of the extended FERM domain serves an auto-inhibitive function that is modulated *in vivo* by conformational changes. These could be induced by post translational modification or membrane association [106]. Interactions with other membrane bound proteins as well as E2 cannot be excluded, either. It is also conceivable that the low-activity conformation possesses is sufficient to produce detectable substrate turnover before MYLIP gets inactivated by either steric hindrance from the auto-ubiquitination or degradation of the protein.

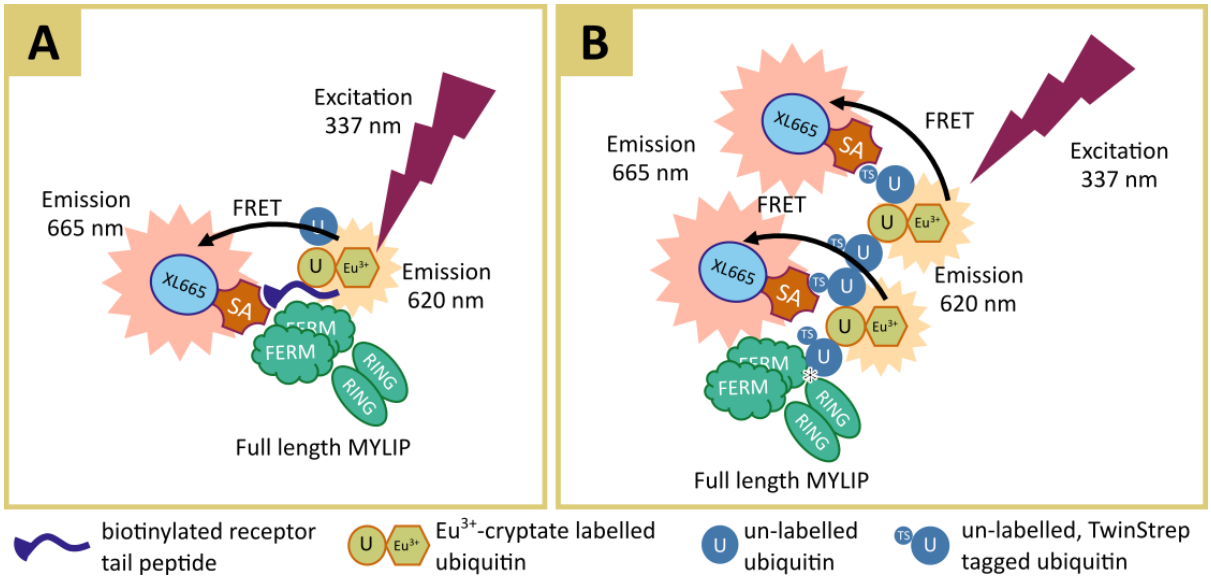


Figure 87: The enzymatic assay produced signal even in the absence of substrate or the presence of known no-substrate peptide. **A:** The expected outcome of the assay. Polyubiquitination on the LDLR peptide leads to formation of FRET pairs when Eu^{3+}-cryptate labelled ubiquitin (yellow, FRET donor) and streptavidin-XL665 (orange and light blue, FRET acceptor) bound to the LDLR peptide's ubiquitin come in proximity. **B:** The most likely model of the actual complex that formed. Eu^{3+}-cryptate labelled ubiquitin (FRET donor) and streptavidin-XL665 (FRET acceptor) bound to the TwinStrep tag of ubiquitin form FRET pairs in absence of biotinylated LDLR peptide. *: MYLIP autoubiquitination sites (K293, K309, K316 or K320).

5.1.6 MYLIP auto-ubiquitination pilot screen.

With the fortuitously discovered MYLIP auto-ubiquitination assay, a subset of the Oxford Dementia Drug Discovery Institute's 10k diversity set was screened in a pilot screen. Data quality was excellent with Z'_{robust} values over 0.8 (**Figure 67A**) and high to low signal ratios around 6:1 for the plates where reagents were dispensed with the automated Multidrop Combi dispense. Even one of the plates that was hand pipetted achieved Z'_{robust} of over 0.7. The last plate was excluded due to aspiration of air bubbles when dispensing reagents. Even if it hadn't been obvious after visible inspection of plates that mis-dispensing has occurred, looking at median assay plate (**Figure 67C**), it was right to exclude rows C,D.

Normalised FRET ratios show a good distribution (**Figure 67B**) with very few "hits". Only 25 compounds out of 2640 (0.95%) showed a reduction in auto-ubiquitination of 70% or greater compared to reference. The lack of any identifiable similarities (**Figure 72**) should not come as a surprise as the pilot screen has been performed with a subset of a diversity library selected to cover a broad chemical space with structurally different compounds. Applying the multidimensional scaling algorithm from ChemMine tools [164] to the entirety of the pilot screen compounds shows an even distribution with no immediately apparent clustering (**Figure 88**). Similarly, the hit compounds are mostly evenly distributed across the entirety of the space covered by the pilot screen.

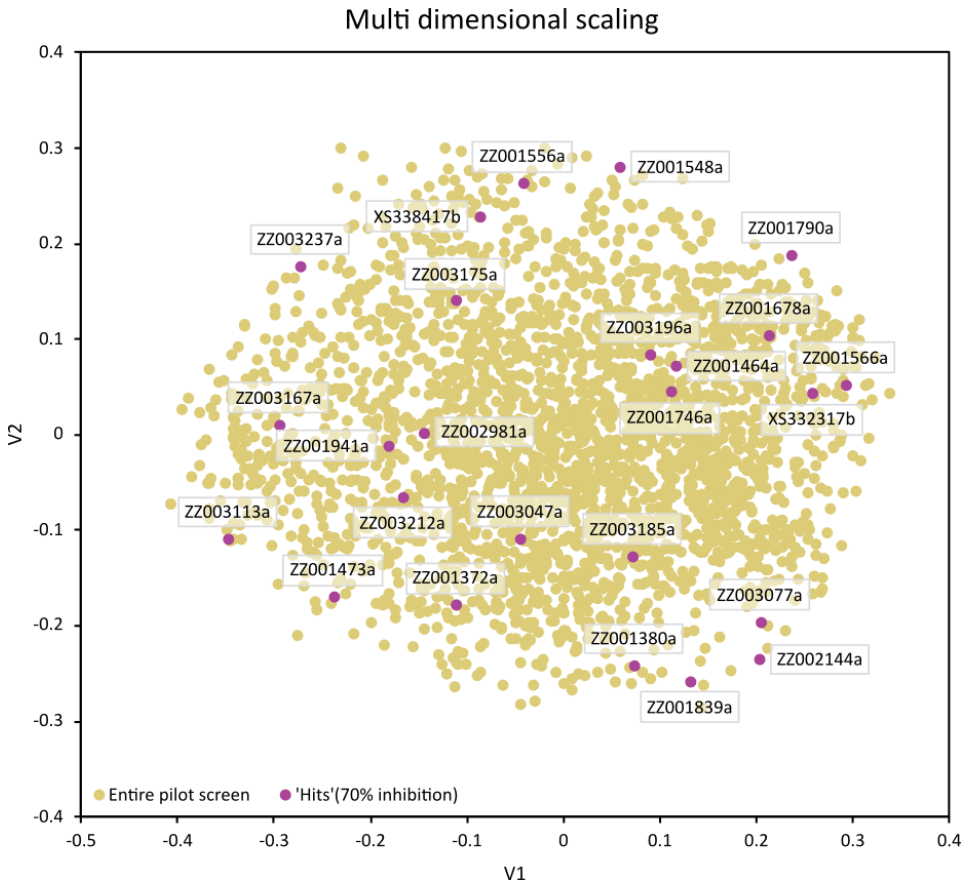


Figure 88: Multidimensional scaling of all-to-all distances of compounds in the MYLIP auto-ubiquitination pilot screen. “Hit” compounds appear to be mostly evenly distributed over the chemical space covered.

5.1.7 Limitations

Since it is necessary to include all enzymes of the ubiquitination cascade in the assay, potential inhibitors have several “points of attack”, complicating the detection of true direct inhibitors of IDOL function. The first point is the binding of ubiquitin to E1 and its activation. The second point is the transfer of activated ubiquitin to E2. The third is the binding of E2 to the RING domain of MYLIP. This represents the first “genuine” interference mechanism, provided the inhibitor binds to MYLIP and not the E2, which might have other E3 ligases it is supplying with ubiquitin. The second genuine interference point would be the transfer of ubiquitin from E2 to the catalytic residue of MYLIP. An inhibitor could conceivably bind near the catalytic residue (C387). The third interference point is the transfer of ubiquitin to the substrate. It is also conceivable to inhibit the homo-dimerisation of MYLIP.

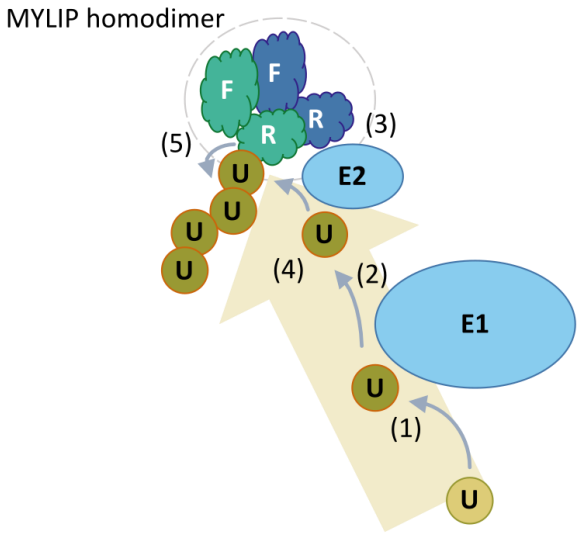


Figure 89: Opportunities to inhibit MYLIP autoubiquitination. (1): Binding to and activation of ubiquitin by E1. (2) Transfer of activated ubiquitin to E2. (3) Binding of E2 to RING domain of MYLIP. (4) Transfer of ubiquitin to RING domain of MYLIP. (5) Transfer of ubiquitin to IDOL autoubiquitination site(s).

5.1.8 Validation of hits

To drive development of selective, high affinity inhibitors, any hits of this pilot screen and any full-scale follow-up must be thoroughly verified. Viable hits would only selectively inhibit the binding of MYLIP to its target protein or to its partner E2 enzyme. Due to the absence of time, this was not pursued in this instance, but several assays can be envisaged:

5.1.8.1 Ubiquitination assay with unrelated E3 ligase.

To filter out any compounds inhibiting any of the steps in the ubiquitination cascade prior to involvement of MYLIP, an assay with an unrelated E3 ligase, but employing the same preceding elements could be utilised. This could be based on the same principles that were originally attempted in this work, namely a substrate with an affinity tag (e.g. hemagglutinin, FLAG, glutathione-S-transferase (GST) or biotin, and either directly FRET labelled ubiquitin or ubiquitin with another affinity, allowing for the use of toolbox reagents to form FRET pairs (e.g. anti-tag-Eu^{3+}-cryptate conjugated antibodies and streptavidin-XL665). The necessary reagents and infrastructure are readily available in our department.

5.1.8.2 Direct binding assays

Another avenue would be to verify the binding of the compounds to MYLIP. This can be achieved by surface plasmon resonance (SPR) or biolayer interferometry (BLI) studies. If the

compounds do not bind to MYLIP or UBE2D isoforms directly, they can be eliminated as valid hits.

SPR: Surface plasmon resonance detects binding of an analyte being flowed over an immobilised ligand on a sensor chip. This is achieved measuring the angle of minimum reflection/maximum absorption of incident light on the gold surface of the sensor chip. The angle is dependent on the thickness of the adsorbed film (i.e. the immobilisation matrix, immobilised ligand, and bound analyte). Relative binding of analyte to ligand can thus be determined and binding characteristics (K_D, k_{on}, k_{off}) can be determined by appropriate experimental design. The BiacoreS200 SPR system in our department uses four flow cells on one chip upon which different proteins can be immobilised and binding to the proteins can be measured. While one flow cell is reserved to detect baseline/background binding to the chip's surface, up to three different proteins can be assayed simultaneously, e.g. isolated FERM domain and isolated RING domain of MYLIP, and the UBE2D isoform utilised in the pilot screen. Using this approach, the mode of action of any valid hit could be identified at the same time. The system accepts 384 well microtiter plates as sample receptacles, allowing for medium throughput in one-concentration-screens and low throughput in the case of K_D determination.

BLI: The OctetRED384 system can employ up to eight sensor tips upon which proteins can be isolated. A more comprehensive discussion of biolayer interferometry is given in section 1.4.3.3 but in brief, binding of an analyte to a ligand immobilised on the sensor's surface leads to a change in the reflection pattern of light passing through the sensor. In contrast to SPR, the sample and reference conditions (e.g. buffer blanks) are stationary on a 384 well microtiter plate and the sensors are dipped into the plate wells. Generally, the same types of experiments can be carried out, but the system is limited in throughput by the microtiter plate.

5.1.9 Future perspectives

The initial results of the pilot screen are encouraging. Only a modest number of compounds scored inhibition values of 70% or greater, relative to baseline. Attempts to optimise the assay and increase specificity must be undertaken. Pre-loading the E2 with ubiquitin, for instance, would remove two non-specific reactions from the cascade, i.e., the activation of ubiquitin by E1 and the transfer from E1 to E2. To reduce batch variability, pre-loaded E2 would have to be separated from unloaded ubiquitin. This could be achieved by letting E1 and E2 react in

presence of excess ubiquitin and ATP, followed by protease digest of the affinity tag on E2. Immobilised metal affinity chromatography could then be used to remove any proteins with intact tag from the mix. Furthermore, generation of a MYLIP mutant lacking auto-ubiquitination sites could drive development of a true substrate ubiquitination assay.

Revisiting the idea of a tripartite complex might also prove fruitful. Considering the low effort required, adding E2 to a MYLIP-LDLR-peptide titration would be a worthwhile effort to investigate if E2 confers conformational changes that increase affinity to LDLR peptides.

Another avenue worth investigating is the plateau that is reached in the continuous assays, see for instance **Figure 55**. While the initial rate of product formation is dependent on the concentration of E1 in the reaction, it appears that the final amount of product is, too. Since ubiquitin and ATP are available in the same amounts in all reactions, one of the enzymes must cease activity at some point. Since it has been shown in this work that full length MYLIP degrades during the expression and/or purification, the stability of MYLIP in the assay bears investigating. Possible approaches could include:

- Setting up the reaction and spiking in fresh MYLIP after the plateau has been reached. Resumption of signal increase could then indicate MYLIP stability as the cause for the plateau.
- Isothermal fluorimetry to detect possible aggregation over time.
- Tracking the appearance of free RING and FERM domain over time by taking samples from a MYLIP solution left at ambient temperature and immediately mixing them with SDS-PAGE loading buffer. If MYLIP does degrade, the relative abundance of intact MYLIP and the separate domains should change.

5.2 Software development discussion

5.2.1 Introduction

Over the course of this work, I developed a versatile platform to build a quickly adaptable data analysis platform on the foundation of a selection of Python libraries. With these, experimental raw data from plate readers and other devices can be cross-referenced with plate maps, either manually entered or extracted from liquid handler transfer reports, to quickly analyse the data with minimal effort from the user. In theory, data could be directly uploaded to the department's database without any user input. In practice and for some

analysis workflows, the user can sight the data and confirm the accuracy of the automated analysis and, if desired, chose different fitting models or exclude failed (sample or reagent dispensing errors) datapoints from the fit. Other workflows are completely hands-off. The individual workflows will be discussed below.

5.2.2 End point single dose (EPSD) workflow

While historically these assays were easy to analyse in our department with a KNIME workflow, BBQ adds value with the assay plate quality control measures, especially the implementation of automatic calculations of Z' and Z'robust. Visualisation of the readings of an assay plate as a heat map also gives the user additional information to judge the quality of the assay and thus the data generated. While again this was possible before by opening the raw data files in MS Excel and applying conditional formatting, BBQ also provides information about the individual wells that the user does not have to cross reference themselves. This workflow was then used to analyse the data of the MYLIP auto-ubiquitination pilot screen conducted in this work.

5.2.3 End Point Dose Response (EPDR) workflow

This workflow offers the same information about plates as the EPSD workflow. In addition, dose response curves are calculated, and the user has the option to remove points from the analysis while retaining data integrity. Analysing these kinds of assays is a lot more time consuming. Prior to the development of BBQ took our team several hours to process a handful of sample plates. Conducting selectivity experiments and running several experiments with different proteins in parallel further exacerbated this. Now, one operator can conceivably carry out the same amount of work in less than an hour, even for a very large number of proteins and compounds, experimental work notwithstanding. The ability to directly upload to the database also saves a significant amount of time since the user does not have to export the data, open the database interface, and then copy and paste the results table into the interface for upload, wait to see if all values are verified.

The biggest limitation of this workflow, however, is that duplicate sample IDs would be treated as if they were replicates, not individual samples. Therefore, it is difficult to run the same compound twice on one plate and obtain separate dose response curves. Instead, more replicates are processed per sample concentration. To solve this problem, unique suffixes

need to be appended to the control ID for each replicate and removed from the results table before uploading to the database. In this case, automatic upload is not possible.

5.2.4 Differential Scanning Fluorimetry (DSF)

This assay was historically performed using 96 well Real time PCR machines. While the department has a Roche LightCycler, which can handle 384 well plates, this has previously been unattainable due to the nature of the workflow. Now, processing this data is trivial and with the addition of the first derivative of the fluorescence as a measure to determine transition points, even data sets with multiple transition points can be analysed.

5.2.5 Cell Based Compound Screen (CBCS)

While this assay still requires an upfront investment in time to extract the "raw data" (i.e., counts of nuclei per well) from the cell culture plate image raw data, discussions with SH and TJ have shown that the time savings are substantial, up to several hours per screen, and allow them to deliver results faster and interrogate the data more easily. In the same vein, the cell-based compound screen analysis workflow allows the phenotypic screening group to perform the data normalisation for one screen in under a minute, leaving the operators more time to focus on the actual results.

5.2.6 Enzymatic Reaction Rate Dose Response (ERDR)

This workflow has been developed in tandem with the assay, therefore there is no historical experience available for comparison. However, the quick development shows the strength of BBQ. Many underlying functions have been previously developed and could be assembled into a functional workflow like Lego pieces. Hopefully this workflow will support the AI-driven Structure-enabled Antiviral Platform (ASAP) consortium's efforts for future pandemic preparedness and drug discovery.

5.2.7 Extensibility

Assay definition files and method editor make it easy to create new workflows without having to hard-code tabs for (meta) data entry. After having established the general structure of an analysis workflow and, it is easy to develop new assay workflows, even if they require *de-novo* programming (see Cell Viability Screen and Enzymatic Reaction rate Dose Response).

Reducing the impact of staff turnover is one of the reasons for developing this software and the assay definition files and method editor, once fully completed, are crucial to achieve this.

Even when the original developer has left, it will still be possible to set up new analysis workflows and connect them to the database, ensuring that new assays do not lead to new bottlenecks in terms of data processing time.

5.2.8 Summary

Once a bespoke BBQ workflow is optimised, the productivity gains are substantial, and the user is free to focus on the actual results as opposed to the process of obtaining them. BBQ can be universally used for most existing Biochemical and Biophysical assay formats. Currently, BBQ is only available for MS Windows PCs and while this is the default operating system in our department, some people do use Linux or iOS devices.

5.2.9 Future directions

While the workflow editor is functionally complete, an interpreter that turns the rule set created by the editor into a cohesive experience for the user still needs to be completed. Al relevant user interface elements have already been developed for the bespoke workflows and generalised classes can easily be developed from them. The same is true for the procedures for data processing. Indeed, some of the existing code base can easily be adapted by adding simple if-then-else branching based on the parameters within the assay definition files. This development should be completed a few weeks to months after submission of this thesis.

In addition to the workflows developed here, discussions with members of the Structural Genomics Consortium in Frankfurt, Germany, have begun at the end of this work to develop bespoke analysis workflows for their NanoBRET assays. Since the Frankfurt site is using a different database software, this is an ideal starting point for the development of an easier database integration, e.g. using database connection definition files similar to the assay definition files.

As discussed above, BBQ is currently distributed amongst the department as a standalone Windows application with no current option for Linux or iOS users. Creating a fully functional, installable Python package (a so called “wheel”) that can be installed on any computer that can support python and is executable from the command line is crucial to reaching a wider user base across different operating systems. A parallel option to reach a wider user base is the development of a browser-based application. Indeed, several services have been designed in our own department, alone or in collaboration with outside partners: ICM-SCARAB has a

web interface, so do SATurn[179], MichelaNglo[180], and Fragalysis (**https://fragalysis.xchem.diamond.ac.uk/**). Browser based applications have the advantage that updates will reach all users at the same time and that they do not require many resources client-side, since all calculations are handled server-side. Developing a web front-end for SCARAB should be the next big goal.

References

[1] Eurostat, “Population structure indicators at national level,” 2023. https://ec.europa.eu/eurostat/databrowser/view/DEMO_PJANIND__custom_964289/bookmark/table?lang=en&bookmarkId=1276db8f-938f-4f55-a801-e0151d01702f (accessed Nov. 22, 2023).

[2] D. Vandenberghe and J. Albrecht, “The financial burden of non-communicable diseases in the European Union: A systematic review,” *Eur. J. Public Health*, vol. 30, no. 4, pp. 833–839, 2020, doi: 10.1093/eurpub/ckz073.

[3] Y. Goryakin *et al.*, “Assessing the future medical cost burden for the European health systems under alternative exposure-to-risks scenarios,” *PLoS One*, vol. 15, no. 9 September, pp. 1–14, 2020, doi: 10.1371/journal.pone.0238565.

[4] C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, and G. Kroemer, “The Hallmarks of Aging,” *Cell*, vol. 153, no. 6, p. 1194, 2013, doi: 10.1016/j.cell.2013.05.039.

[5] V. Kuan *et al.*, “Data-driven identification of ageing-related diseases from electronic health records,” *Sci. Rep.*, vol. 11, no. 1, pp. 1–17, 2021, doi: 10.1038/s41598-021-82459-y.

[6] D. G. Le Couteur and J. Thillainadesan, “What Is an Aging-Related Disease? An Epidemiological Perspective,” *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.*, vol. 77, no. 11, pp. 2168–2174, 2022, doi: 10.1093/gerona/glac039.

[7] R. Beaglehole *et al.*, “Priority actions for the non-communicable disease crisis,” *Lancet*, vol. 377, no. 9775, pp. 1438–1447, 2011, doi: 10.1016/S0140-6736(11)60393-0.

[8] N. Leigh-Hunt *et al.*, “An overview of systematic reviews on the public health consequences of social isolation and loneliness,” *Public Health*, vol. 152, pp. 157–171, 2017, doi: 10.1016/j.puhe.2017.07.035.

[9] World Health Organisation, “WHO Fact Sheet Dementia,” 2023. https://www.who.int/news-room/fact-sheets/detail/dementia.

[10] A. Alzheimer, “Über eine eigenartige Erkrankung der Hirnrinde,” *Allg. Zeitschrift fur Psychiatr. und Psych. Medizin*, vol. 64, pp. 146–148, 1907.

[11] R. A. Stelzmann, H. Norman Schnitzlein, and F. Reed Murtagh, “An english translation of alzheimer’s 1907 paper, ‘über eine eigenartige erkankung der hirnrinde,’” *Clin. Anat.*, vol. 8, no. 6, pp. 429–431, 1995, doi: 10.1002/ca.980080612.

[12] Eurostat, “Causes of death - deaths by country of residence and occurrence,” *Cardiovascular Diseases Statistics*, 2021. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth_cd_aro&lang=en (accessed Feb. 17, 2021).

[13] Alzheimer Europe, *Dementia in Europe Yearbook 2019: Estimating the prevalence of dementia Europe*. Luxembourg: Alzheimer Europe, 2019.

[14] L. Jönsson, A. Tate, O. Frisell, and A. Wimo, “The Costs of Dementia in Europe: An Updated Review and Meta-analysis,” *Pharmacoeconomics*, vol. 41, no. 1, pp. 59–75, 2023, doi: 10.1007/s40273-022-01212-z.

[15] Eurostat, “Living conditions in Europe - income distribution and income inequality,” 2023. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Living_conditions_in_Europe_-_income_distribution_and_income_inequality (accessed Nov. 22, 2023).

[16] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, “Clinical diagnosis of Alzheimer’s disease,” *Neurology*, vol. 34, no. 7, pp. 939–944, Jul. 1984, doi: 10.1212/WNL.34.7.939.

[17] G. G. Glenner and C. W. Wong, “Alzheimer’s disease and Down’s syndrome: Sharing of a unique cerebrovascular amyloid fibril protein,” *Biochem. Biophys. Res. Commun.*, vol. 122, no. 3, pp. 1131–1135, 1984, doi: 10.1016/0006-291X(84)91209-9.

[18] A. Anand, A. A. Patience, N. Sharma, and N. Khurana, “The present and future of pharmacotherapy of Alzheimer’s disease: A comprehensive review,” *Eur. J. Pharmacol.*, vol. 815, no. August, pp. 364–375, 2017, doi: 10.1016/j.ejphar.2017.09.043.

[19] T. G. Beach, S. E. Monsell, L. E. Phillips, and W. Kukull, “Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010,” *J. Neuropathol. Exp. Neurol.*, vol. 71, no. 4, pp. 266–273, 2012, doi: 10.1097/NEN.0b013e31824b211b.

[20] A. E. Budson and P. R. Solomon, “New diagnostic criteria for Alzheimer’s disease and mild cognitive impairment for the practical neurologist,” *Pract. Neurol.*, vol. 12, no. 2, pp. 88–96, 2012, doi: 10.1136/practneurol-2011-000145.

[21] B. Olsson *et al.*, “CSF and blood biomarkers for the diagnosis of Alzheimer’s disease: a systematic review and meta-analysis,” *Lancet Neurol.*, vol. 15, no. 7, pp. 673–684, 2016, doi: 10.1016/S1474-4422(16)00070-3.

[22] Roche, “Roche’s Elecsys Amyloid Plasma Panel Granted FDA Breakthrough Device Designation to Enable a Timely Diagnosis of Alzheimer’s Disease,” 2022. .

[23] E. Karran and B. De Strooper, “The amyloid hypothesis in Alzheimer disease: new insights from new therapeutics,” *Nat. Rev. Drug Discov.*, vol. 21, no. 4, pp. 306–318, 2022, doi: 10.1038/s41573-022-00391-w.

[24] M. Gatz *et al.*, “Role of genes and environments for explaining Alzheimer disease,” *Arch. Gen. Psychiatry*, vol. 63, no. 2, pp. 168–174, 2006, doi: 10.1001/archpsyc.63.2.168.

[25] Q. Jiang *et al.*, “ApoE Promotes the Proteolytic Degradation of Ab,” *Neuron*, vol. 58, no. 5, pp. 681–693, 2008, doi: 10.1016/j.neuron.2008.04.010.

[26] U.S. Food and Drug Administration, “FDA Approves Expanded Use of Treatment for Patients With Severe Alzheimer’s Disease,” 2006. https://web.archive.org/web/20090710031727/http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/2006/ucm108768.htm (accessed Nov. 17, 2023).

[27] U.S. Food and Drug Administration, “Approval Letter - Reminyl,” 2001. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2001/21-169_Reminyl_Approv.pdf (accessed Nov. 17, 2023).

[28] U.S. Food and Drug Administration, “Approval Letter - Rivastigmina,” 2000. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2000/20823_Exelon_Approv.pdf (accessed Nov. 15, 2023).

[29] U.S. Food and Drug Administration, “Approval Letter - Namenda,” 2003. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2003/21-487_Namenda_Approv.pdf (accessed Nov. 17, 2023).

[30] National Institute for Health and Care Excellence, “Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer’s disease -Technology appraisal guidance [TA217],” 2011. https://www.nice.org.uk/Guidance/TA217 (accessed Nov. 15, 2023).

[31] Euorpean Medicines Agency, “Ebixa - EPAR summary for the public,” 2011. https://www.ema.europa.eu/en/documents/overview/ebixa-epar-summary-public_en.pdf.

[32] M. J. H. J. Dekker *et al.*, “Alignment of European Regulatory and Health Technology Assessments: A Review of Licensed Products for Alzheimer’s Disease,” *Front. Med.*, vol. 6, no. May, 2019, doi: 10.3389/fmed.2019.00073.

[33] U.S. Food and Drug Administration, “FDA Grants Accelerated Approval for Alzheimer’s Drug,” 2021. https://www.fda.gov/news-events/press-announcements/fda-grants-accelerated-approval-alzheimers-drug.

[34] D. S. Knopman, D. T. Jones, and M. D. Greicius, “Failure to demonstrate efficacy of aducanumab: An analysis of the EMERGE and ENGAGE trials as reported by Biogen, December 2019,” *Alzheimer’s Dement.*, vol. 17, no. 4, pp. 696–701, 2021, doi: 10.1002/alz.12213.

[35] A. Mullard, “FDA approval for Biogen’s aducanumab sparks Alzheimer disease firestorm,” *Nat. Rev. Drug Discov.*, vol. 20, no. 7, p. 496, 2021, doi: 10.1038/d41573-021-00099-3.

[36] European Medicines Agency, “Aduhelm: Withdrawal of the marketing authorisation application,” 2022. https://web.archive.org/web/20230821135656/https://www.ema.europa.eu/en/medicines/human/withdrawn-applications/aduhelm (accessed Nov. 21, 2023).

[37] U.S. Food and Drug Administration, “Leqembi (lecanemab-irmb) injection - accessdata.fda.gov.” https://www.accessdata.fda.gov/drugsatfda_docs/label/2023/761269Orig1s001lbl.pdf (accessed Nov. 22, 2023).

[38] R. J. Bateman *et al.*, “Two Phase 3 Trials of Gantenerumab in Early Alzheimer’s

Disease.," *N. Engl. J. Med.*, vol. 389, no. 20, pp. 1862–1876, 2023, doi: 10.1056/NEJMoa2304430.

[39] C. H. van Dyck *et al.*, "Lecanemab in early Alzheimer's Disease," *N. Engl. J. Med.*, vol. 388, no. 1, pp. 9–21, 2021, doi: 10.1056/nejmoa2212948.

[40] J. R. Sims *et al.*, "Donanemab in Early Symptomatic Alzheimer Disease: The TRAILBLAZER-ALZ 2 Randomized Clinical Trial," *Jama*, vol. 330, no. 6, pp. 512–527, 2023, doi: 10.1001/jama.2023.13239.

[41] U.S. Food and Drug Administration, "FDA Grants Accelerated Approval for Alzheimer's Disease Treatment," 2023. https://www.fda.gov/news-events/press-announcements/fda-grants-accelerated-approval-alzheimers-disease-treatment (accessed Nov. 17, 2023).

[42] J. Y. Hur, "γ-Secretase in Alzheimer's disease," *Exp. Mol. Med.*, vol. 54, no. 4, pp. 433–446, 2022, doi: 10.1038/s12276-022-00754-8.

[43] H. Thomas *et al.*, "Global Atlas of Cardiovascular Disease 2000-2016: The Path to Prevention and Control," *Glob. Heart*, vol. 13, no. 3, pp. 143–163, 2018, doi: 10.1016/j.gheart.2018.09.511.

[44] R. Luengo-Fernandez *et al.*, "Cardiovascular disease burden due to productivity losses in European Society of Cardiology countries," *Eur. Hear. J. - Qual. Care Clin. Outcomes*, pp. 1–9, 2023, doi: 10.1093/ehjqcco/qcad031.

[45] World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization, *Global Atlas on Cardiovascular Disease Prevention and Control*. Geneva: World Health Organization in collaboration with the World Heart Federation and the World Stroke Organization, 2011.

[46] R. Ross, "Atherogenesis - An Inflammatory Disease," *N. Engl. J. Med.*, vol. 340, no. 2, pp. 115–126, 1999, doi: 10.1056/NEJM199901143400207.

[47] Ross Russell, "The pathogenesis of atherosclerosis: a perspective for the 1990s," *Nature*, vol. 362, no. April, p. 801, 1993.

[48] J. Fan and T. Watanabe, "Inflammatory Reactions in the Pathogenesis of Atherosclerosis," *J. Atheroscler. Thromb.*, vol. 10, no. 2, pp. 63–71, 2003, doi:

10.5551/jat.10.63.

[49] S. K. and I. E., “Functional rafts in cell membranes.,” *Nature*, vol. 387, no. 6633, pp. 569–572, 1997.

[50] J. R. Silvius, “Role of cholesterol in lipid raft formation: Lessons from lipid model systems,” *Biochim. Biophys. Acta - Biomembr.*, vol. 1610, no. 2, pp. 174–183, 2003, doi: 10.1016/S0005-2736(03)00016-6.

[51] J. Hu, Z. Zhang, W. J. Shen, and S. Azhar, “Cellular cholesterol delivery, intracellular processing and utilization for biosynthesis of steroid hormones,” *Nutr. Metab.*, vol. 7, no. Table 1, pp. 7–9, 2010, doi: 10.1186/1743-7075-7-47.

[52] M. S. Brown and J. L. Goldstein, “Lipoprotein metabolism in the macrophage: implications for cholesterol deposition in atherosclerosis,” *Annu. Rev. Biochem.*, vol. 52, pp. 223–261, 1983, doi: 10.1161/01.atv.0000217281.06795.61.

[53] C. Yokoyama *et al.*, “SREBP-1 , a Basic-Helix-Loop-Helix-Leucine Zipper Protein That Controls Transcription of the Low Density Lipoprotein Receptor Gene,” *Cell*, vol. 75, no. 1, pp. 187–197, 1993.

[54] D. Eberlé, B. Hegarty, P. Bossard, P. Ferré, and F. Foufelle, “SREBP transcription factors: Master regulators of lipid homeostasis,” *Biochimie*, vol. 86, no. 11, pp. 839–848, 2004, doi: 10.1016/j.biochi.2004.09.018.

[55] A. Nohturfft, M. S. Brown, and J. L. Goldstein, “Topology of SREBP cleavage-activating protein, a polytopic membrane protein with a sterol-sensing domain,” *J. Biol. Chem.*, vol. 273, no. 27, pp. 17243–17250, 1998, doi: 10.1074/jbc.273.27.17243.

[56] T. Yang *et al.*, “Crucial step in cholesterol homeostasis: Sterols promote binding of SCAP to INSIG-1, a membrane protein that facilitates retention of SREBPs in ER,” *Cell*, vol. 110, no. 4, pp. 489–500, 2002, doi: 10.1016/S0092-8674(02)00872-3.

[57] D. Yabe, M. S. Brown, and J. L. Goldstein, “Insig-2, a second endoplasmic reticulum protein that binds SCAP and blocks export of sterol regulatory element-binding proteins,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 20, pp. 12753–12758, 2002, doi: 10.1073/pnas.162488899.

[58] K. N. Maxwell, R. E. Soccio, E. M. Duncan, E. Sehayek, and J. L. Breslow, “Novel putative

SREBP and LXR target genes identified by microarray analysis in liver of cholesterol-fed mice," *J. Lipid Res.*, vol. 44, no. 11, pp. 2109–2119, 2003, doi: 10.1194/jlr.M300203-JLR200.

[59] P. Costet *et al.*, "Hepatic PCSK9 expression is regulated by nutritional status via insulin and sterol regulatory element-binding protein 1c," *J. Biol. Chem.*, vol. 281, no. 10, pp. 6211–6218, 2006, doi: 10.1074/jbc.M508582200.

[60] J. R. Smith, T. F. Osborne, M. S. Brown, J. L. Goldstein, and G. Gil, "Multiple sterol regulatory elements in promoter for hamster 3-hydroxy-3-methylglutaryl-coenzyme A synthase," *J. Biol. Chem.*, vol. 263, no. 34, pp. 18480–18487, 1988, doi: 10.1016/s0021-9258(19)81383-2.

[61] C. C. Y. Chang, H. Y. Huh, K. M. Cadigan, and T. Y. Chang, "Molecular cloning and functional expression of human acyl-coenzyme A:cholesterol acyltransferase cDNA in mutant Chinese hamster ovary cells," *J. Biol. Chem.*, vol. 268, no. 28, pp. 20747–20755, 1993, doi: 10.1016/s0021-9258(19)36846-2.

[62] C. C. Y. Chang *et al.*, "Regulation and immunolocalization of acyl-coenzyme A:cholesterol acyltransferase in mammalian cells as studied with specific antibodies," *J. Biol. Chem.*, vol. 270, no. 49, pp. 29532–29540, 1995, doi: 10.1074/jbc.270.49.29532.

[63] Y. Zhang, C. Yu, J. Liu, T. A. Spencer, C. C. Y. Chang, and T. Y. Chang, "Cholesterol is superior to 7-ketocholesterol or 7α-hydroxycholesterol as an allosteric activator for acyl-coenzyme A:Cholesterol acyltransferase 1," *J. Biol. Chem.*, vol. 278, no. 13, pp. 11642–11647, 2003, doi: 10.1074/jbc.M211559200.

[64] A. Miyazaki *et al.*, "Expression of ACAT-1 protein in human atherosclerotic lesions and cultured human monocytes-macrophages," *Arterioscler. Thromb. Vasc. Biol.*, vol. 18, no. 10, pp. 1568–1574, 1998, doi: 10.1161/01.ATV.18.10.1568.

[65] T. Y. Hui, L. M. Olivier, S. Kang, and R. A. Davis, "Microsomal triglyceride transfer protein is essential for hepatic secretion of apoB-100 and apoB-48 but not triglyceride," *J. Lipid Res.*, vol. 43, no. 5, pp. 785–793, 2002, doi: 10.1016/s0022-2275(20)30121-8.

[66] D. W. Zhang *et al.*, "Binding of proprotein convertase subtilisin/kexin type 9 to epidermal growth factor-like repeat A of low density lipoprotein receptor decreases

receptor recycling and increases degradation," *J. Biol. Chem.*, vol. 282, no. 25, pp. 18602–18612, 2007, doi: 10.1074/jbc.M702027200.

[67] L. Shan, L. Pang, R. Zhang, N. J. Murgolo, H. Lan, and J. A. Hedrick, "PCSK9 binds to multiple receptors and can be functionally inhibited by an EGF-A peptide," *Biochem. Biophys. Res. Commun.*, vol. 375, no. 1, pp. 69–73, 2008, doi: 10.1016/j.bbrc.2008.07.106.

[68] S. Poirier *et al.*, "The proprotein convertase PCSK9 induces the degradation of low density lipoprotein receptor (LDLR) and its closest family members VLDLR and ApoER2," *J. Biol. Chem.*, vol. 283, no. 4, pp. 2363–2372, 2008, doi: 10.1074/jbc.M708098200.

[69] J. L. Goldstein and M. S. Brown, "Regulation of the mevalonate pathway," *Nature*, vol. 343, no. 6257, pp. 425–430, 1990, doi: 10.1038/343425a0.

[70] M. Nakanishi, J. L. Goldstein, and M. S. Brown, "Multivalent control of 3-hydroxy-3-methylglutaryl coenzyme A reductase. Mevalonate-derived product inhibits translation of mRNA and accelerates degradation of enzyme," *J. Biol. Chem.*, vol. 263, no. 18, pp. 8929–8937, 1988, doi: 10.1016/s0021-9258(18)68397-8.

[71] M. S. Brown and J. L. Goldstein, "Multivalent feedback regulation of HMG CoA reductase, a control mechanism coordinating isoprenoid synthesis and cell growth," *J. Lipid Res.*, vol. 21, no. 5, pp. 505–517, 1980, doi: 10.1016/s0022-2275(20)42221-7.

[72] B. A. Janowski, P. J. Willy, T. R. Devi, J. R. Falck, and D. J. Mangelsdorf, "An oxysterol signalling pathway mediated by the nuclear receptor LXRa," *Nature*, vol. 383, no. 5, pp. 728–731, 1996.

[73] J. J. Repa *et al.*, "Regulation of absorption and ABC1-mediated efflux of cholesterol by RXR heterodimers," *Science (80-.).*, vol. 289, no. 5484, pp. 1524–1529, 2000, doi: 10.1126/science.289.5484.1524.

[74] N. Zelcer, C. Hong, R. Boyadjian, and P. Tontonoz, "LXR Regulates Cholesterol Uptake Through Idol-Dependent Ubiquitination of the LDL Receptor," *Science (80-.).*, vol. 325, no. 5936, pp. 100–104, 2009, doi: 10.1126/science.1168974.

[75] L. A. Green and A. S. Tischler, "Establishment of a noradrenergic clonal line of rat

adrenal pheochromocytoma cells which respond to nerve growth factor," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 73, no. 7, pp. 2424–2428, 1976, doi: 10.1073/pnas.73.7.2424.

[76] C. Hong *et al.*, "The E3 ubiquitin ligase IDOL induces the degradation of the low density lipoprotein receptor family members VLDLR and ApoER2," *J. Biol. Chem.*, vol. 285, no. 26, pp. 19720–19726, 2010, doi: 10.1074/jbc.M110.123729.

[77] L. Zhang *et al.*, "The IDOL-UBE2D complex mediates sterol-dependent degradation of the LDL receptor," *Genes Dev.*, vol. 25, no. 12, pp. 1262–1274, 2011, doi: 10.1101/gad.2056211.

[78] A. C. Calkin *et al.*, "FERM-dependent E3 ligase recognition is a conserved mechanism for targeted degradation of lipoprotein receptors," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 50, pp. 20107–20112, 2011, doi: 10.1073/pnas.1111589108.

[79] O. Wiklund, C. A. Dyer, B. P. Tsao, and L. K. Curtiss, "Stoichiometric binding of apolipoprotein B-specific monoclonal antibodies to low density lipoproteins," *J. Biol. Chem.*, vol. 260, no. 20, pp. 10956–10960, 1985, doi: 10.1016/s0021-9258(17)39129-9.

[80] M. Dashti, W. Kulik, F. Hoek, E. C. Veerman, M. P. Peppelenbosch, and F. Rezaee, "A phospholipidomic analysis of all defined human plasma lipoproteins," *Sci. Rep.*, vol. 1, pp. 1–11, 2011, doi: 10.1038/srep00139.

[81] D. R. Janero, P. Siuta-Mangano, K. W. Miller, and M. D. Lane, "Synthesis, Processing, and Secretion of Hepatic Very Low Density Lipoprotein," *J. Cell. Biochem.*, vol. 24, no. 2, pp. 131–152, 1984, doi: 10.1002/jcb.240240205.

[82] J. P. Segrest, M. K. Jones, H. De Loof, C. G. Brouillette, Y. V. Venkatachalapathi, and G. M. Anantharamaiah, "The amphipathic helix in the exchangeable apolipoproteins: A review of secondary structure and function," *J. Lipid Res.*, vol. 33, no. 2, pp. 141–166, 1992, doi: 10.1016/s0022-2275(20)41536-6.

[83] J. P. Segrest, M. K. Jones, H. De Loof, and N. Dashti, "Structure of apolipoprotein B-100 in low density lipoproteins," *J. Lipid Res.*, vol. 42, no. 9, pp. 1346–1367, 2001, doi: 10.1016/s0022-2275(20)30267-4.

[84] R. M. Lawn *et al.*, "The Tangier disease gene product ABC1 controls the cellular

apolipoprotein-mediated lipid removal pathway," *J. Clin. Invest.*, vol. 104, no. 8, 1999, doi: 10.1172/JCI8119.

[85] T. U. Consortium, "UniProt: the universal protein knowledgebase in 2021," *Nucleic Acids Res.*, vol. 49, no. D1, pp. D480–D489, 2021, doi: 10.1093/nar/gkaa1100.

[86] M. M. Islam, I. Hlushchenko, and S. G. Pfisterer, "Low-Density Lipoprotein Internalization, Degradation and Receptor Recycling Along Membrane Contact Sites," *Front. Cell Dev. Biol.*, vol. 10, no. January, pp. 1–6, 2022, doi: 10.3389/fcell.2022.826379.

[87] X. Wang, R. Sato, M. S. Brown, X. Hua, and J. L. Goldstein, "SREBP-1 , a Membrane-Bound Transcription Factor Released by Sterol-Regulated Proteolysis," *Cell*, vol. 77, no. 1, pp. 53–62, 1994, doi: 10.1016/0092-8674(94)90234-8.

[88] V. Sorrentino, L. Scheer, A. Santos, E. Reits, B. Bleijlevens, and N. Zelcer, "Distinct functional domains contribute to degradation of the Low density lipoprotein receptor (LDLR) by the E3 ubiquitin ligase inducible degrader of the LDLR (IDOL)," *J. Biol. Chem.*, vol. 286, no. 34, pp. 30190–30199, 2011, doi: 10.1074/jbc.M111.249557.

[89] J. K. Nelson *et al.*, "The deubiquitylase USP2 regulates the ldlr pathway by counteracting the E3-Ubiquitin Ligase IDOL," *Circ. Res.*, vol. 118, no. 3, pp. 410–419, 2016, doi: 10.1161/CIRCRESAHA.115.307298.

[90] M. S. Brown and J. L. Goldstein, "A Receptor-Mediated Pathway for Cholesterol Homeostasis," *Science (80-.).*, vol. 232, no. 5, pp. 34–47, 1986, doi: 10.1126/science.3513311.

[91] S. S. Reddy, T. E. Connor, E. J. Weeber, and W. Rebeck, "Similarities and differences in structure, expression, and functions of VLDLR and ApoER2," *Mol. Neurodegener.*, vol. 6, no. 1, pp. 1–10, 2011, doi: 10.1186/1750-1326-6-30.

[92] M. Futema *et al.*, "Analysis of the frequency and spectrum of mutations recognised to cause familial hypercholesterolaemia in routine clinical practice in a UK specialist hospital lipid clinic," *Atherosclerosis*, vol. 229, no. 1, pp. 161–168, 2013, doi: 10.1016/j.atherosclerosis.2013.04.011.

[93] D. Gabcova-Balaziova *et al.*, "Molecular-genetic aspects of familial

hypercholesterolemia," *Endocr. Regul.*, vol. 49, no. 3, pp. 164–181, 2015, doi: 10.4149/endo_2015_03_164.

[94] C. K. Garcia *et al.*, "Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein," *Science (80-.).*, vol. 292, no. 5520, pp. 1394–1398, 2001, doi: 10.1126/science.1060458.

[95] G. He, S. Gupta, M. Yi, P. Michaely, H. H. Hobbs, and J. C. Cohen, "ARH is a modular adaptor protein that interacts with the LDL receptor, clathrin, and AP-2," *J. Biol. Chem.*, vol. 277, no. 46, pp. 44044–44049, 2002, doi: 10.1074/jbc.M208539200.

[96] J. Stewart, G. Manmathan, and P. Wilkinson, "Primary prevention of cardiovascular disease: A review of contemporary guidance and literature," *JRSM Cardiovasc. Dis.*, vol. 6, p. 204800401668721, 2017, doi: 10.1177/2048004016687211.

[97] J. K. Grant and C. E. Orringer, "Coronary and Extra-coronary Subclinical Atherosclerosis to Guide Lipid-Lowering Therapy," *Curr. Atheroscler. Rep.*, no. 0123456789, 2023, doi: 10.1007/s11883-023-01161-8.

[98] K. H. Beyer, "Chlorothiazide: How the thiazides evolved as antihypertensive therapy," *Hypertension*, vol. 22, no. 3, pp. 388–391, 1993, doi: 10.1161/01.HYP.22.3.388.

[99] G. W. K. Wong and J. M. Wright, "Blood pressure lowering efficacy of nonselective beta-blockers for primary hypertension," *Cochrane Database Syst. Rev.*, vol. 2014, no. 2, 2014, doi: 10.1002/14651858.CD007452.pub2.

[100] G. G. Schwartz *et al.*, "Alirocumab and Cardiovascular Outcomes after Acute Coronary Syndrome," *N. Engl. J. Med.*, vol. 379, no. 22, pp. 2097–2107, 2018, doi: 10.1056/nejmoa1801174.

[101] M. S. Sabatine *et al.*, "Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease," *N. Engl. J. Med.*, vol. 376, no. 18, pp. 1713–1722, 2017, doi: 10.1056/nejmoa1615664.

[102] U.S. Food and Drug Administration, "Leqvio (inclisiran) injection," 2021. .

[103] K. K. Ray *et al.*, "Long-term efficacy and safety of inclisiran in patients with high cardiovascular risk and elevated LDL cholesterol (ORION-3): results from the 4-year open-label extension of the ORION-1 trial," *Lancet Diabetes Endocrinol.*, vol. 11, no. 2,

pp. 109–119, 2023, doi: 10.1016/S2213-8587(22)00353-9.

[104] P. A. Olsson, L. Korhonen, E. A. Mercer, and D. Lindholm, “MIR is a novel ERM-like protein that interacts with myosin regulatory light chain and inhibits neurite outgrowth,” *J. Biol. Chem.*, vol. 274, no. 51, pp. 36288–36392, 1999, doi: 10.1074/jbc.274.51.36288.

[105] P. A. Olsson, B. C. Bornhauser, L. Korhonen, and D. Lindholm, “Neuronal expression of the ERM-like protein MIR in rat brain and its localization to human chromosome 6,” *Biochem. Biophys. Res. Commun.*, vol. 279, no. 3, pp. 879–883, 2000, doi: 10.1006/bbrc.2000.4028.

[106] L. Martinelli *et al.*, “Structural analysis of the LDL receptor–interacting FERM domain in the E3 ubiquitin ligase IDOL reveals an obscured substrate-binding site,” *J. Biol. Chem.*, vol. 295, no. 39, pp. 13570–13583, 2020, doi: 10.1074/jbc.RA120.014349.

[107] D. C. Bosanquet, L. Ye, K. G. Harding, and W. G. Jiang, “FERM family proteins and their importance in cellular movements and wound healing (Review),” *Int. J. Mol. Med.*, vol. 34, no. 1, pp. 3–12, 2014, doi: 10.3892/ijmm.2014.1775.

[108] B. C. Bornhauser, P. A. Olsson, and D. Lindholm, “MSAP is a novel MIR-interacting protein that enhances neurite outgrowth and increases myosin regulatory light chain,” *J. Biol. Chem.*, vol. 278, no. 37, pp. 35412–35420, 2003, doi: 10.1074/jbc.M306271200.

[109] B. C. Bornhauser, C. Johansson, and D. Lindholm, “Functional activities and cellular localization of the ezrin, radixin, moesin (ERM) and RING zinc finger domains in MIR,” *FEBS Lett.*, vol. 553, no. 1–2, pp. 195–199, 2003, doi: 10.1016/S0014-5793(03)01010-X.

[110] K. Nagano *et al.*, “PDGF regulates the actin cytoskeleton through hnRNP-K-mediated activation of the ubiquitin E3-ligase MIR,” *EMBO J.*, vol. 25, no. 9, pp. 1871–1882, 2006, doi: 10.1038/sj.emboj.7601059.

[111] R. A. Blake *et al.*, “SU6656, a Selective Src Family Kinase Inhibitor, Used To Probe Growth Factor Signaling,” *Mol. Cell. Biol.*, vol. 20, no. 23, pp. 9018–9027, 2000, doi: 10.1128/mcb.20.23.9018-9027.2000.

[112] S. B. Joseph *et al.*, “Synthetic LXR ligand inhibits the development of atherosclerosis in mice,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 11, pp. 7604–7609, 2002, doi:

10.1073/pnas.112059299.

[113] P. G. Traves *et al.*, “Selective activation of liver X receptors by acanthoic acid-related diterpenes,” *Mol. Pharmacol.*, vol. 71, no. 6, pp. 1545–1553, 2007, doi: 10.1124/mol.106.031906.

[114] D. P. Aden, A. Fogel, S. Plotkin, I. Damjanov, and B. B. Knowles, “Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line,” *Nature*, vol. 282, pp. 615–616, 1979, doi: 10.1038/282615a0.

[115] D. H. Lee and A. L. Goldberg, “Proteasome inhibitors: valuable new tools for cell biologists,” *Trends Cell Biol.*, vol. 8, no. 10, pp. 397–403, 1998, doi: 10.1016/S0962-8924(98)01346-4.

[116] D. Komander and M. Rape, “The ubiquitin code,” *Annu. Rev. Biochem.*, vol. 81, pp. 203–229, 2012, doi: 10.1146/annurev-biochem-060310-170328.

[117] N. Foot, T. Henshall, and S. Kumar, “Ubiquitination and the regulation of membrane proteins,” *Physiol. Rev.*, vol. 97, no. 1, pp. 253–281, 2017, doi: 10.1152/physrev.00012.2016.

[118] P. S. Freemont, I. M. Hanson, and J. Trowsdale, “A Novel Cystein-Rich Sequence Motif,” *Cell*, vol. 64, no. 3, pp. 483–484, 1991, doi: 10.1016/0092-8674(91)90229-R.

[119] P. Barlow, B. Luisi, A. Milner, M. Elliot, and R. Everett, “Structure of the C3HC4 Domain by 1H-nuclear Magnetic Resonance Spectroscopy: A New Structural Class of Zinc-finger,” *J. Mol. Biol.*, vol. 237, no. 2, pp. 201–211, 1994, doi: doi.org/10.1006/jmbi.1994.1222.

[120] E. Branigan, J. C. Penedo, and R. T. Hay, “Ubiquitin transfer by a RING E3 ligase occurs from a closed E2 ~ ubiquitin conformation,” *Nat. Commun.*, vol. 11, p. 2846, 2020, doi: 10.1038/s41467-020-16666-y.

[121] T. Yao and R. E. Cohen, “A cryptic protease couples deubiquitination and degradation by the proteasome,” *Nature*, vol. 419, no. 6905, pp. 403–407, 2002, doi: 10.1038/nature01071.

[122] J. Choi, J. Kim, J. Gao, C. Hong, J. Kim, and P. Tontonoz, “The E3 ubiquitin ligase Idol controls brain LDL receptor levels, ApoE clearance and Ab amyloidosis,” *Sci. Transl.*

Med., vol. 7, no. 314, 2015, doi: 10.1126/scitranslmed.aad1904.

[123] J. L. Jankowsky *et al.*, "Mutant presenilins specifically elevate the levels of the 42 residue b -amyloid peptide in vivo : evidence for augmentation of a 42-specific c secretase," *Hum. Mol. Genet.*, vol. 13, no. 2, pp. 159–170, 2004, doi: 10.1093/hmg/ddh019.

[124] J. M. Basak, P. B. Verghese, H. Yoon, J. Kim, and D. M. Holtzman, "Low-density Lipoprotein Receptor Represents an Apolipoprotein E-independent Pathway of Ab Uptake and Degradation by Astrocytes," *J. Biol. Chem.*, vol. 287, no. 17, pp. 13959–13971, 2012, doi: 10.1074/jbc.M111.288746.

[125] J. Gao *et al.*, "The E3 ubiquitin ligase IDOL regulates synaptic ApoER2 levels and is important for plasticity and learning," *Elife*, vol. 6, pp. 1–24, 2017, doi: 10.7554/eLife.29178.

[126] J. Gao *et al.*, "Therapeutic IDOL Reduction Ameliorates Amyloidosis and Improves Cognitive Function in APP/PS1 Mice," *Mol. Cell. Biol.*, vol. 40, no. 8, 2020, doi: 10.1128/mcb.00518-19.

[127] U. Beffert *et al.*, "ApoE Receptor 2 Controls Neuronal Survival in the Adult Brain," *Curr. Biol.*, vol. 16, no. 24, pp. 2446–2452, 2006, doi: 10.1016/j.cub.2006.10.029.

[128] V. Sorrentino *et al.*, "Identification of a loss-of-function inducible degrader of the low-density lipoprotein receptor variant in individuals with low circulating low-density lipoprotein," *Eur. Heart J.*, vol. 34, no. 17, pp. 1292–1297, 2013, doi: 10.1093/eurheartj/ehs472.

[129] D. Adi *et al.*, "IDOL G51S Variant Is Associated With High Blood Cholesterol and Increases Low-Density Lipoprotein Receptor Degradation," *Arterioscler. Thromb. Vasc. Biol.*, no. December, pp. 1–12, 2019, doi: 10.1161/ATVBAHA.119.312589.

[130] D. Weissglas-Volkov *et al.*, "The N342S MYLIP polymorphism is associated with high total cholesterol and increased LDL receptor degradation in humans," *J. Clin. Invest.*, vol. 121, no. 8, pp. 3062–3071, 2011, doi: 10.1172/JCI45504.

[131] N. M. van Loon *et al.*, "Inactivation of the E3 Ubiquitin Ligase IDOL Attenuates Diet-Induced Obesity and Metabolic Dysfunction in Mice," *Arterioscler. Thromb. Vasc. Biol.*,

vol. 38, no. 8, pp. 1785–1795, 2018, doi: 10.1161/ATVBAHA.118.311168.

[132] E. K. Leitch *et al.*, "Inhibition of low-density lipoprotein receptor degradation with a cyclic peptide that disrupts the homodimerization of IDOL E3 ubiquitin ligase," *Chem. Sci.*, vol. 9, no. 27, pp. 5957–5966, 2018, doi: 10.1039/c8sc01186a.

[133] S. Ito *et al.*, "CNPY2 inhibits MYLIP-mediated AR protein degradation in prostate cancer cells," *Oncotarget*, vol. 9, no. 25, pp. 17645–17655, 2018, doi: 10.18632/oncotarget.24824.

[134] L. Zhao, Y. Zhao, Y. He, and Y. Mao, "miR-19b promotes breast cancer metastasis through targeting MYLIP and its related cell adhesion molecules," *Oncotarget*, vol. 8, no. 38, pp. 64330–64343, 2017, doi: 10.18632/oncotarget.19278.

[135] H. Wang, L. Ma, X. Pan, Z. Du, and Y. Chen, "Novel associations of SNPs MYLIP rs3757354 and ABCA1 2230806 gene with early-onset-preeclampsia: A case-control candidate genetic study," *Pregnancy Hypertens.*, vol. 23, no. December 2020, pp. 185–190, 2021, doi: 10.1016/j.preghy.2020.12.005.

[136] T. Christott *et al.*, "Discovery of a Selective Inhibitor for the YEATS Domains of ENL/AF9," *SLAS Discov. Adv. Life Sci. R&D*, vol. 24, no. 2, pp. 133–141, 2018, doi: 10.1177/2472555218809904.

[137] Y. Xiong *et al.*, "Discovery of a Potent and Selective Fragment-like Inhibitor of Methyllysine Reader Protein Spindlin 1 (SPIN1)," *J. Med. Chem.*, vol. 62, no. 20, 2019, doi: 10.1021/acs.jmedchem.9b00522.

[138] V. Fagan *et al.*, "A Chemical Probe for Tudor Domain Protein Spindlin1 to Investigate Chromatin Function," *J. Med. Chem.*, vol. 62, no. 20, 2019, doi: 10.1021/acs.jmedchem.9b00562.

[139] B. Raux *et al.*, "Discovery of PFI-6, a small-molecule chemical probe for the YEATS domain of MLLT1 and MLLT3," *Bioorg. Med. Chem. Lett.*, no. August, p. 129546, 2023, doi: 10.1016/j.bmcl.2023.129546.

[140] A. T. Londregan *et al.*, "Discovery of High-Affinity Small-Molecule Binders of the Epigenetic Reader YEATS4," *J. Med. Chem.*, vol. 66, no. 1, pp. 460–472, 2023, doi: 10.1021/acs.jmedchem.2c01421.

[141] C. S. Lebakken *et al.*, "Development and applications of a broad-coverage, TR-FRET-based kinase binding assay platform," *J. Biomol. Screen.*, vol. 14, no. 8, pp. 924–935, 2009, doi: 10.1177/1087057109339207.

[142] A. Velazquez-Campoy, S. A. Leavitt, and E. Freire, "Characterization of Protein–Protein Interactions by Isothermal Titration Calorimetry," *Methods Mol. Biol.*, vol. 261, pp. 35–54, 2004, doi: 10.1007/978-1-4939-2425-7_11.

[143] Higher Education Statistics Agency, "Where do they work, come from and go to?," 2023. https://www.hesa.ac.uk/data-and-analysis/staff/location (accessed Nov. 17, 2023).

[144] University of Toronto, "UK-Canadian consortium commits $95 million to international health research project," *EurekAlert!*, 2003. https://www.eurekalert.org/news-releases/566997 (accessed Nov. 20, 2023).

[145] F. H. Niesen, H. Berglund, and M. Vedadi, "The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability," *Nat. Protoc.*, vol. 2, no. 9, pp. 2212–2221, 2007, doi: 10.1038/nprot.2007.321.

[146] M. R. Berthold *et al.*, "KNIME: The Konstanz Information Miner," in *Data Analysis, Machine Learning and Applications. Studies in Classification, Data Analysis, and Knowledge Organization*, C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker, Eds. Springer, Berlin, Heidelberg, 2008, pp. 319–326.

[147] "Anaconda 3." Anaconda Inc., [Online]. Available: https://www.anaconda.com/.

[148] J.-H. Zhang, T. D. Y. Chung, and K. R. Oldenburg, "A Simple Statistical Parameter of Use in Evaluation and Validation of High Throughput Screening Assays," *J. Biomulecular Screen.*, vol. 4, no. 2, pp. 64–73, 1999, doi: 10.1177/108705719900400206.

[149] X. D. Zhang, "A pair of new statistical parameters for quality control in RNA interference high-throughput screening assays," *Genomics*, vol. 89, no. 4, pp. 552–561, 2007, doi: 10.1016/j.ygeno.2006.12.014.

[150] C. Strain-Dammerell, P. Mahajan, O. Gileadi, and N. A. Burgess-Brown, "Medium-Throughput Production of Recombinant Human Proteins: Ligation-Independent Cloning," in *Structural Genomics. Methods in Molecular Biology*, Y. W. Chen, Ed.

Totowa, NJ: Humana Press, 2014, pp. 55–72.

[151] C. Aslanidis and P. J. de Jong, “Ligation-independent cloning of PCR products (LIC-PCR),” *Nucleic Acids Res.*, vol. 18, no. 20, pp. 6069–6074, 1990, doi: 10.1093/nar/18.20.6069.

[152] D. T. Jones, “Protein secondary structure prediction based on position-specific scoring matrices,” *J. Mol. Biol.*, vol. 292, no. 2, pp. 195–202, 1999, doi: 10.1006/jmbi.1999.3091.

[153] D. T. Jones and D. Cozzetto, “Structural bioinformatics DISOPRED3 : precise disordered region predictions with annotated protein-binding activity,” *Bioinformatics*, vol. 31, no. 3, pp. 857–863, 2015, doi: 10.1093/bioinformatics/btu744.

[154] D. W. A. Buchan and D. T. Jones, “The PSIPRED Protein Analysis Workbench: 20 years on,” *Nucleic Acids Res. Res.*, vol. 47, no. W1, pp. W402–W407, 2019, doi: 10.1093/nar/gkz297.

[155] A. F. Carvalho *et al.*, “High-yield expression in Escherichia coli and purification of mouse ubiquitin-activating enzyme E1,” *Mol. Biotechnol.*, vol. 51, no. 3, pp. 254–261, 2012, doi: 10.1007/s12033-011-9463-x.

[156] R. Stevens, L. Stevens, and N. C. Price, “The stabilities of various thiol compounds used in protein purifications,” *Biochem. Educ.*, vol. 11, no. 2, p. 70, 1983, doi: 10.1016/0307-4412(83)90048-1.

[157] J. Peppard, F. Glickman, Y. He, S. I. Hu, J. Doughty, and R. Goldberg, “Development of a high-throughput screening assay for inhibitors of aggrecan cleavage using luminescent oxygen channeling (AlphaScreen™),” *J. Biomol. Screen.*, vol. 8, no. 2, pp. 149–156, 2003, doi: 10.1177/1087057103252308.

[158] E. F. Ullman *et al.*, “Luminescent oxygen channeling immunoassay: Measurement of particle binding kinetics by chemiluminescence,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 91, no. 12, pp. 5426–5430, 1994, doi: 10.1073/pnas.91.12.5426.

[159] H. Bazin, M. Préaudat, E. Trinquet, and G. Mathis, “Homogeneous time resolved fluorescence resonance energy transfer using rare earth cryptates as a tool for probing molecular interactions in biology,” *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, vol. 57, no. 11, pp. 2197–2211, 2001, doi: 10.1016/S1386-1425(01)00493-0.

[160] W. A. Lea and A. Simeonov, “Fluorescence polarization assays in small molecule screening,” *Expert Opin. Drug Discov.*, vol. 6, no. 1, pp. 17–32, 2011, doi: 10.1517/17460441.2011.537322.

[161] M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske, and S. Duhr, “Molecular interaction studies using microscale thermophoresis,” *Assay Drug Dev. Technol.*, vol. 9, no. 4, pp. 342–353, 2011, doi: 10.1089/adt.2011.0380.

[162] N. Yabuki, S. Watanabe, T. Kudoh, S. Nihira, and C. Miyamoto, “Application of homogeneous time-resolved fluorescence (HTRF^{TM}) to monitor poly-ubiquitination of wild-type p53,” *Comb. Chem. High Throughput Screen.*, vol. 2, no. 5, pp. 279–287, 1999, doi: 10.2174/1386207302666220204200224.

[163] B. R. Terlouw, S. P. J. M. Vromans, and M. H. Medema, “PIKAChU: a Python-based informatics kit for analysing chemical units,” *J. Cheminform.*, vol. 14, no. 1, pp. 1–17, 2022, doi: 10.1186/s13321-022-00616-5.

[164] T. W. H. Backman, Y. Cao, and T. Girke, “ChemMine tools: An online service for analyzing and clustering small molecules,” *Nucleic Acids Res.*, vol. 39, no. SUPPL. 2, pp. 486–491, 2011, doi: 10.1093/nar/gkr320.

[165] P. Tol, “Paul Tol’s Notes,” 2021. https://web.archive.org/web/20230128071616/https://personal.sron.nl/~pault/ (accessed Jan. 28, 2023).

[166] The pandas develoment team, “pandas-dev/pandas: Pandas.” Feb. 2020.

[167] C. R. Harris *et al.*, “Array programming with NumPy,” *Nature*, vol. 585, no. 7825, pp. 357–362, 2020, doi: 10.1038/s41586-020-2649-2.

[168] P. Virtanen *et al.*, “SciPy 1.0: fundamental algorithms for scientific computing in Python,” *Nat. Methods*, vol. 17, no. 3, pp. 261–272, 2020, doi: 10.1038/s41592-019-0686-2.

[169] L. H. Negri and C. Vestri, “PeakUtils,” 2017. https://github.com/lucashn/peakutils.

[170] J. D. Hunter, “Matplotlib: A 2D graphics environment,” *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 90–95, 2007, doi: 10.1109/MCSE.2007.55.

[171] MolSoft and The Structural Genomics Consortium, “ICM-Scarab - Data Capture,

Integrator and Miner." [Online]. Available: https://molsoft.com/icm-scarab.html.

[172] Oracle Corporation, "oracledb." [Online]. Available: https://oracle.github.io/python-oracledb/.

[173] J. Kitchin, "Nonlinear curve fitting with parameter confidence intervals," 2013. https://kitchingroup.cheme.cmu.edu/blog/2013/02/12/Nonlinear-curve-fitting-with-parameter-confidence-intervals/ (accessed Nov. 26, 2023).

[174] L. L. Maggiora, C. W. Smith, and Z. Y. Zhang, "A General Method for the Preparation of Internally Quenched Fluorogenic Protease Substrates Using Solid-Phase Peptide Synthesis," *J. Med. Chem.*, vol. 35, no. 21, pp. 3727–3730, 1992, doi: 10.1021/jm00099a001.

[175] W. P. Lu and L. Fei, "A logarithmic approximation to initial rates of enzyme reactions," *Anal. Biochem.*, vol. 316, no. 1, pp. 58–65, 2003, doi: 10.1016/S0003-2697(03)00034-4.

[176] M. Moustakim *et al.*, "Discovery of an MLLT1/3 YEATS Domain Chemical Probe," *Angew. Chemie - Int. Ed.*, vol. 57, pp. 16302–16307, 2018, [Online]. Available: https://chemrxiv.org/articles/Discovery_of_an_MLLT1_3_YEATS_Domain_Chemical_Probe/7090547%0Ahttps://ndownloader.figshare.com/files/13039349%0Ahttps://ndownloader.figshare.com/files/13039352.

[177] Thomas Christott; Moses Moustakim; James Bennett; Carmen Coxon; Octovia Monteiro; Charline Giroud; Paul E Brennan; Oleg Fedorov, "Human Mixed- Lineage Leukemia, Translocated to 1 (MLLT1); A Target Enabling Package," *Zenodo*, 2019. .

[178] M. Hedhammar and S. Hober, "Zbasic-A novel purification tag for efficient protein recovery," *J. Chromatogr. A*, vol. 1161, no. 1–2, pp. 22–28, 2007, doi: 10.1016/j.chroma.2007.05.091.

[179] D. Damerell, C. Strain-Dammerell, S. Garsot, S. P. Joyche, P. Barrett, and B. D. Marsden, "SATurn: a modular bioinformatics framework for the design of robust maintainable web-based and standalone applications," *Bioinformatics*, vol. 35, no. 2, pp. 349–351, 2019, doi: 10.1093/bioinformatics/bty549.

[180] M. P. Ferla, A. P. Pagnamenta, D. Damerell, J. C. Taylor, and B. D. Marsden,

"MichelaNglo: sculpting protein views on web pages without coding," *Bioinformatics*, vol. 36, no. 10, pp. 3268–3270, 2020, doi: 10.1093/bioinformatics/btaa104.

Appendix

DNA templates

```
ATGCTGTGTTATGTGACGAGGCCGGACGCGGTGCTGATGGAGGTGGAGGTGGAGGCGAAAGCCAAC
GGCGAGGACTGCCTCAACCAGGTGTGCAGGCGACTGGGAATCATAGAAGTTGACTATTTTGGACTG
CAGTTTACGGGTAGCAAAGGTGAAAGTTTATGGCTAAACCTGAGAAACCGGATCTCCCAGCAGATG
GATGGGCTAGCCCCTTACAGGCTTAAACTTAGAGTCAAGTTCTTCGTGGAGCCTCATCTCATCTTA
CAGGAGCAGACTAGGCATATCTTTTTCTTGCACATCAAGGAGGCCCTCTTGGCAGGCCACCTCTTG
TGTTCCCCAGAGCAGGCAGTGGAACTCAGTGCCCTCCTGGCCCAGACCAAGTTTGGAGACTACAAC
CAGAACACTGCCAAGTATAACTATGAGGAGCTCTGTGCCAAGGAGCTCTCCTCTGCCACCTTGAAC
AGCATTGTTGCAAAACATAAGGAGTTGGAGGGGACCAGCCAGGCTTCAGCTGAATACCAAGTTTTG
CAGATTGTGTCGGCAATGGAAAACTATGGCATAGAATGGCATTCTGTGCGGGATAGCGAAGGGCAG
AAACTGCTCATTGGGGTTGGACCTGAAGGAATCTCAATTTGTAAAGATGACTTTAGCCCAATTAAT
AGGATAGCTTATCCTGTGGTGCAGATGGCCACCCAGTCAGGAAAGAATGTATATTTGACGGTCACC
AAGGAATCTGGGAACAGCATCGTGCTCTTGTTTAAAATGATCAGCACCAGGGCGGCCAGCGGGCTC
TACCGAGCGATAACAGAGACGCACGCATTCTACAGGTGTGACACAGTGACCAGCGCCGTGATGATG
CAGTATAGCCGTGACTTGAAGGGCCACTTGGCATCTCTGTTTCTGAATGAAAACATTAACCTTGGC
AAGAAATATGTCTTTGATATTAAAAGAACATCAAAGGAGGTGTATGACCATGCCAGGAGGGCTCTG
TACAATGCTGGCGTTGTGGACCTCGTTTCAAGAAACAACCAGAGCCCTTCACACTCGCCTCTGAAG
TCCTCAGAAAGCAGCATGAACTGCAGCAGCTGCGAGGGCCTCAGCTGCCAGCAGACCCGGGTGCTG
CAGGAGAAGCTACGCAAGCTGAAGGAAGCCATGCTGTGCATGGTGTGCTGCGAGGAGGAGATCAAC
TCCACCTTCTGTCCCTGTGGCCACACTGTGTGCTGTGAGAGCTGCGCCGCCCAGCTACAGTCATGT
CCCGTCTGCAGGTCGCGTGTGGAGCATGTCCAGCACGTCTATCTGCCAACGCACACCAGTCTTCTC
AATCTGACTGTAATCTAA
```

Figure A1: MYLIP DNA template

```
ATGGACGCGCTCAAGTCGGCGGGGCGGGCGCTGATCCGGAGCCCCAGCTTGGCCAAGCAGAGCTGG
GGGGGCGGTGGCCGGCACCGCAAGCTGCCTGAGAACTGGACAGACACGCGGGAGACGCTGCTGGAG
GGGATGCTGTTCAGCCTCAAGTACCTGGGCATGACGCTAGTGGAGCAGCCCAAGGGTGAGGAGCTG
TCGGCCGCCGCCATCAAGAGGATCGTGGCTACAGCTAAGGCCAGTGGGAAGAAGCTGCAGAAGGTG
ACTCTGAAGGTGTCGCCACGGGGAATTATCCTGACAGACAACCTCACCAACCAGCTCATTGAGAAC
GTGTCCATATACAGGATCTCCTATTGCACAGCAGACAAGATGCACGACAAGGTGTTTGCATACATC
GCCCAGAGCCAGCACAACCAGAGCCTCGAGTGCCACGCCTTCCTCTGCACCAAGCGGAAGATGGCA
CAGGCTGTTACCCTCACCGTAGCCCAGGCCTTCAAAGTCGCCTTTGAGTTTTGGCAGGTGTCCAAG
GAAGAGAAAGAGAAGAGGGACAAAGCCAGCCAAGAGGGAGGGGACGTCCTGGGGGCCCGCCAAGAC
TGCACCCCCCCCTTGAAGAGCTTGGTCGCCACTGGGAACCTGCTGGACTTAGAGGAGACGGCTAAG
GCCCCGCTGTCCACGGTCAGCGCCAACACCACCAACATGGACGAGGTGCCGCGGCCACAAGCCTTG
AGTGGCAGCAGTGTTGTCTGGGAGCTGGATGATGGCCTGGATGAAGCGTTTTCGAGGCTTGCCCAG
TCTCGGACAAACCCTCAGGTCCTGGACACTGGCCTGACAGCCCAGGACATGCATTACGCCCAGTGC
CTCTCGCCTGTCGACTGGGACAAGCCTGACAGCAGCGGCACAGAGCAGGATGACCTCTTCAGCTTC
TGA
```

Figure A2: LDLRAP1/ARH DNA template

```
ATGGACGCGCTCAAGTCGGCGGGGCGGGCGCTGATCCGGAGCCCCAGCTTGGCCAAGCAGAGCTGG
GGGGGCGGTGGCCGGCACCGCAAGCTGCCTGAGAACTGGACAGACACGCGGGAGACGCTGCTGGAG
GGGATGCTGTTCAGCCTCAAGTACCTGGGCATGACGCTAGTGGAGCAGCCCAAGGGTGAGGAGCTG
TCGGCCGCCGCCATCAAGAGGATCGTGGCTACAGCTAAGGCCAGTGGGAAGAAGCTGCAGAAGGTG
ACTCTGAAGGTGTCGCCACGGGGAATTATCCTGACAGACAACCTCACCAACCAGCTCATTGAGAAC
GTGTCCATATACAGGATCTCCTATTGCACAGCAGACAAGATGCACGACAAGGTGTTTGCATACATC
GCCCAGAGCCAGCACAACCAGAGCCTCGAGTGCCACGCCTTCCTCTGCACCAAGCGGAAGATGGCA
CAGGCTGTTACCCTCACCGTAGCCCAGGCCTTCAAAGTCGCCTTTGAGTTTTGGCAGGTGTCCAAG
GAAGAGAAAGAGAAGAGGGACAAAGCCAGCCAAGAGGGAGGGGACGTCCTGGGGGCCCGCCAAGAC
TGCACCCCCCCCTTGAAGAGCTTGGTCGCCACTGGGAACCTGCTGGACTTAGAGGAGACGGCTAAG
GCCCCGCTGTCCACGGTCAGCGCCAACACCACCAACATGGACGAGGTGCCGCGGCCACAAGCCTTG
AGTGGCAGCAGTGTTGTCTGGGAGCTGGATGATGGCCTGGATGAAGCGTTTTCGAGGCTTGCCCAG
TCTCGGACAAACCCTCAGGTCCTGGACACTGGCCTGACAGCCCAGGACATGCATTACGCCCAGTGC
CTCTCGCCTGTCGACTGGGACAAGCCTGACAGCAGCGGCACAGAGCAGGATGACCTCTTCAGCTTC
TGA
```

Figure A3: DAB1 DNA template

ATGTCTAACGAAGTAGAAACAAGTGCAACCAATGGTCAGCCCGACCAACAGGCCGCACCAAAAGCA
CCCTCAAAGAAGGAAAAAAAGAAAGGCCCTGAAAAGACAGATGAATATCTCTTAGCAAGGTTCAAA
GGCGATGGTGTAAAATATAAGGCCAAGCTGATTGGCATTGATGATGTGCCAGATGCAAGAGGGGAT
AAAATGAGCCAAGACTCTATGATGAAACTAAAGGGAATGGCGGCAGCTGGTCGGTCTCAGGGACAA
CACAAACAAAGGATCTGGGTCAACATTTCCCTTTCTGGGATAAAAATAATTGATGAGAAAACTGGG
GTAATAGAGCATGAACATCCAGTAAATAAGATTTCTTTCATTGCCCGTGATGTGACAGACAACCGG
GCATTTGGTTACGTGTGTGGAGGAGAAGGCCAGCATCAGTTTTTTGCCATAAAAACCGGGCAACAG
GCTGAACCATTAGTTGTTGATCTTAAAGACCTTTTTCAAGTTATCTATAATGTAAAGAAAAAGGAA
GAAGAAAAGAAAAAGATAGAGGAAGCCAGCAAAGCAGTTGAGAATGGGAGTGAGGCCCTAATGATT
CTAGATGACCAAACTAACAAACTGAAATCGGGTGTTGACCAGATGGATTTGTTTGGGGACATGTCT
ACACCTCCTGACCTAAATAGTCCAACAGAAAGCAAAGATATCCTGTTAGTGGATCTAAACTCTGAA
ATCGACACCAATCAGAATTCTTTAAGAGAAAATCCATTCTTAACAAACGGCATCACCTCCTGTTCT
CTTCCTCGACCAACGCCTCAGGCATCCTTCTTGCCTGAAAATGCCTTTTCTGCCAATCTCAACTTC
TTTCCCACCCCTAATCCTGATCCTTTCCGTGACGATCCTTTCACACAGCCAGACCAATCGACACCT
TCTTCGTTTGATTCTCTCAAATCTCCAGATCAGAAGAAAGAGAATTCGAGTAGCTCGTCTACTCCG
CTGAGTAATGGGCCCCTGAATGGTGATGTTGACTACTTTGGTCAGCAATTTGACCAGATCTCTAAC
CGGACTGGCAAACAGGAAGCTCAGGCAGGCCCATGGCCCTTTTCAAGTTCGCAAACCCAGCCAGCA
GTGAGAACTCAAAATGGGGTATCTGAAAGAGAACAGAACGGCTTCTCTGTCAAATCCTCCCCGAAC
CCTTTTGTGGGAAGCCCTCCCAAAGGACTGTCCATACAGAATGGCGTAAAGCAGGACTTGGAAAGC
TCTGTCCAGTCCTCACCACATGACTCCATAGCCATTATCCCACCTCCACAAAGTACCAAACCAGGA
AGAGGCAGAAGGACTGCTAAGTCTTCAGCCAATGACTTGCTTGCATCAGACATCTTTGCTCCTCCC
GTCTCAGAACCTTCAGGCCAGGCGTCACCCACAGGACAACCTACAGCCCTGCAGCCCAACCCTCTG
GATCTCTTCAAAACAAGTGCTCCTGCCCCAGTGGGGCCCCTGGTGGGTCTAGGTGGTGTAACTGTC
ACACTCCCTCAGGCAGGACCATGGAACACAGCATCTTTGGTCTTCAATCAGTCCCCTTCAATGGCT
CCGGGAGCCATGATGGGTGGTCAACCTTCAGGTTTTAGTCAGCCCGTCATTTTTGGTACAAGTCCA
GCTGTTTCAGGTTGGAACCAGCCTTCACCCTTTGCAGCCTCAACTCCCCCTCCAGTGCCTGTTGTC
TGGGGCCCTTCTGCATCTGTGGCACCCAATGCTTGGTCAACAACAAGCCCTTTGGGGAATCCTTTT
CAGAGCAATATTTTTCCAGCTCCTGCTGTGTCCACTCAGCCCCCATCCATGCACTCCTCTCTCCTG
GTCACTCCTCCTCAGCCACCTCCCAGAGCTGGCCCTCCCAAGGACATCTCCAGTGATGCCTTCACT
GCCTTAGACCCACTTGGGGATAAAGAGATCAAGGATGTGAAAGAAATGTTTAAGGATTTCCAACTG
CGGCAGCCACCTGCTGTGCCCGCGCGGAAGGGAGAGCAGACTTCTTCTGGGACTTTGAGTGCCTTT
GCCAGTTATTTCAACAGCAAGGTTGGCATTCCTCAGGAGAATGCAGACCATGATGACTTTGATGCT
AATCAACTATTGAACAAGATCAATGAACCACCAAAGCCAGCTCCCAGACAAGTTTCCCTGCCAGTT
ACCAAATCTACTGACAATGCATTTGAGAACCCTTTCTTTAAAGATTCTTTTGGTTCATCACAAGCC
TCTGTGGCTTCTTCTCAACCTGTATCTTCTGAGATGTATAGGGATCCATTTGGAAATCCTTTTGCC
TAA

Figure A4: DAB2 DNA template

Vector Properties

Table A1: Resistance markers, affinity tags and protease site of vectors used.

Vector Name	Antibiotic Resistance Marker	Tags for Purification	Protease Site
pNIC28-Bsa4	Kanamycin	N-terminal His_6	TEV
pNIC-CTH0	Kanamycin	C-terminal His_6	TEV
pNIC-ZB	Kanamycin	N-terminal ZB+His_6	TEV
pNIC-Bio3	Kanamycin	N-terminal His_6+C-terminal Avi	TEV
pNIC-NHStIIT	Kanamycin	N-terminal His_6+Twin-Strep	TEV
pSUMO-LIC	Kanamycin	N-terminal SUMO+His_6	SUMO

Table A2: Restriction sites, matching dNTP pairs for vector an insert T4 polymerase treatment,

Vector Name	Restriction Site for LIC	dNTP for Vector	dNTP for Insert
pNIC28-Bsa4	BsaI	dGTP	dCTP
pNIC-CTH0	BfuAI	dCTP	dGTP
pNIC-ZB	BsaI	dGTP	dCTP
pNIC-Bio3	BsaI	dGTP	dCTP
pNIC-NHStIIT	BsaI	dGTP	dCTP
pSUMO-LIC	BsaI	dTTP	dATP

Table A3: Primer extension for inserts and bp added during PCR

Vector Name	5' LIC Primer Extension	3' LIC Primer Extension	bp added during PCR
pNIC28-Bsa4	TACTTCCAATCCATG	TATCCACCTTTACTG*TCA*	~200
pNIC-CTH0	TTAAGAAGGAGATATACTATG	GATTGGAAGTAGAGGTTCTCTGC	~200
pNIC-ZB	TACTTCCAATCCATG	TATCCACCTTTACTG*TCA*	~400
pNIC-Bio3	TACTTCCAATCCATG	TATCCACCTTTACTGCT	~300
pNIC-NHStIIT	TACTTCCAATCCATG	TATCCACCTTTACTG*TCA*	~350
pSUMO-LIC	ACCAGGAGCAAACGGGAGGTATG	CAGACCGCCACCGCT*TCA*	~500

Colony screen primers for vectors used in this work:

pLIC-F: TGTGAGCGGATAACAATTCC

pLIC-R: AGCAGCCAACTCAGCTTCC

MYLIP protein constructs

Table A4: Complete list of designed MYLIP constructs. "MYLIPA-" prefix is omitted from construct and primer IDs.

ID	Start / Stop	Expected mass (Da)		Vector	Forward primer	Reverse primer	Status	Cloner
		(tagged)	(no tag)					
c001	1-273	33104.9	30639.2	pNIC28-Bsa4	f001	r034	Cloned	LS
c002	1-276	33571.4	31105.8	pNIC28-Bsa4	f001	r033	Cloned	LS
c003	1-280	33989.9	31524.2	pNIC28-Bsa4	f001	r028	Cloned	LS
c004	1-298	35967.2	33501.5	pNIC28-Bsa4	f001	r025	Cloned	LS
c005	1-300	36227.5	33761.9	pNIC28-Bsa4	f001	r022	Cloned	LS
c006	1-312	37614.1	35148.5	pNIC28-Bsa4	f001	r020	Cloned	LS
c007	1-324	39096.8	36631.1	pNIC28-Bsa4	f001	r018	Cloned	LS
c008	1-326	39305	36839.4	pNIC28-Bsa4	f001	r014	Cloned	LS
c009	1-329	39688.5	37222.8	pNIC28-Bsa4	f001	r012	Cloned	LS
c010	1-333	40150	37684.3	pNIC28-Bsa4	f001	r009	Cloned	LS
c011	1-335	40306.2	37840.5	pNIC28-Bsa4	f001	r005	Cloned	LS
c012	1-338	40633.6	38167.9	pNIC28-Bsa4	f001	r001	Cloned	LS
c013	1-273	32328	31418.1	pNIC-CTH0	f002	r035	Cloned	LS
c014	1-276	32794.6	31884.6	pNIC-CTH0	f002	r032	Cloned	LS
c015	1-280	33213	32303.1	pNIC-CTH0	f002	r030	Cloned	LS
c016	1-298	35190.3	34280.4	pNIC-CTH0	f002	r027	Cloned	LS
c017	1-300	35450.7	34540.7	pNIC-CTH0	f002	r024	Cloned	LS
c018	1-312	36837.3	35927.4	pNIC-CTH0	f002	r021	Cloned	LS
c019	1-324	38319.9	37410	pNIC-CTH0	f002	r017	Cloned	LS
c020	1-326	38528.2	37618.2	pNIC-CTH0	f002	r015	Cloned	LS
c021	1-329	38911.6	38001.7	pNIC-CTH0	f002	r010	Cloned	LS
c022	1-333	39373.1	38463.2	pNIC-CTH0	f002	r007	Cloned	LS
c023	1-335	39529.3	38619.4	pNIC-CTH0	f002	r006	Cloned	LS
c024	1-338	39856.7	38946.8	pNIC-CTH0	f002	r002	Cloned	LS
c025	1-273	39597.2	30639.2	pNIC-ZB	f001	r034	Cloned	LS
c026	1-276	40063.8	31105.8	pNIC-ZB	f001	r033	Cloned	LS
c027	1-280	40482.2	31524.2	pNIC-ZB	f001	r028	Cloned	LS
c028	1-298	42459.5	33501.5	pNIC-ZB	f001	r025	Cloned	LS
c029	1-300	42719.9	33761.9	pNIC-ZB	f001	r022	Cloned	LS
c030	1-312	44106.5	35148.5	pNIC-ZB	f001	r020	Cloned	LS
c031	1-324	45589.1	36631.1	pNIC-ZB	f001	r018	Cloned	LS
c032	1-326	45797.4	36839.4	pNIC-ZB	f001	r014	Cloned	LS
c033	1-329	46180.8	37222.8	pNIC-ZB	f001	r012	Cloned	LS
c034	1-333	46642.3	37684.3	pNIC-ZB	f001	r009	Cloned	LS

Table A4: Complete list of designed MYLIP constructs. "MYLIPA-" prefix is omitted from construct and primer IDs.

ID	Start / Stop	Expected mass (Da)		Vector	Forward primer	Reverse primer	Status	Cloner
		(tagged)	(no tag)					
c035	1-335	46798.5	37840.5	pNIC-ZB	f001	r005	Cloned	LS
c036	1-338	47125.9	38167.9	pNIC-ZB	f001	r001	Cloned	LS
c037	1-273	43225.3	30665.3	pSUMO-LIC	f003	r036	Cloned	LS
c038	1-276	43691.9	31131.9	pSUMO-LIC	f003	r031	Cloned	LS
c039	1-280	44110.3	31550.3	pSUMO-LIC	f003	r029	Cloned	LS
c040	1-299	46087.6	33527.6	pSUMO-LIC	f003	r026	Cloned	LS
c041	1-301	46348	33788	pSUMO-LIC	f003	r023	Cloned	LS
c042	1-312	47734.6	35174.6	pSUMO-LIC	f003	r019	Cloned	LS
c043	1-324	49217.2	36657.2	pSUMO-LIC	f003	r016	Cloned	LS
c044	1-326	49425.5	36865.5	pSUMO-LIC	f003	r013	Cloned	LS
c045	1-330	49808.9	37248.9	pSUMO-LIC	f003	r011	Cloned	LS
c046	1-333	50270.4	37710.4	pSUMO-LIC	f003	r008	Cloned	LS
c047	1-335	50426.6	37866.6	pSUMO-LIC	f003	r004	Cloned	LS
c048	1-338	50754	38194	pSUMO-LIC	f003	r003	Cloned	LS
c049	1-445	52463.3	49997.6	pNIC28-Bsa4	f006	r039	Cloned	TC
c050	1-445	51686.4	50776.5	pNIC-CTH0	f005	r040	Insert ready	
c051	1-445	58955.6	49997.6	pNIC-ZB	f006	r039	Insert ready	
c052	1-445	62470.6	49910.5	pSUMO-LIC	f004	r037	Cloned	TC
c053	1-445	55402.4	52388.2	pNIC-Bio3	f006	r038	Cloned	TC
c054	1-441	52036.7	49571.1	pNIC28-Bsa4	f006	r043	Cloned	TC
c055	1-441	51259.9	50349.9	pNIC-CTH0	f005	r042	Insert ready	
c056	1-441	58529.1	49571.1	pNIC-ZB	f006	r043	Insert ready	
c057	1-441	62044	49484	pSUMO-LIC	f004	r041	Insert ready	
c058	1-345	41419.4	38953.7	pNIC28-Bsa4	f006	r046	Insert ready	
c059	1-345	40642.5	39732.6	pNIC-CTH0	f005	r047	Insert ready	
c060	1-345	47911.7	38953.7	pNIC-ZB	f006	r046	Insert ready	
c061	1-345	51426.7	38866.6	pSUMO-LIC	f004	r044	Insert ready	
c062	1-340	40819.8	38354.1	pNIC28-Bsa4	f006	r049	Insert ready	
c063	1-340	40042.9	39133	pNIC-CTH0	f005	r050	Insert ready	
c064	1-340	47312.1	38354.1	pNIC-ZB	f006	r049	Insert ready	
c065	1-340	50827.1	38267	pSUMO-LIC	f004	r048	Insert ready	
c066	1-283	34249.2	31783.5	pNIC28-Bsa4	f006	r052	Insert ready	
c067	1-283	33472.3	32562.4	pNIC-CTH0	f005	r053	Insert ready	
c068	1-283	40741.5	31783.5	pNIC-ZB	f006	r052	Insert ready	
c069	1-283	44256.4	31696.4	pSUMO-LIC	f004	r051	Insert ready	
c070	1-279	33890.8	31425.1	pNIC28-Bsa4	f006	r055	Insert ready	
c071	1-279	33113.9	32204	pNIC-CTH0	f005	r056	Insert ready	

Table A4: Complete list of designed MYLIP constructs. "MYLIPA-" prefix is omitted from construct and primer IDs.

ID	Start / Stop	Expected mass (Da)		Vector	Forward primer	Reverse primer	Status	Cloner
		(tagged)	(no tag)					
c072	1-279	40383.1	31425.1	pNIC-ZB	f006	r055	Insert ready	
c073	1-279	43898	31338	pSUMO-LIC	f004	r054	Insert ready	
c074	1-277	33674.6	31208.9	pNIC28-Bsa4	f006	r058	Insert ready	
c075	1-277	32897.7	31987.8	pNIC-CTH0	f005	r059	Insert ready	
c076	1-277	40166.9	31208.9	pNIC-ZB	f006	r058	Insert ready	
c077	1-277	43681.9	31121.8	pSUMO-LIC	f004	r057	Insert ready	
c078	183-345	20971.8	18506.2	pNIC28-Bsa4	f009	r046	Insert ready	
c079	183-345	20194.9	19285	pNIC-CTH0	f008	r047	Insert ready	
c080	183-345	27464.2	18506.2	pNIC-ZB	f009	r046	Insert ready	
c081	183-345	30847.9	18287.9	pSUMO-LIC	f007	r044	Insert ready	
c082	183-345	23911	20896.8	pNIC-Bio3	f009	r045	Insert ready	
c083	183-445	32015.7	29550.1	pNIC28-Bsa4	f009	r039	Insert ready	
c084	183-445	31238.9	30328.9	pNIC-CTH0	f008	r040	Insert ready	
c085	183-445	38508.1	29550.1	pNIC-ZB	f009	r039	Insert ready	
c086	183-445	41891.8	29331.8	pSUMO-LIC	f007	r037	Insert ready	
c087	183-445	34954.9	31940.7	pNIC-Bio3	f009	r038	Insert ready	
c088	1-298	35232.2	34322.3	pNIC-CTH0			Verified	CG
c089	1-300	42761.8	33803.8	pNIC-ZB			Verified	CG
c090	1-335	46840.4	37882.5	pNIC-ZB			Verified	CG

LS: Cloned via biotech pipeline by Leela Shrestha

TC: Thomas Christott

CG: Charline Giroud

MYLIP Primers

Table A5: MYLIP forward primers

ID	Sequence
MYLIPA-f001	TACTTCCAATCCATGCTGTGTTATGTGACGAGGCC
MYLIPA-f002	TTAAGAAGGAGATATACTATGCTGTGTTATGTGACGAGGCCG
MYLIPA-f003	ACCAGGAGCAAACGGGAGGTATGCTGTGTTATGTGACGA
MYLIPA-f004	ACCAGGAGCAAACGGGAGGTATGCTGTGTTATGTGACGA
MYLIPA-f005	TTAAGAAGGAGATATACTATGCTGTGTTATGTGACGAGGCCG
MYLIPA-f006	TACTTCCAATCCATGCTGTGTTATGTGACGAGGCC

Table A5: MYLIP forward primers	
ID	**Sequence**
MYLIPA-f007	ACCAGGAGCAAACGGGAGGTGAAAACTATGGCATAGAATGGC
MYLIPA-f008	TTAAGAAGGAGATATACTATGGAAAACTATGGCATAGAATGGCATTCTGTGC
MYLIPA-f009	TACTTCCAATCCATGGAAAACTATGGCATAGAATGGCATTCT

Table A6: MYLIP forward primers	
ID	**Sequence**
MYLIPA-r001	TATCCACCTTTACTGTCAGAGGTCCACAACGCCAGC
MYLIPA-r002	GATTGGAAGTAGAGGTTCTCTGCGAGGTCCACAACGCCAGC
MYLIPA-r003	CAGACCGCCACCGCTTCATAGGAGGTCCACAACGCCAGC
MYLIPA-r004	CAGACCGCCACCGCTTCATAGAACGCCAGCATTGTACAG
MYLIPA-r005	TATCCACCTTTACTGTCAAACGCCAGCATTGTACAGA
MYLIPA-r006	GATTGGAAGTAGAGGTTCTCTGCAACGCCAGCATTGTACAG
MYLIPA-r007	GATTGGAAGTAGAGGTTCTCTGCAGCATTGTACAGAGCCCT
MYLIPA-r008	CAGACCGCCACCGCTTCATAGAGCATTGTACAGAGCCCT
MYLIPA-r009	TATCCACCTTTACTGTCAAGCATTGTACAGAGCCCT
MYLIPA-r010	GATTGGAAGTAGAGGTTCTCTGCAGCCCTCCTGGCATGGTC
MYLIPA-r011	CAGACCGCCACCGCTTCATAGAGCCCTCCTGGCATGGTC
MYLIPA-r012	TATCCACCTTTACTGTCAAGCCCTCCTGGCATGGTC
MYLIPA-r013	CAGACCGCCACCGCTTCATAGGGCATGGTCATACACCTC
MYLIPA-r014	TATCCACCTTTACTGTCAGGCATGGTCATACACCTC
MYLIPA-r015	GATTGGAAGTAGAGGTTCTCTGCGGCATGGTCATACACCTC
MYLIPA-r016	CAGACCGCCACCGCTTCATAGGTCATACACCTCCTTTGATGT
MYLIPA-r017	GATTGGAAGTAGAGGTTCTCTGCGTCATACACCTCCTTTGATGT
MYLIPA-r018	TATCCACCTTTACTGTCAGTCATACACCTCCTTTGATGTTCT
MYLIPA-r019	CAGACCGCCACCGCTTCATAGGACATATTTCTTGCCAAGGTT
MYLIPA-r020	TATCCACCTTTACTGTCAGACATATTTCTTGCCAAGGTTAATGTTTTCA
MYLIPA-r021	GATTGGAAGTAGAGGTTCTCTGCGACATATTTCTTGCCAAGGTT
MYLIPA-r022	TATCCACCTTTACTGTCAAAACAGAGATGCCAAGTGGC
MYLIPA-r023	CAGACCGCCACCGCTTCATAGAAACAGAGATGCCAAGTGG
MYLIPA-r024	GATTGGAAGTAGAGGTTCTCTGCAAACAGAGATGCCAAGTGG
MYLIPA-r025	TATCCACCTTTACTGTCAAGATGCCAAGTGGCCCTT
MYLIPA-r026	CAGACCGCCACCGCTTCATAGAGATGCCAAGTGGCCCTT
MYLIPA-r027	GATTGGAAGTAGAGGTTCTCTGCAGATGCCAAGTGGCCCTT
MYLIPA-r028	TATCCACCTTTACTGTCACACTGTGTCACACCTGTA
MYLIPA-r029	CAGACCGCCACCGCTTCATAGCACTGTGTCACACCTGTA
MYLIPA-r030	GATTGGAAGTAGAGGTTCTCTGCCACTGTGTCACACCTGTA
MYLIPA-r031	CAGACCGCCACCGCTTCATAGCCTGTAGAATGCGTGCGT
MYLIPA-r032	GATTGGAAGTAGAGGTTCTCTGCCCTGTAGAATGCGTGCGT
MYLIPA-r033	TATCCACCTTTACTGTCACCTGTAGAATGCGTGCGT

MYLIPA-r034	TATCCACCTTTACTGTCATGCGTGCGTCTCTGTTAT
MYLIPA-r035	GATTGGAAGTAGAGGTTCTCTGCTGCGTGCGTCTCTGTTAT
MYLIPA-r036	CAGACCGCCACCGCTTCATAGTGCGTGCGTCTCTGTTAT
MYLIPA-r037	CAGACCGCCACCGCTTCAGATTACAGTCAGATTGAGAAGAC
MYLIPA-r038	TATCCACCTTTACTGCTGATTACAGTCAGATTGAGAAGACTGG
MYLIPA-r039	TATCCACCTTTACTGTCAGATTACAGTCAGATTGAGAAGACTGGT
MYLIPA-r040	GATTGGAAGTAGAGGTTCTCTGCGATTACAGTCAGATTGAGAAGAC
MYLIPA-r041	CAGACCGCCACCGCTTCAATTGAGAAGACTGGTGTGC
MYLIPA-r042	GATTGGAAGTAGAGGTTCTCTGCATTGAGAAGACTGGTGTGC
MYLIPA-r043	TATCCACCTTTACTGTCAATTGAGAAGACTGGTGTGCG
MYLIPA-r044	CAGACCGCCACCGCTTCAGCTCTGGTTGTTTCTTGAAA
MYLIPA-r045	TATCCACCTTTACTGCTGCTCTGGTTGTTTCTTGAAACG
MYLIPA-r046	TATCCACCTTTACTGTCAGCTCTGGTTGTTTCTTGAAACG
MYLIPA-r047	GATTGGAAGTAGAGGTTCTCTGCGCTCTGGTTGTTTCTTGAAA
MYLIPA-r048	CAGACCGCCACCGCTTCATGAAACGAGGTCCACAAC
MYLIPA-r049	TATCCACCTTTACTGTCATGAAACGAGGTCCACAACG
MYLIPA-r050	GATTGGAAGTAGAGGTTCTCTGCTGAAACGAGGTCCACAAC
MYLIPA-r051	CAGACCGCCACCGCTTCAGGCGCTGGTCACTGTGTC
MYLIPA-r052	TATCCACCTTTACTGTCAGGCGCTGGTCACTGTGTC
MYLIPA-r053	GATTGGAAGTAGAGGTTCTCTGCGGCGCTGGTCACTGTGTC
MYLIPA-r054	CAGACCGCCACCGCTTCATGTGTCACACCTGTAGAATG
MYLIPA-r055	TATCCACCTTTACTGTCATGTGTCACACCTGTAGAATGC
MYLIPA-r056	GATTGGAAGTAGAGGTTCTCTGCTGTGTCACACCTGTAGAATG
MYLIPA-r057	CAGACCGCCACCGCTTCAACACCTGTAGAATGCGTG
MYLIPA-r058	TATCCACCTTTACTGTCAACACCTGTAGAATGCGTGC
MYLIPA-r059	GATTGGAAGTAGAGGTTCTCTGCACACCTGTAGAATGCGTG

MYLIP test expressions

Table A7: MYLIPA-c001 - c048 test expression results

Expression ID	Construct ID	Elution yield
MYLIPA-e001	MYLIPA-c001	0: None
MYLIPA-e002	MYLIPA-c003	5: High
MYLIPA-e003	MYLIPA-c004	3: Medium
MYLIPA-e004	MYLIPA-c005	0: None
MYLIPA-e005	MYLIPA-c006	1: Low
MYLIPA-e006	MYLIPA-c007	0: None
MYLIPA-e007	MYLIPA-c008	1: Low
MYLIPA-e008	MYLIPA-c009	1: Low
MYLIPA-e009	MYLIPA-c010	3: Medium
MYLIPA-e010	MYLIPA-c011	0: None
MYLIPA-e011	MYLIPA-c012	1: Low
MYLIPA-e012	MYLIPA-c013	0: None
MYLIPA-e013	MYLIPA-c014	5: High
MYLIPA-e014	MYLIPA-c015	0: None
MYLIPA-e015	MYLIPA-c016	5: High
MYLIPA-e016	MYLIPA-c017	0: None
MYLIPA-e017	MYLIPA-c018	1: Low
MYLIPA-e018	MYLIPA-c019	0: None
MYLIPA-e019	MYLIPA-c020	3: Medium
MYLIPA-e020	MYLIPA-c021	0: None
MYLIPA-e021	MYLIPA-c022	3: Medium
MYLIPA-e022	MYLIPA-c023	0: None
MYLIPA-e023	MYLIPA-c024	3: Medium
MYLIPA-e024	MYLIPA-c025	3: Medium
MYLIPA-e025	MYLIPA-c026	5: High
MYLIPA-e026	MYLIPA-c027	5: High
MYLIPA-e027	MYLIPA-c028	5: High
MYLIPA-e028	MYLIPA-c029	5: High
MYLIPA-e029	MYLIPA-c030	3: Medium
MYLIPA-e030	MYLIPA-c031	3: Medium
MYLIPA-e031	MYLIPA-c032	3: Medium
MYLIPA-e032	MYLIPA-c033	0: None
MYLIPA-e033	MYLIPA-c034	5: High
MYLIPA-e034	MYLIPA-c035	5: High
MYLIPA-e035	MYLIPA-c036	5: High
MYLIPA-e036	MYLIPA-c037	5: High
MYLIPA-e037	MYLIPA-c038	5: High

Table A7: MYLIPA-c001 - c048 test expression results

Expression ID	Construct ID	Elution yield
MYLIPA-e038	MYLIPA-c039	3: Medium
MYLIPA-e039	MYLIPA-c040	5: High
MYLIPA-e040	MYLIPA-c041	0: None
MYLIPA-e041	MYLIPA-c042	1: Low
MYLIPA-e042	MYLIPA-c043	0: None
MYLIPA-e043	MYLIPA-c044	1: Low
MYLIPA-e044	MYLIPA-c045	0: None
MYLIPA-e045	MYLIPA-c046	5: High
MYLIPA-e046	MYLIPA-c047	0: None
MYLIPA-e047	MYLIPA-c048	5: High



Figure A5: OF-5-VAR PCR#1 results: Pre-cast 1% agarose gel (Invitrogen E-Gel 96). MYLIPA-c049: A01, MYLIPA-c052: A04, MYLIPA-c053: A05, MYLIPA-c054: A06

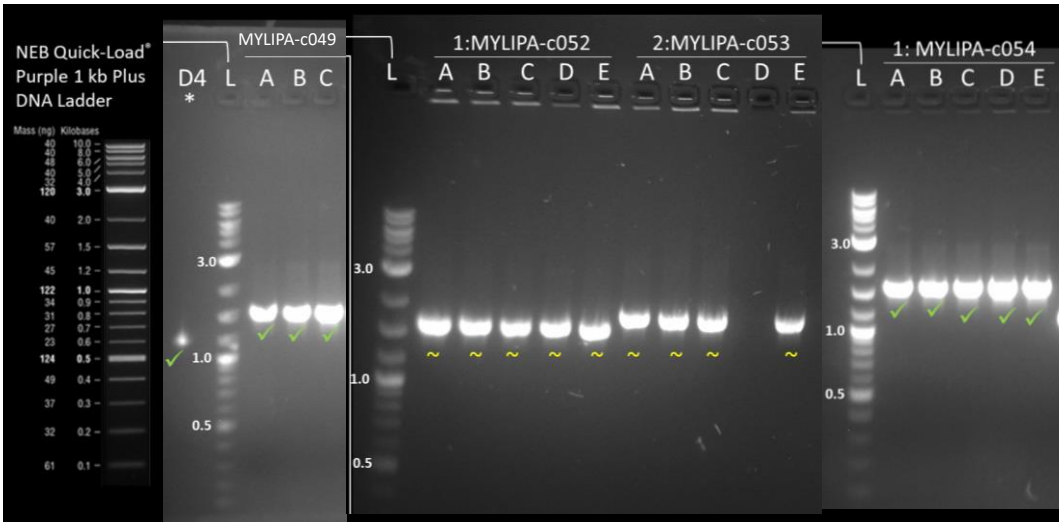


Figure A6: Composite image of Colony screen PCR agarose gel (1.5%) for MYLIPA-c049, MYLIPA-c052, MYLIPA-c053, MYLIP-c054.

Full scale MYLIP expressions and purifications

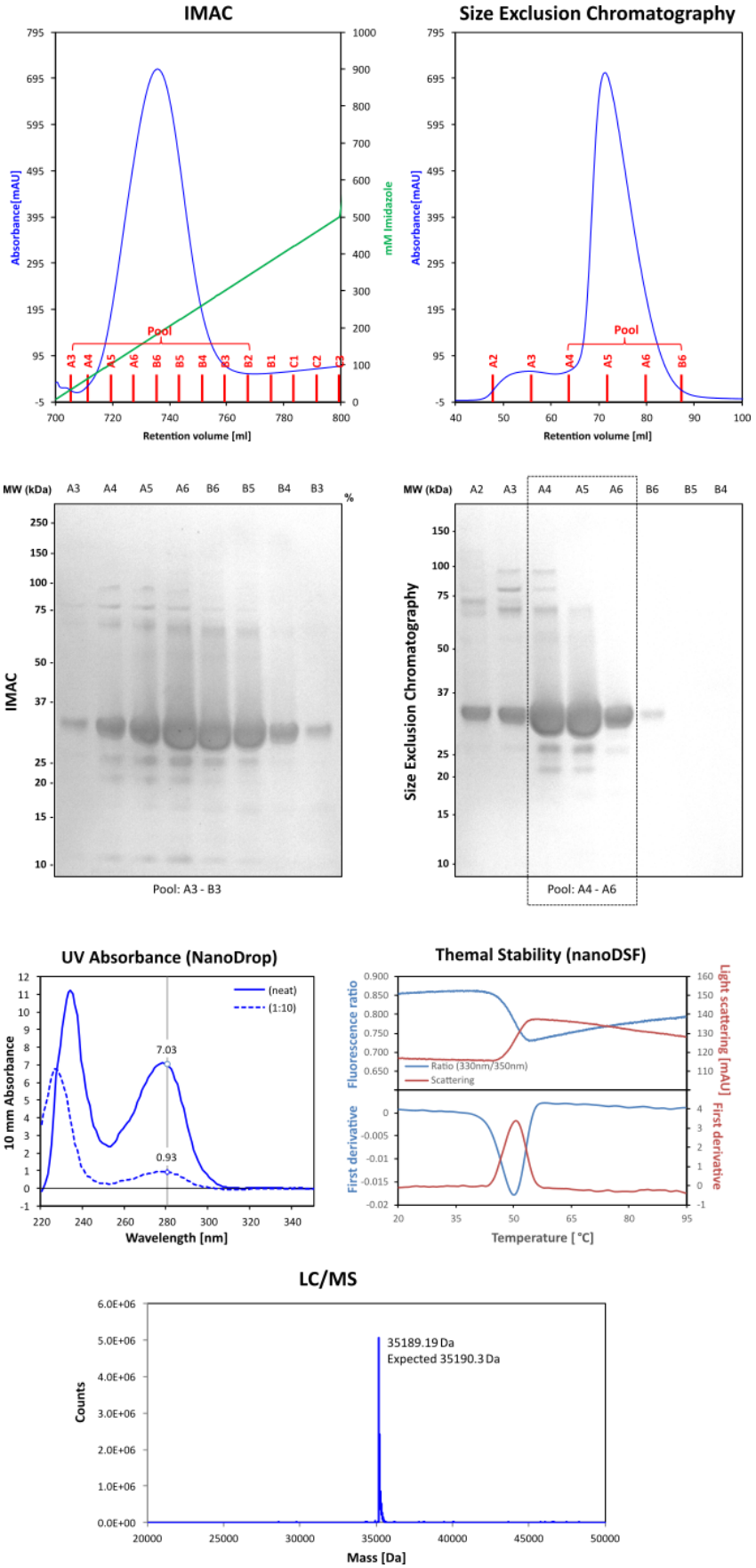


Figure A7: Purification details of MYLIPA-p003(c016:M1-S298;pNIC-CTH0).

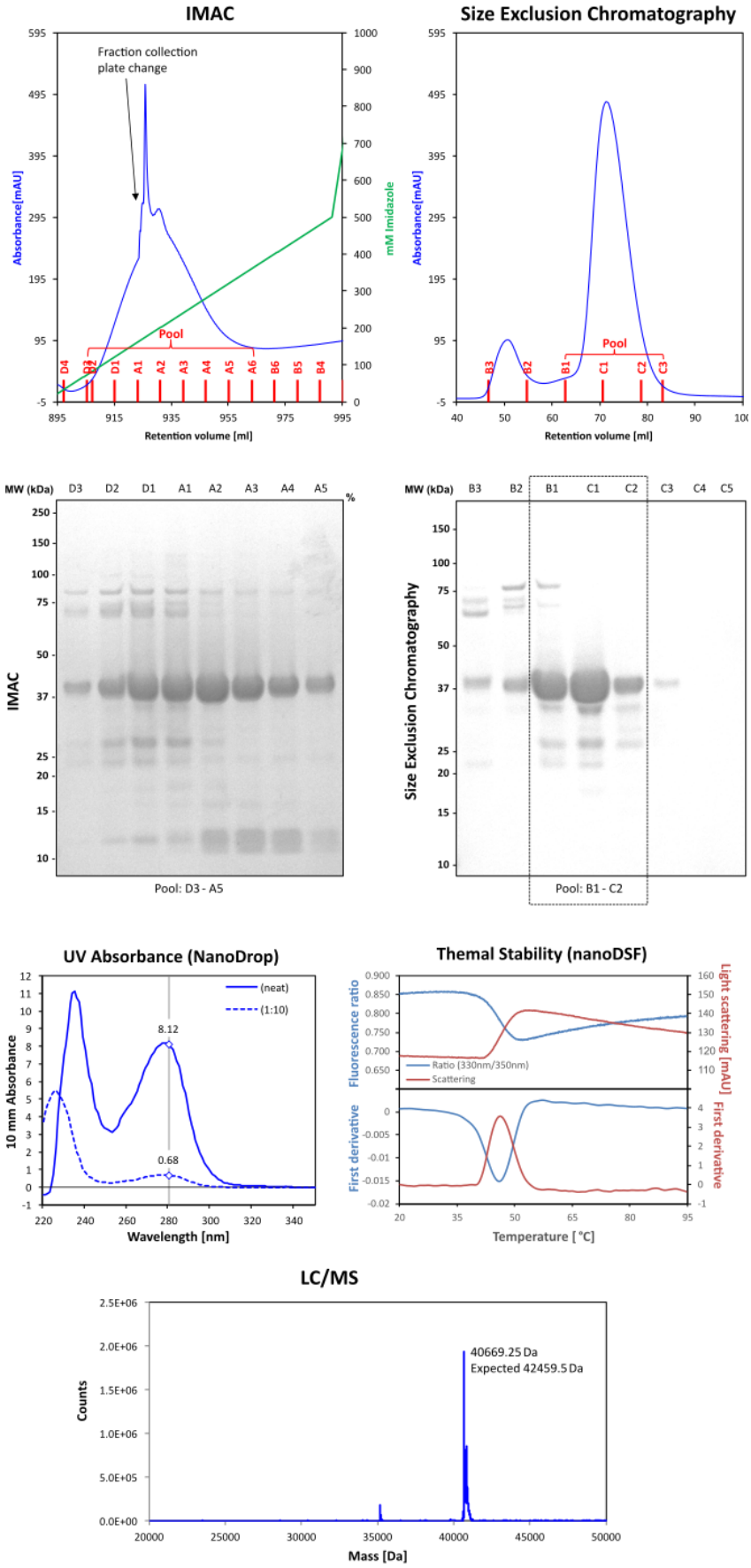


Figure A8: Purification details of MYLIPA-p004(c028:M1-S298,pNIC-ZB). Observed mass corresponds to C-terminal truncation (last 16 amino acids) (ΔMass = 1790.65 Da)

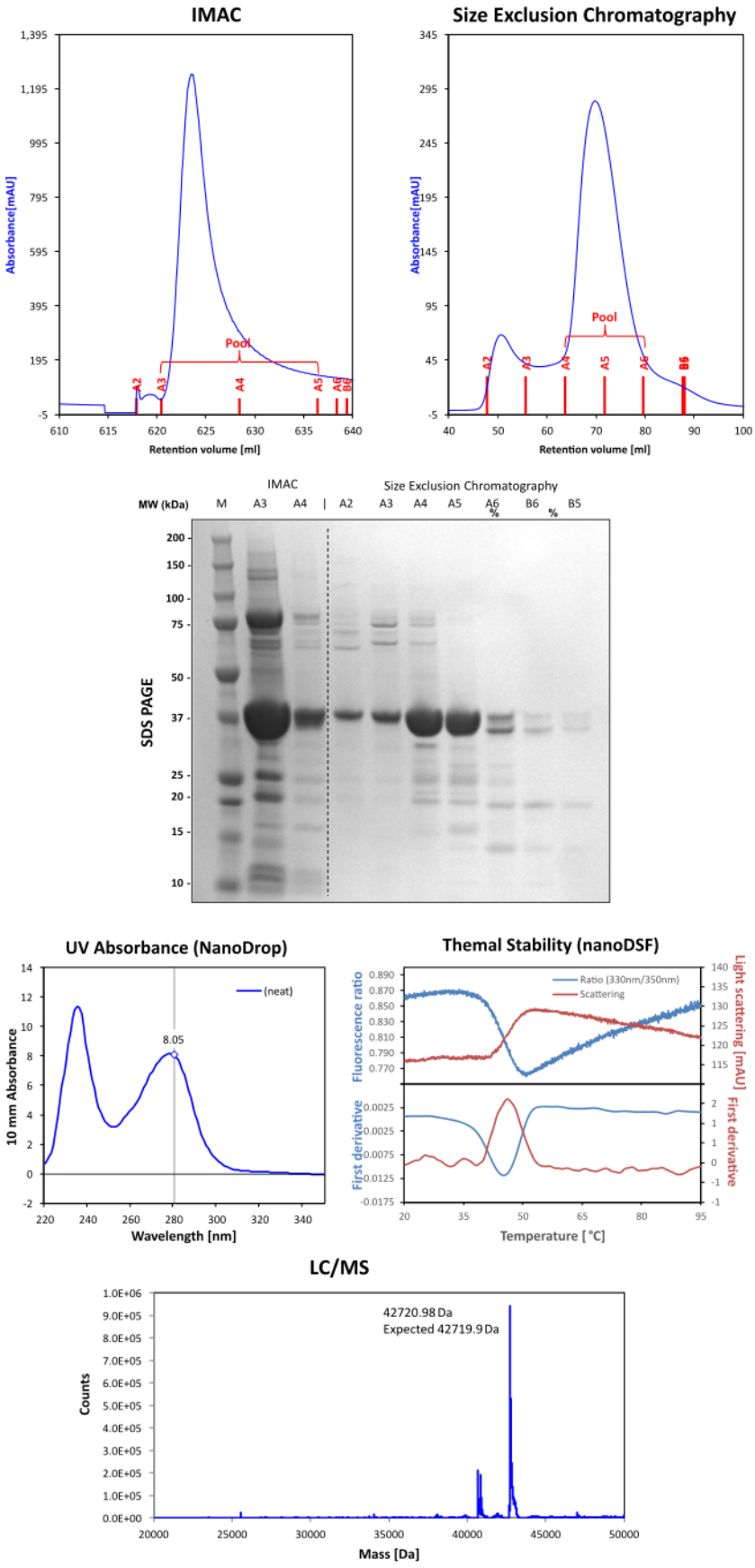


Figure A9: Purification details of MYLIPA-p005(c029:M1-F300;pNIC-ZB).

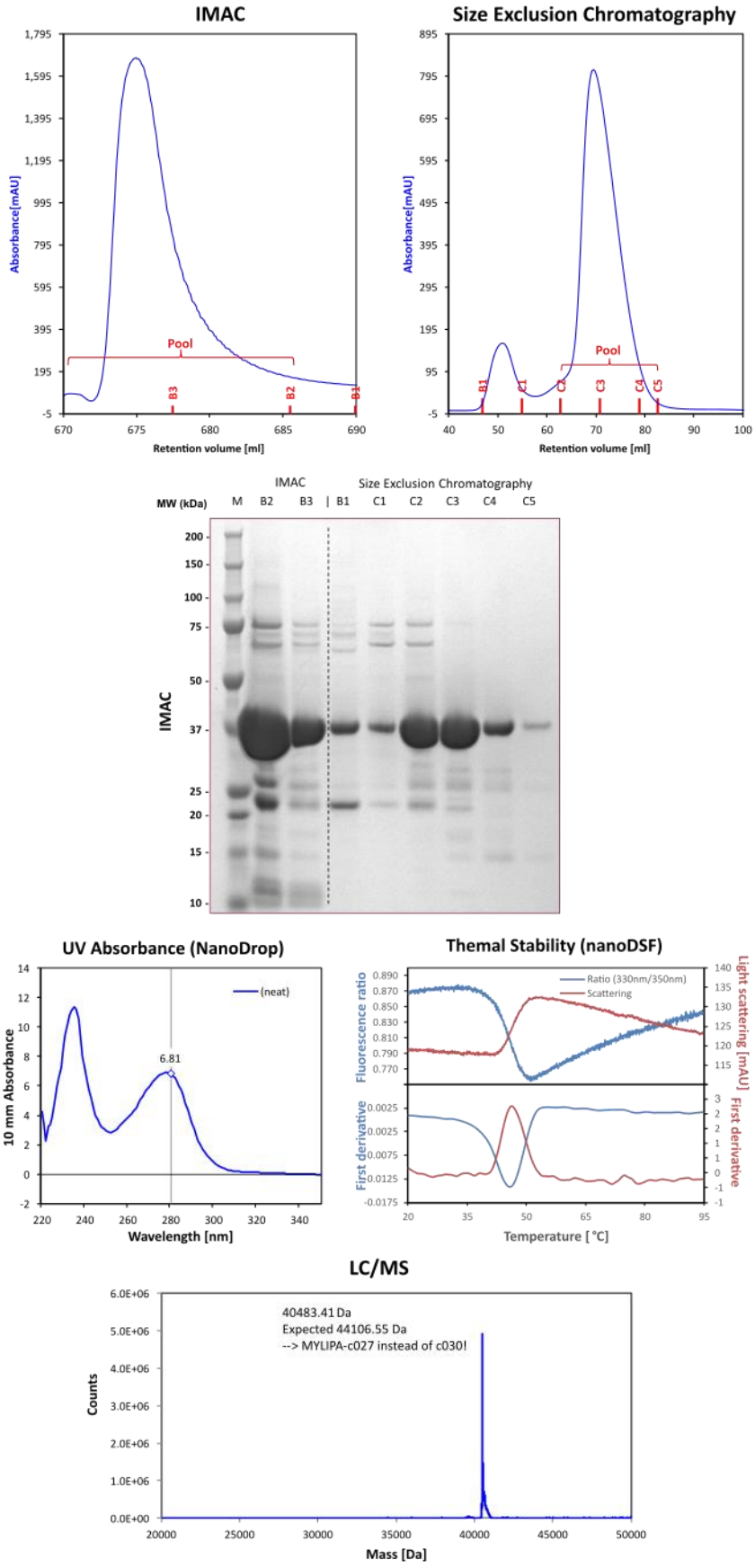


Figure A10: Purification details of MYLIPA-p006(c027:M1-V280;pNIC-ZB).

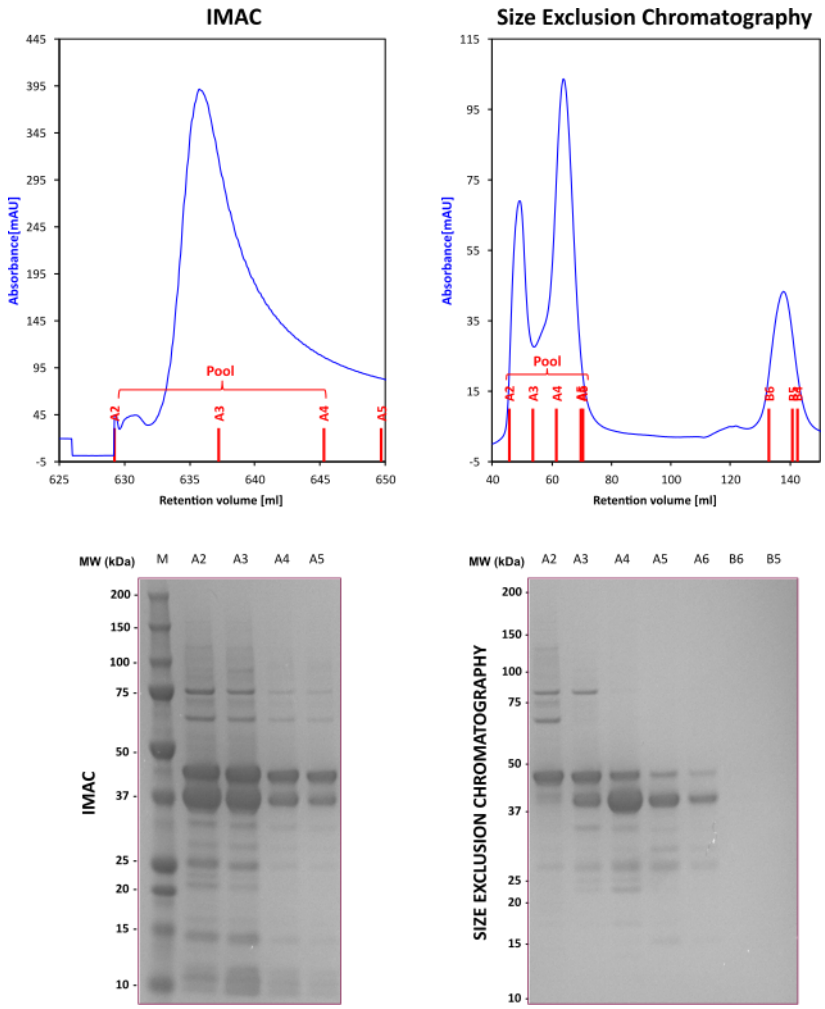


Figure A11: Purification details for MYLIPA-p009(c035: M1-V335; pNIC-ZB). Intact mass and thermal stability were not determined due to precipitation after His_6-tag cleavage.

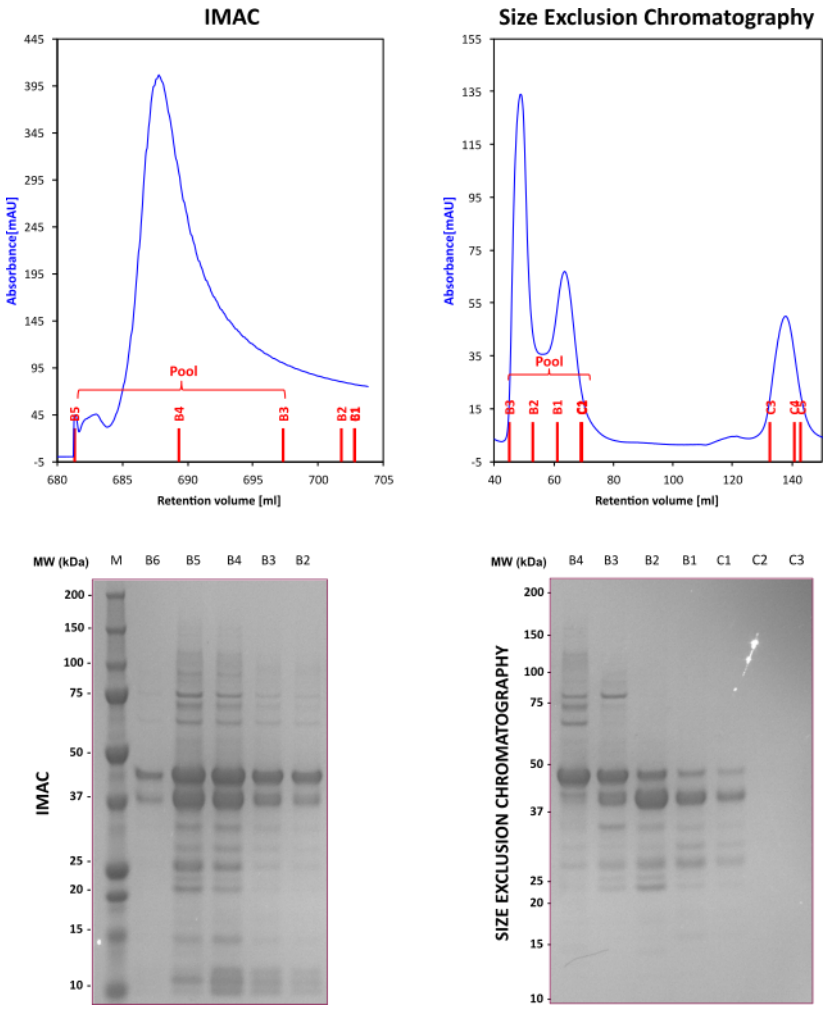


Figure A12: Purification details for MYLIPA-p010(c036:M1-L338;pNIC-ZB). Intact mass and thermal stability were not determined due to precipitation after His_6-tag cleavage.

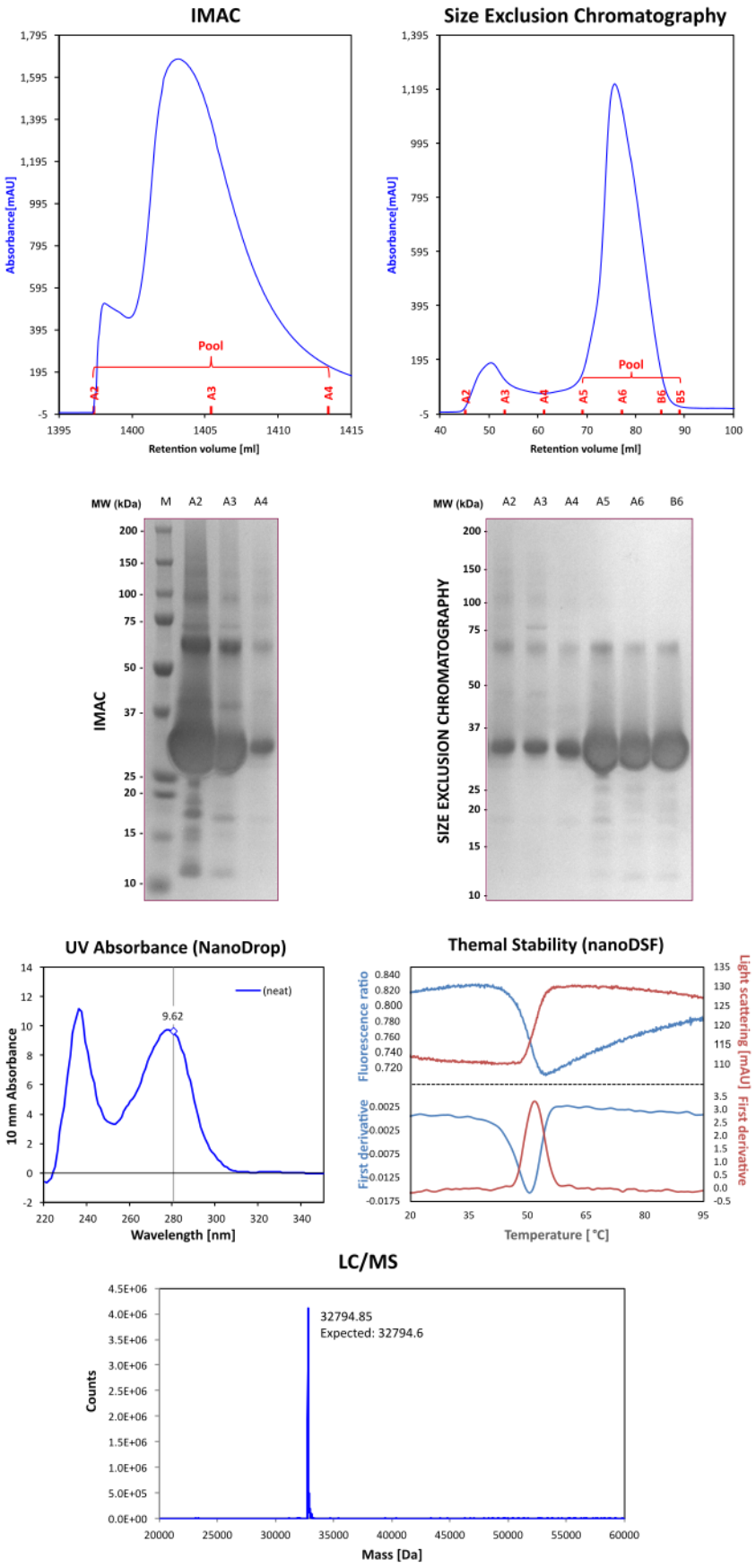


Figure A13: Purification details for MYLIPA-p011(c014:M1-R276;pNIC-CTH0).

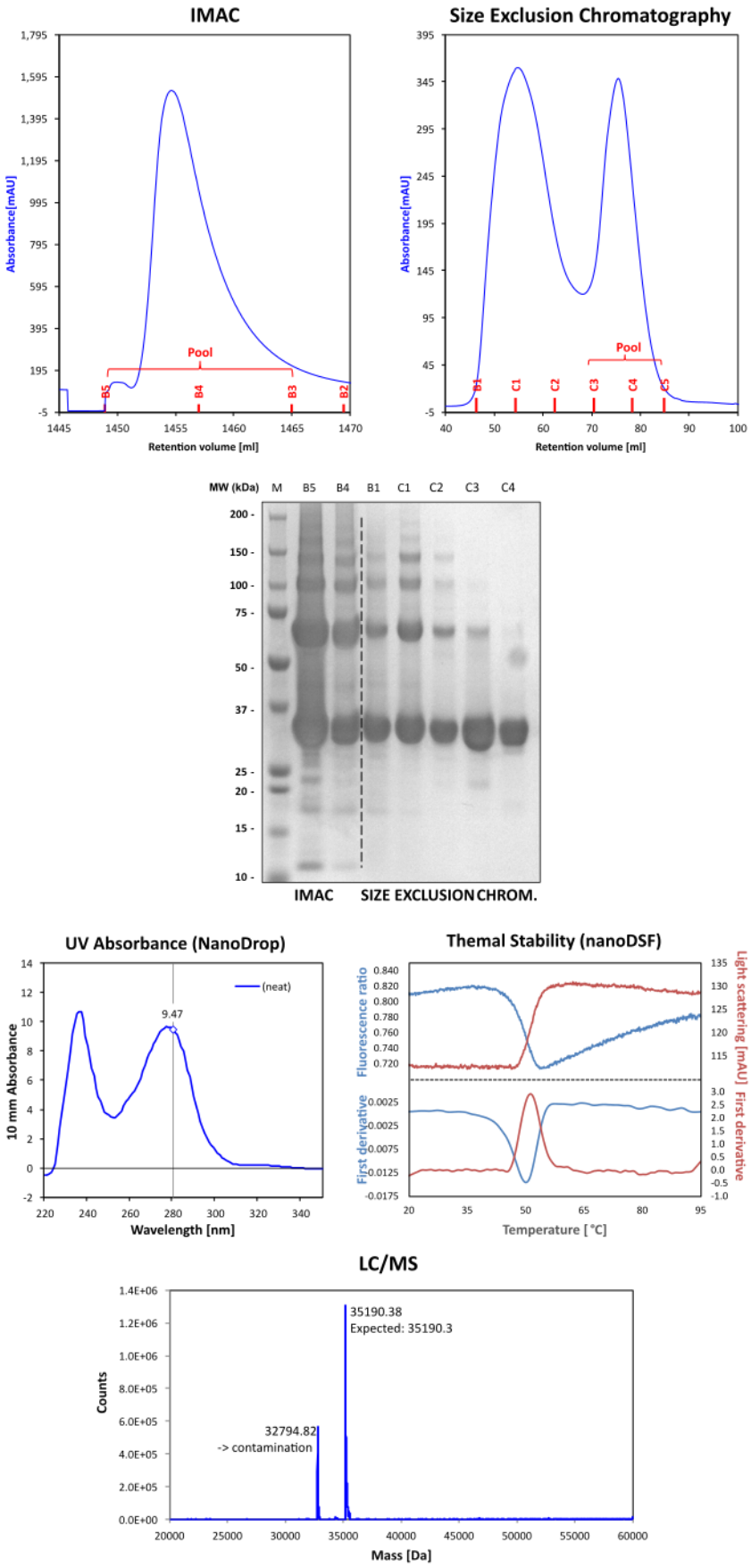


Figure A14: Purification details for MYLIPA-p012(c016;M1-S298;pNIC-CTH0).

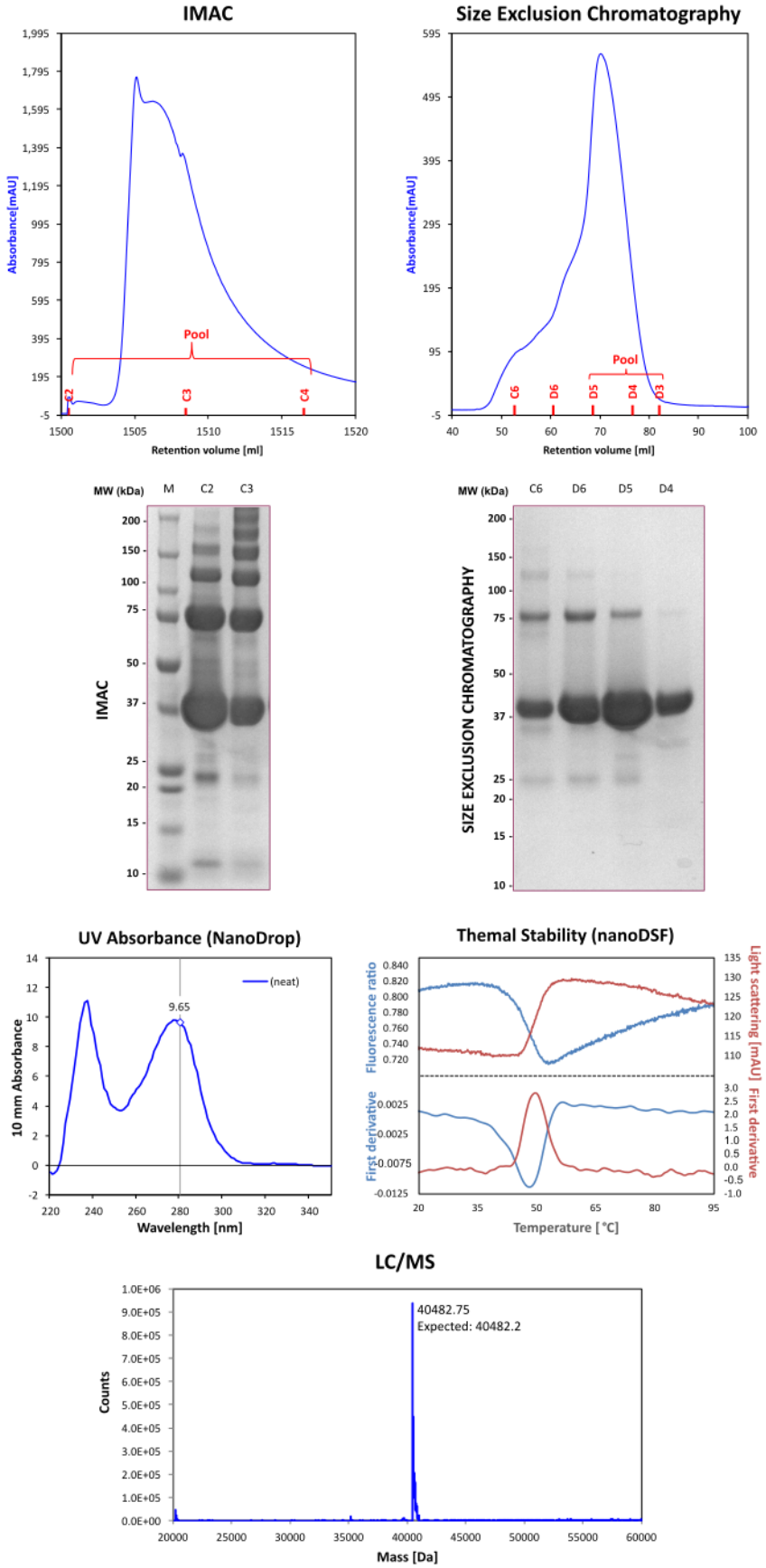


Figure A15: Purification details for MYLIPA-p013(c027;M1-V280;pNIC-ZB).

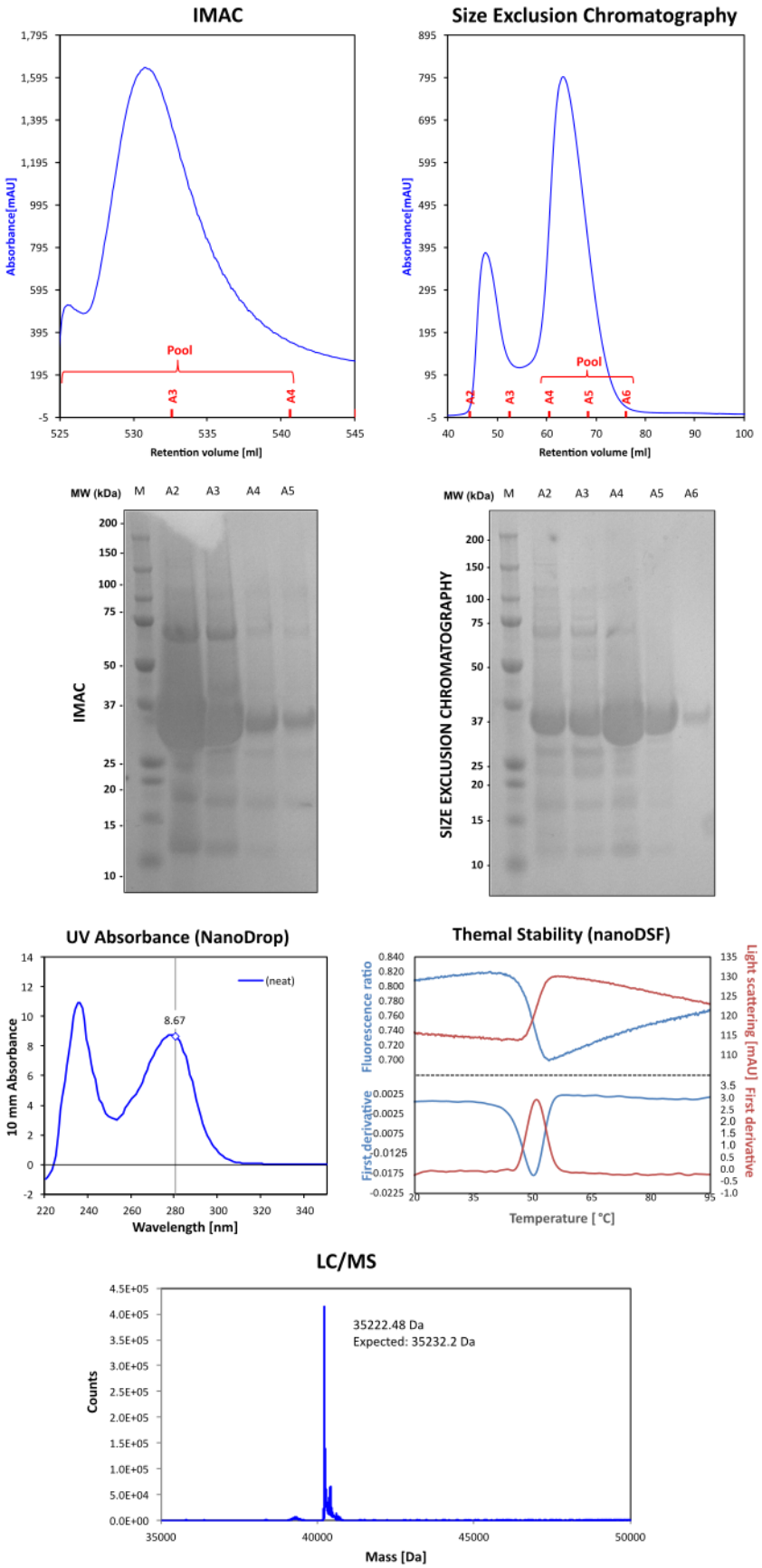


Figure A16: Purification details for MYLIPA-p018(c088:M1-S298;pNIC-CTH0).

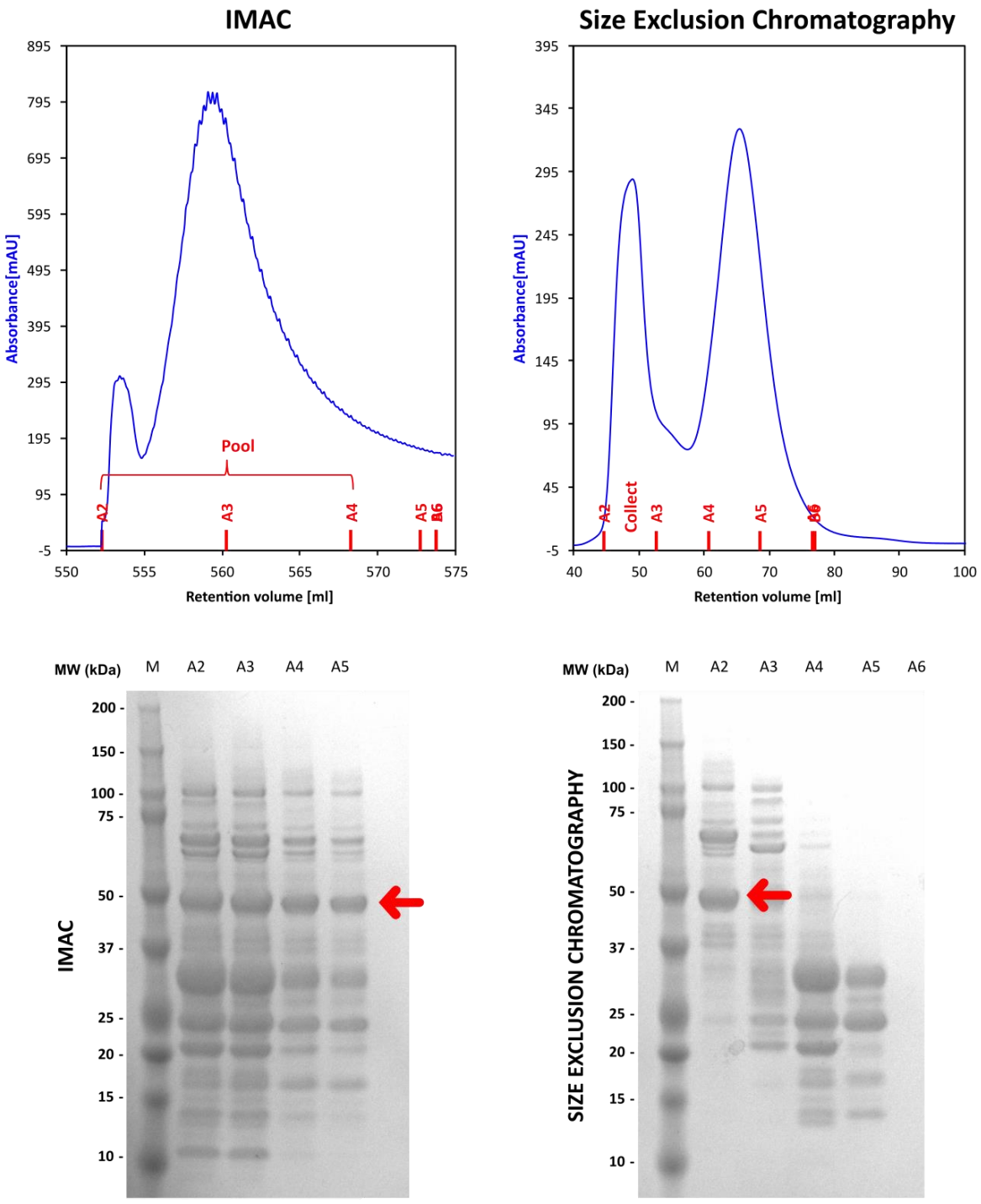


Figure A17: Purification details for MYLIPA-p022(c049:M1-I445;pNIC28-Bsa4).

LDLR adapter protein constructs

Table A8: List of LDLR adapter PTB domain constructs

Construct ID	Boundaries	Expected mass (Da)		Vector	Forward Primer	Reverse Primer	Status	Cloner
		(tagged)	(no tag)					
DAB1A-c001	23-174	19734.7	17269.0	pNIC28-Bsa4	DAB1A-f004	DAB1A-r004	Cloned	CG
DAB1A-c002	23-174	18957.8	18047.9	pNIC-CTH0	DAB1A-f003	DAB1A-r005	Cloned	CG
DAB1A-c003	23-174	22125.3	19659.6	pNIC-Bio3	DAB1A-f004	DAB1A-r006		
DAB1A-c004	23-174	22900	20434.3	pNIC-Bio3-LL	DAB1A-f004	DAB1A-r006		
DAB1A-c005	20-175	20275.3	17809.7	pNIC28-Bsa4	DAB1A-f001	DAB1A-r002		
DAB1A-c006	20-175	19498.5	18588.5	pNIC-CTH0	DAB1A-f002	DAB1A-r001		
DAB1A-c007	20-175	22665.9	20200.3	pNIC-Bio3	DAB1A-f001	DAB1A-r003		
DAB1A-c008	20-175	23440.6	20975.0	pNIC-Bio3-LL	DAB1A-f001	DAB1A-r003		
DAB2A-c001	33-181	19464.3	16998.6	pNIC28-Bsa4	DAB2A-f001	DAB2A-r005	Cloned	CG
DAB2A-c002	33-181	18687.4	17777.5	pNIC-CTH0	DAB2A-f002	DAB2A-r006	Cloned	CG
DAB2A-c003	33-181	21854.9	19389.2	pNIC-Bio3	DAB2A-f001	DAB2A-r004		
DAB2A-c004	33-181	22629.6	20163.9	pNIC-Bio3-LL	DAB2A-f001	DAB2A-r004		
DAB2A-c005	33-191	20535.4	18069.6	pNIC28-Bsa4	DAB2A-f001	DAB2A-r002		
DAB2A-c006	33-191	19758.5	18848.6	pNIC-CTH0	DAB2A-f002	DAB2A-r001		
DAB2A-c007	33-191	22926	20460.3	pNIC-Bio3	DAB2A-f001	DAB2A-r003		
DAB2A-c008	33-191	23700.7	21235.1	pNIC-Bio3-LL	DAB2A-f001	DAB2A-r003		
LDLRAP1A-c001	44-175	14960.6	17426.3	pNIC28-Bsa4	LDLRAP1A-f003	LDLRAP1A-r003		

LDLRAP1A-c002	4-175	15739.5	16649.4	pNIC-CTH0	LDLRAP1A-f004	LDLRAP1A-r002
LDLRAP1A-c003	4-175	17351.2	19816.9	pNIC-Bio3	LDLRAP1A-f003	LDLRAP1A-r001
LDLRAP1A-c004	4-175	18125.9	20591.6	pNIC-Bio3-LL	LDLRAP1A-f003	LDLRAP1A-r001
LDLRAP1A-c005	43-174	14986.7	17452.3	pNIC28-Bsa4	LDLRAP1A-f001	LDLRAP1A-r006
LDLRAP1A-c006	43-174	15765.6	16675.5	pNIC-CTH0	LDLRAP1A-f002	LDLRAP1A-r005
LDLRAP1A-c007	43-174	17377.3	19842.9	pNIC-Bio3	LDLRAP1A-f001	LDLRAP1A-r004
LDLRAP1A-c008	43-174	18151.99	20617.6	pNIC-Bio3-LL	LDLRAP1A-f001	LDLRAP1A-r004

CG: Charline Giroud

Table A9: DAB1 forward primers

ID	Sequence
DAB1A-f001	TACTTCCAATCCATGAGAAAGAAAGGTCAGGATCGCA
DAB1A-f002	TTAAGAAGGAGATATACTATGAGAAAGAAAGGTCAGGATCGCAGTGA
DAB1A-f003	TTAAGAAGGAGATATACTATGGGTCAGGATCGCAGTGAAGC
DAB1A-f004	TACTTCCAATCCATGGGTCAGGATCGCAGTGAA

Table A10: DAB1 reverse primers

ID	Sequence
DAB1A-r004	TATCCACCTTTACTGTCATGCCTTTTTTTCTAATTCTTCTCTTTGCTTCA
DAB1A-r005	GATTGGAAGTAGAGGTTCTCTGCTGCCTTTTTTTCTAATTCTTCTCT

Table A11: DAB2 forward primers

ID	Sequence
DAB1A-r001	GATTGGAAGTAGAGGTTCTCTGCTTGTGCCTTTTTTTCTAATTCTTC
DAB1A-r002	TATCCACCTTTACTGTCATTGTGCCTTTTTTTCTAATTCTTCTCTTTGC
DAB1A-r003	TATCCACCTTTACTGCTTTGTGCCTTTTTTTCTAATTCTTCTCTTTGC

DAB1A-r004	TATCCACCTTTACTGTCATGCCTTTTTTTCTAATTCTTCTCTTTGCTTCA
DAB1A-r005	GATTGGAAGTAGAGGTTCTCTGCTGCCTTTTTTTCTAATTCTTCTCT
DAB1A-r006	TATCCACCTTTACTGCTTGCCTTTTTTTCTAATTCTTCTCTTTGCT

Table A12: DAB2 reverse primers	
ID	**Sequence**
DAB2A-r001	GATTGGAAGTAGAGGTTCTCTGCATTCTCAACTGCTTTGCTGG
DAB2A-r002	TATCCACCTTTACTGTCAATTCTCAACTGCTTTGCTGGC
DAB2A-r003	TATCCACCTTTACTGCTATTCTCAACTGCTTTGCTGGC
DAB2A-r004	TATCCACCTTTACTGCTCTTTTTCTTTTCTTCTTCCTTTTTCTTTACATTATAG ATAACTTGAAAAAGG
DAB2A-r005	TATCCACCTTTACTGTCACTTTTTCTTTTCTTCTTCCTTTTTCTTTACATTATA GATAACTTGAAAAAGG
DAB2A-r006	GATTGGAAGTAGAGGTTCTCTGCCTTTTTCTTTTCTTCTTCCTTTTTCT

Table A13: LDLRAP1 forward primers	
ID	**Sequence**
LDLRAP1A-f001	TACTTCCAATCCATGCTGGAGGGGATGCTGTTC
LDLRAP1A-f002	TTAAGAAGGAGATATACTATGCTGGAGGGGATGCTGTTCA
LDLRAP1A-f003	TACTTCCAATCCATGGAGGGGATGCTGTTCAGC
LDLRAP1A-f004	TTAAGAAGGAGATATACTATGGAGGGGATGCTGTTCAGCC

Table A14: LDLRAP1 reverse primers	
ID	**Sequence**
LDLRAP1A-r001	TATCCACCTTTACTGCTGGACACCTGCCAAAACTC
LDLRAP1A-r002	GATTGGAAGTAGAGGTTCTCTGCGGACACCTGCCAAAACTC
LDLRAP1A-r003	TATCCACCTTTACTGTCAGGACACCTGCCAAAACTC
LDLRAP1A-r004	TATCCACCTTTACTGCTCACCTGCCAAAACTCAAAGG
LDLRAP1A-r005	GATTGGAAGTAGAGGTTCTCTGCCACCTGCCAAAACTCAAAG
LDLRAP1A-r006	TATCCACCTTTACTGTCACACCTGCCAAAACTCAAAGG
LDLRAP1A-r001	TATCCACCTTTACTGCTGGACACCTGCCAAAACTC

LDLR adapter proteins purification

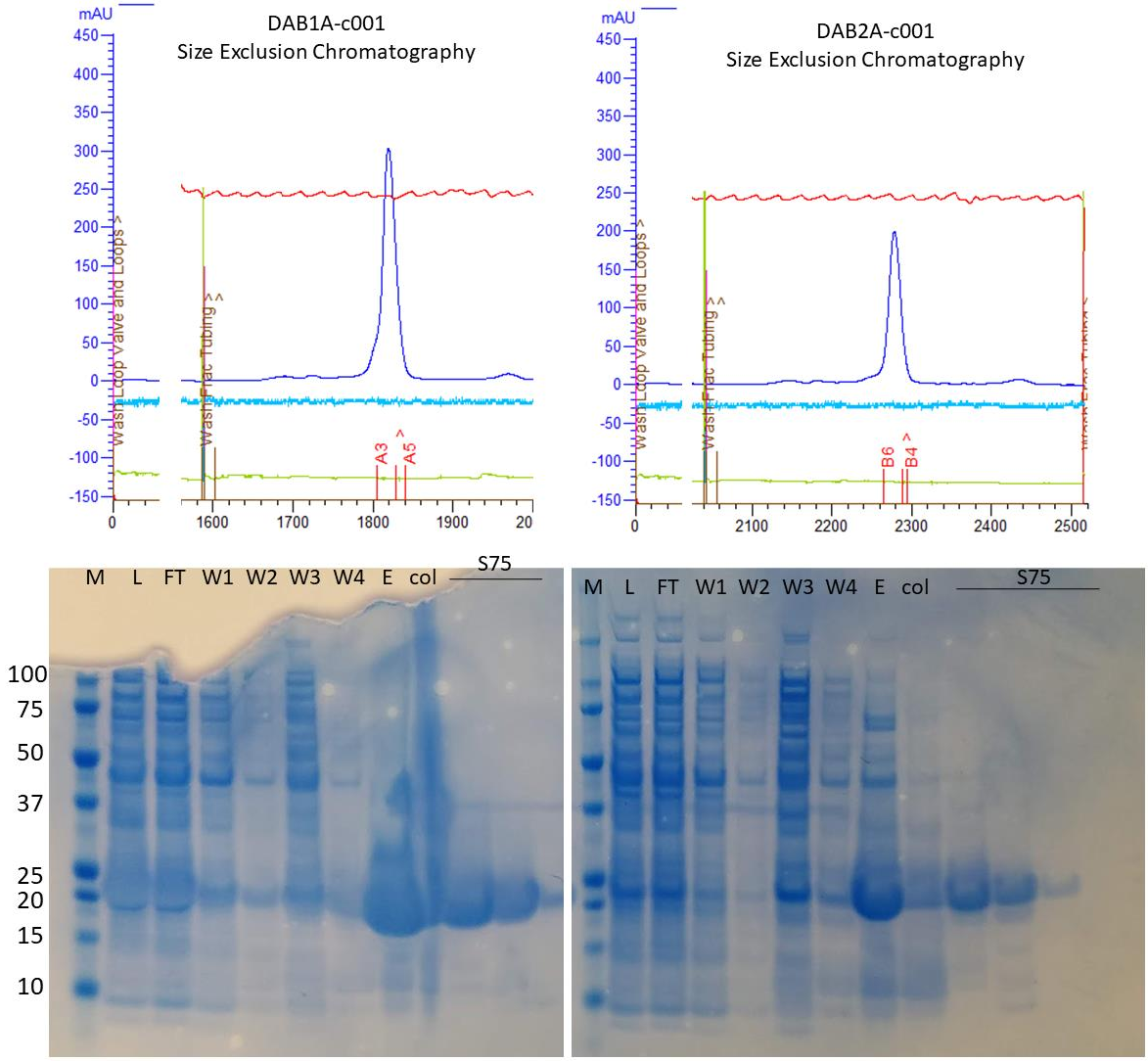


Figure A18: Purification details for DAB1A-p001 and DAB2-p001. **Top:** Size exclusion chromatography of pooled fractions after IMAC elution. **Bottom:** SDS-PAGE of IMAC and SEC(S75) fractions. M: Marker. L: Lysate. F: Flow Through. W1 – W3: Washes. E: Elution. col: Column clean with 2 M imidazole, p

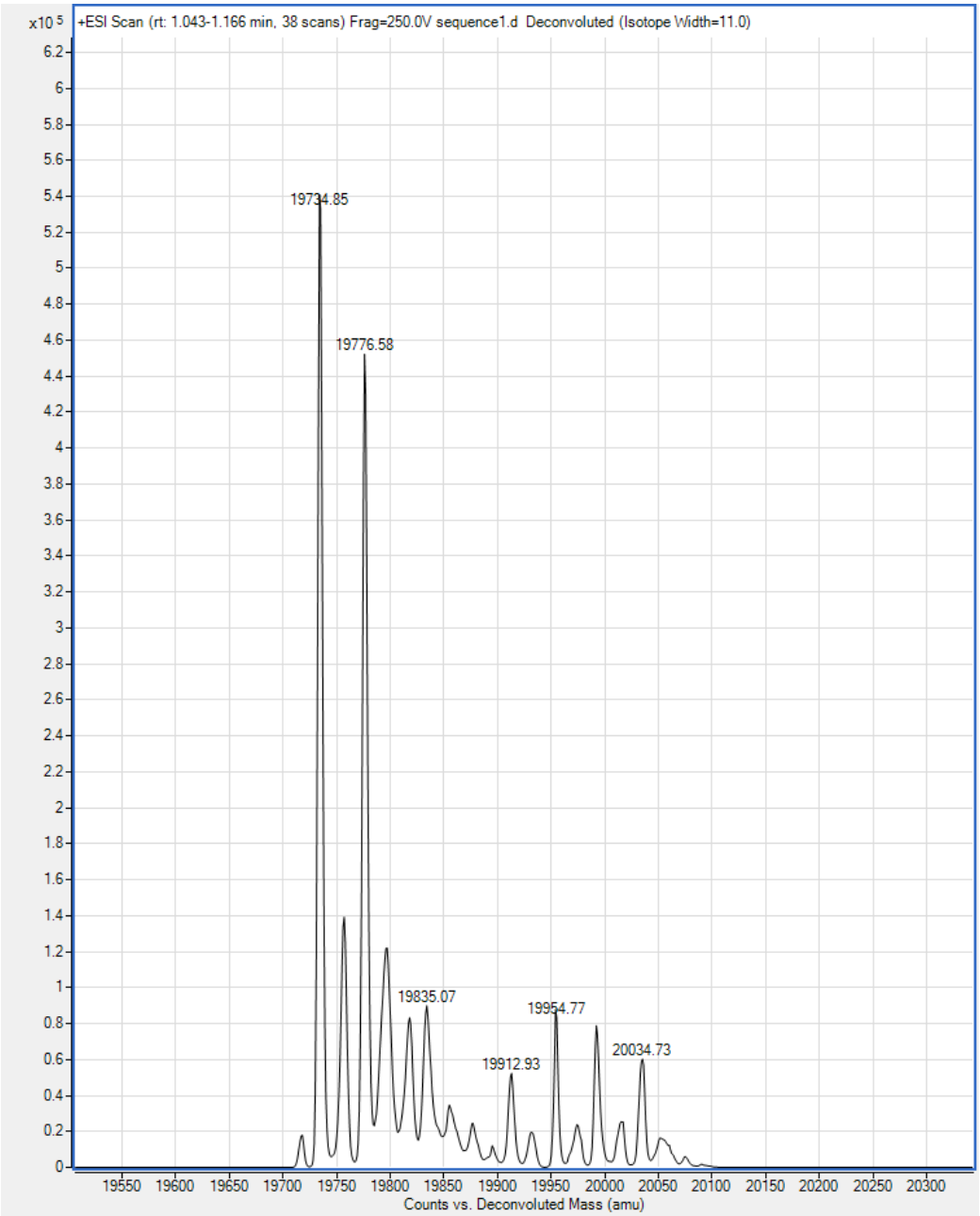


Figure A 19: Deconvoluted LC/MS spectrum for DAB1A-p001. Expected mass was 19734.7 Da, observed masses were 19734.84 Da and 19776.58 Da.

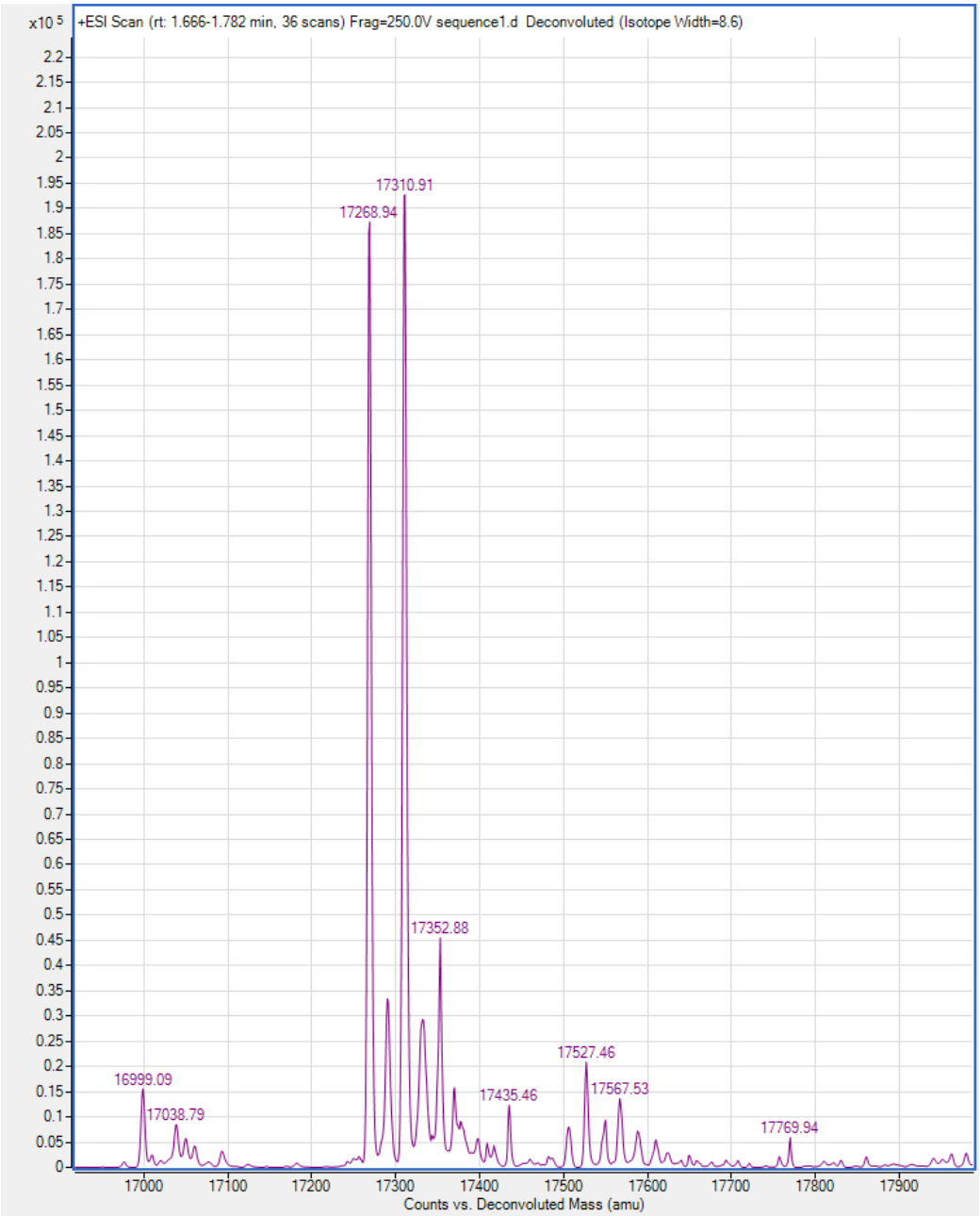


Figure A20: Deconvoluted LC/MS spectrum for DAB1A-p002. Expected mass was 17269.0 Da, observed masses are 17268.94 Da and 17310.91 Da.

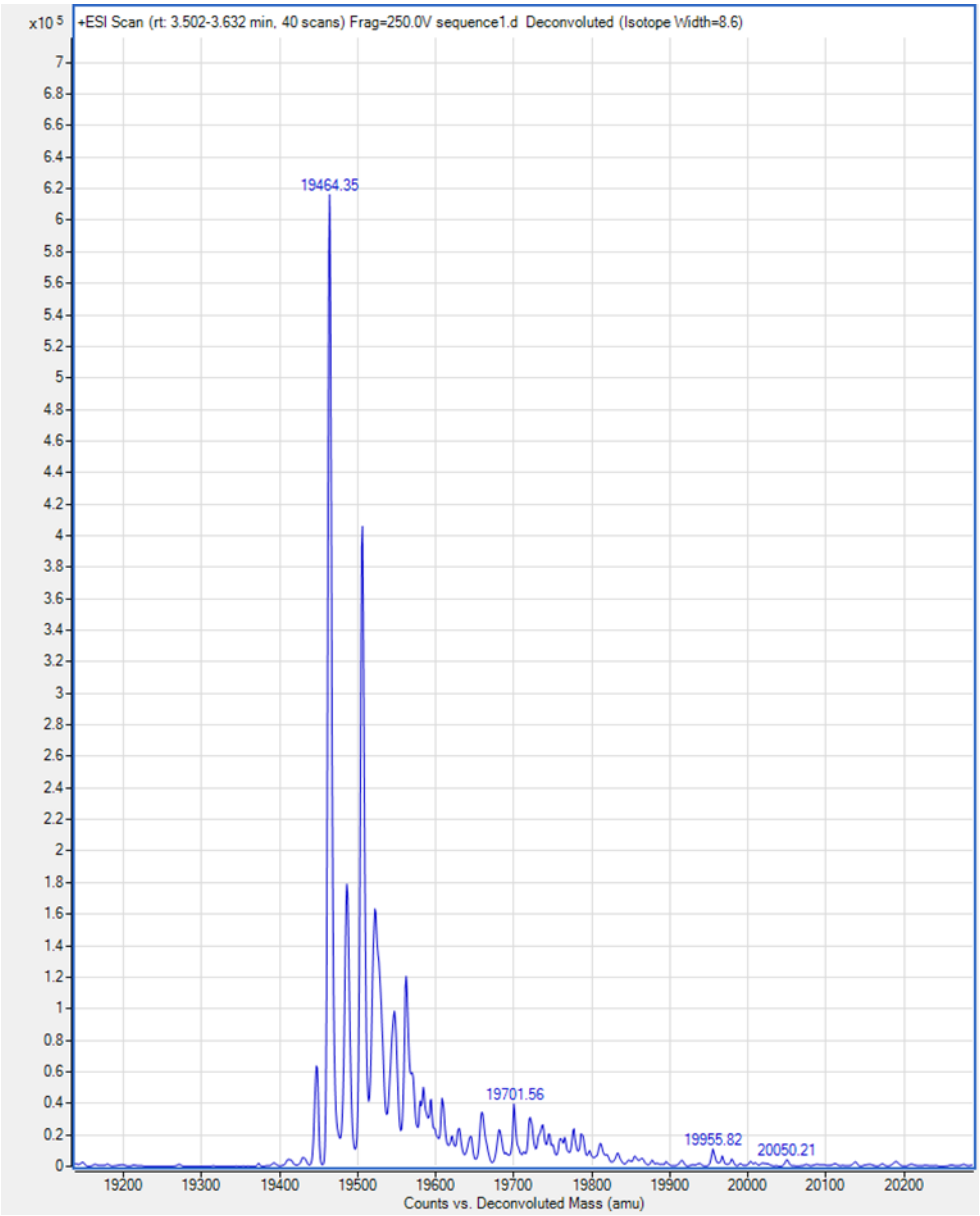


Figure A21: Deconvoluted LC/MS spectrum for DAB2A-p001. Expected mass was 19464.3 Da, observed was 19464.35 Da.

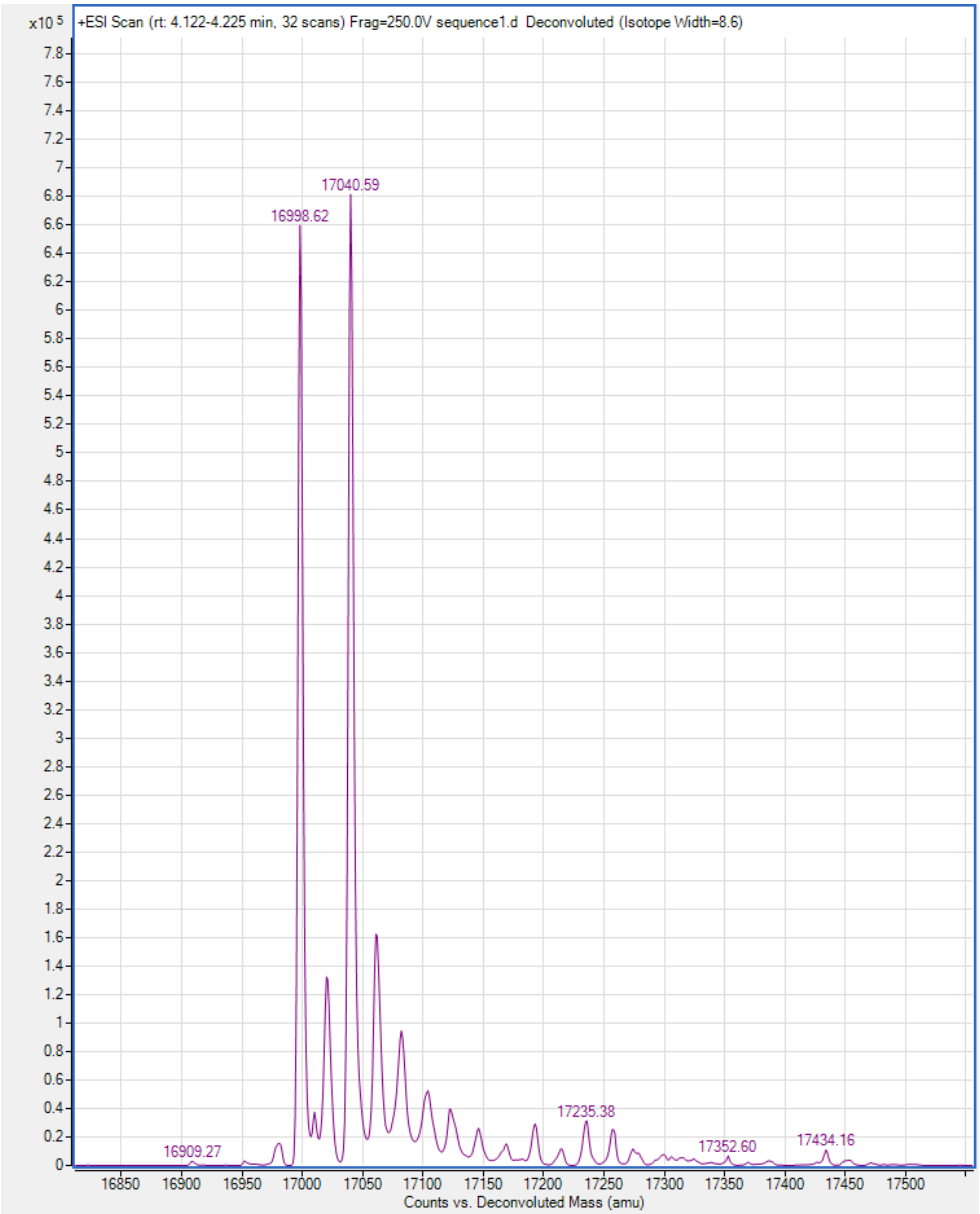


Figure A22: Deconvoluted LC/MS spectrum for DAB2A-p002. Expected mass was 16998.6 Da, observed masses were 16998.62 Da and 17040.59 Da.

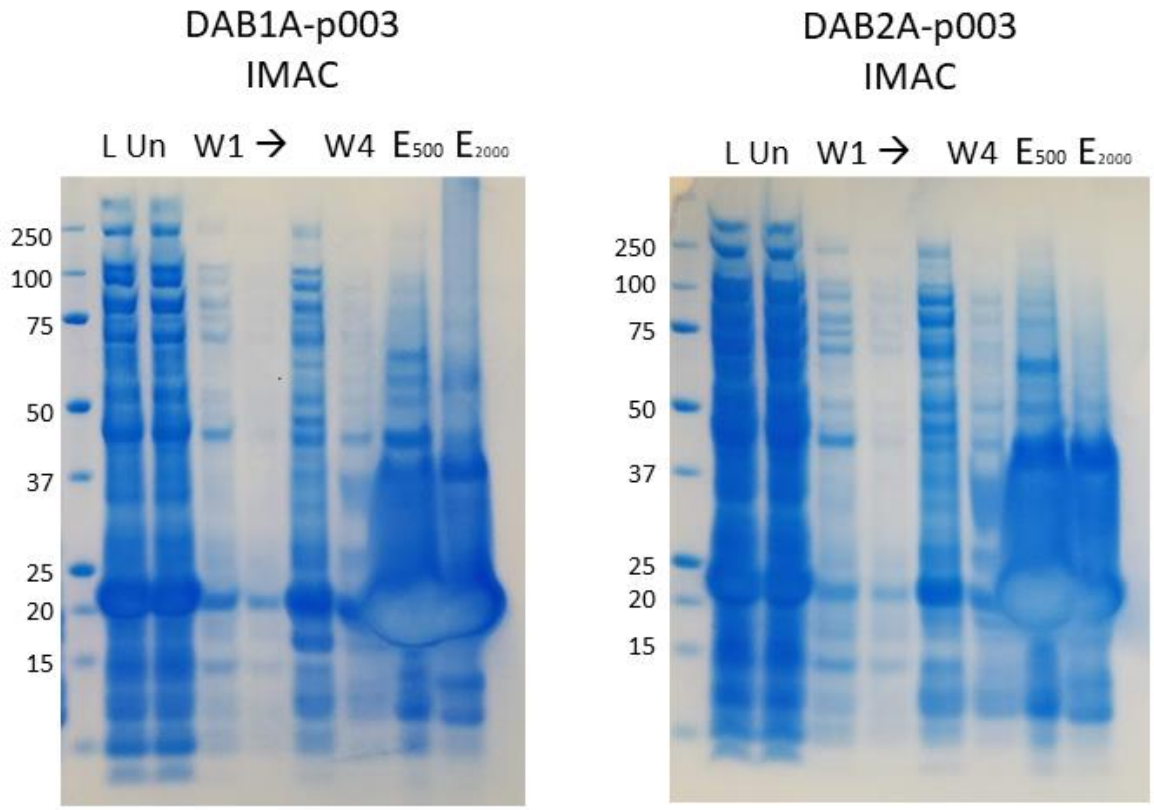


Figure A23: Purification details DAB1A-p003 and DAB2A-p003. SDS-PAGE after IMAC (**L**: Lysate. **Un**: Unbound fraction/flow through. **W1**->**W4**: Washes. **E**: Elution with imidazole concentration in mM.

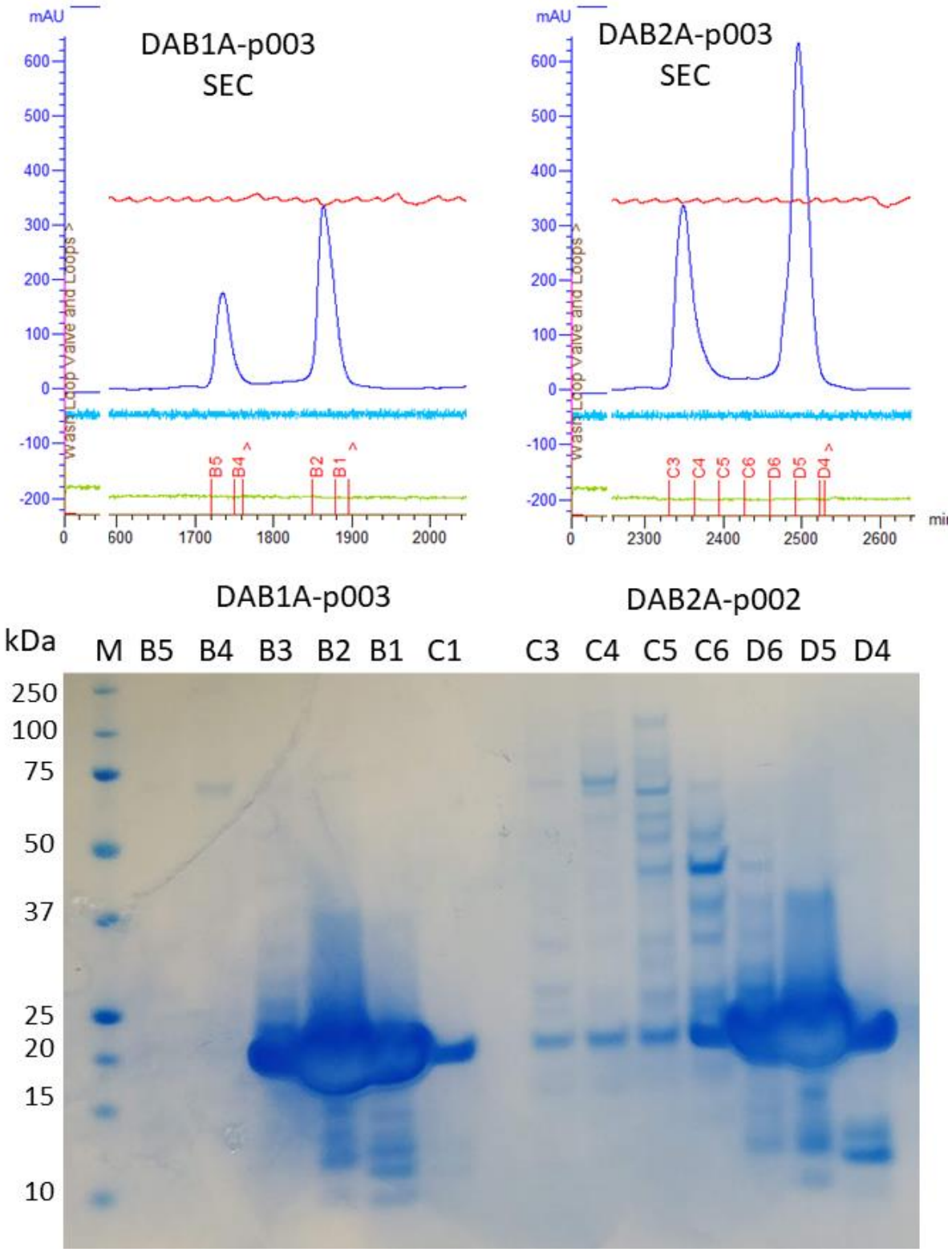


Figure A24: Purification details DAB1A-p003 and DAB2A-p003. **Top row:** SEC chromatograms. Bottom: SDS-PAGE of fractions collected from SEC. Fractions B2, B1, C1 were pooled for DAB1A-p003. Fractions D6, D5, D4 were pooled for DAB2A-p003.

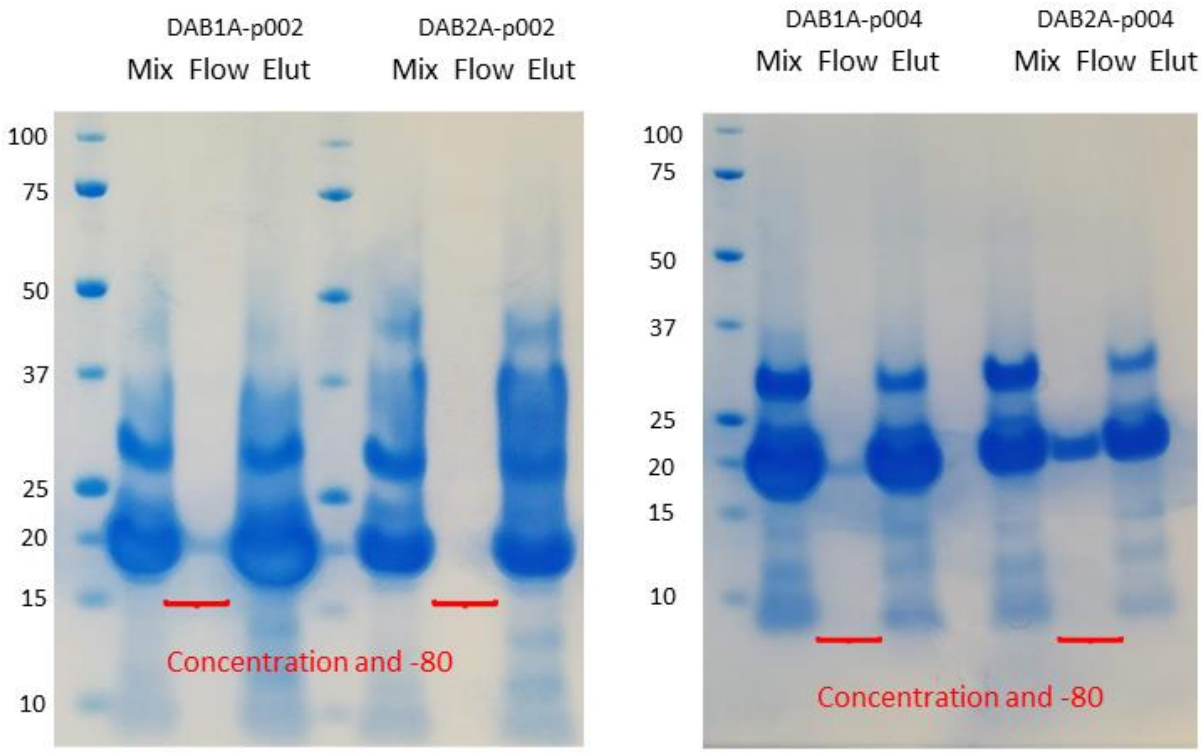


Figure A25: PTB domains SDS-PAGE after TEV digest. Mix: protein + Tev. Flow: Unbound fraction/flow through after reverse IMAC. Elut: Eluted fraction after reverse IMAC.

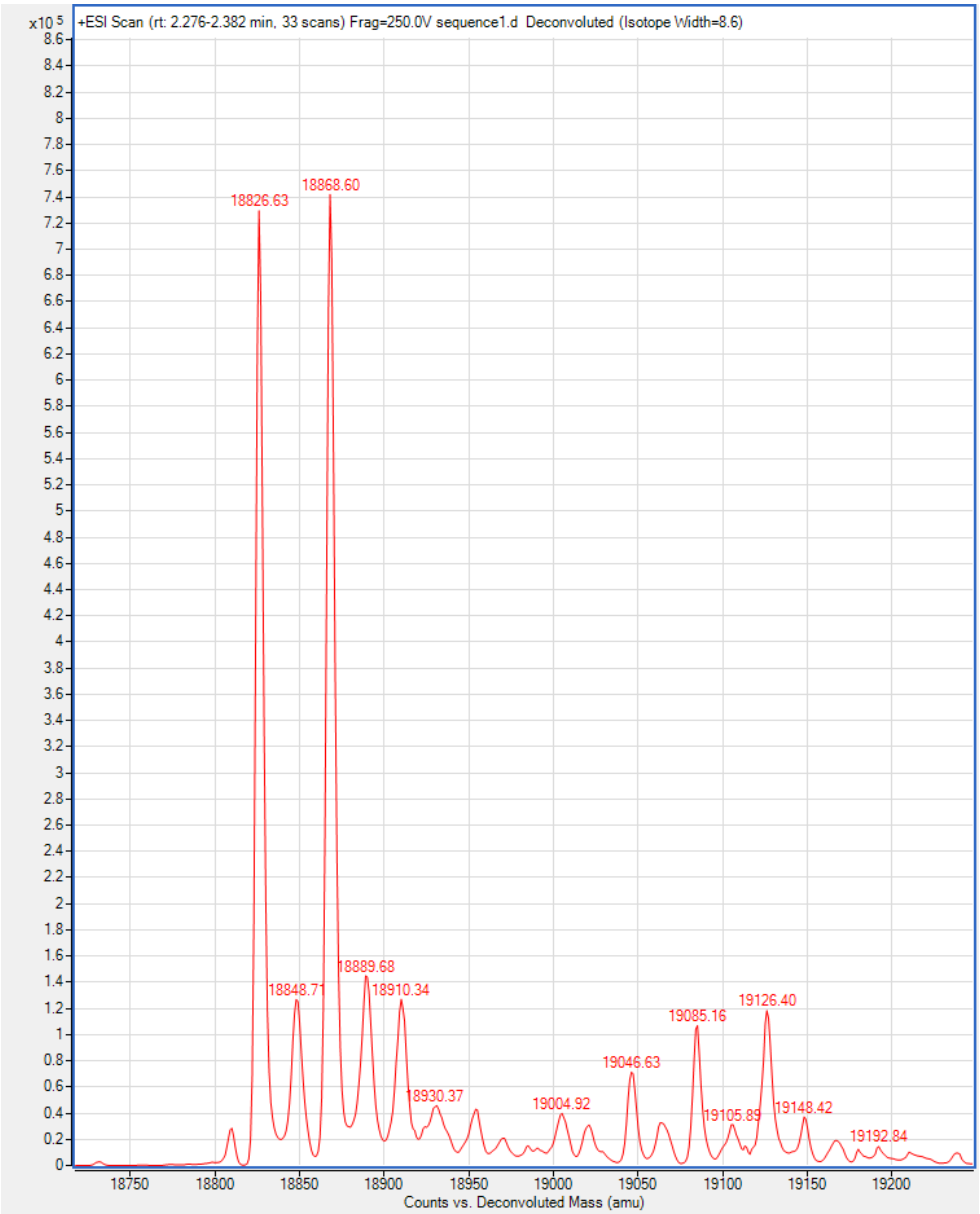


Figure A26: Deconvoluted LC/MS spectrum for DAB1A-p003. Expected mass was 18957.8 Da. Observed masses were 18826.63 Da (consistent with N-terminal methionine loss) and 18868.6 Da.

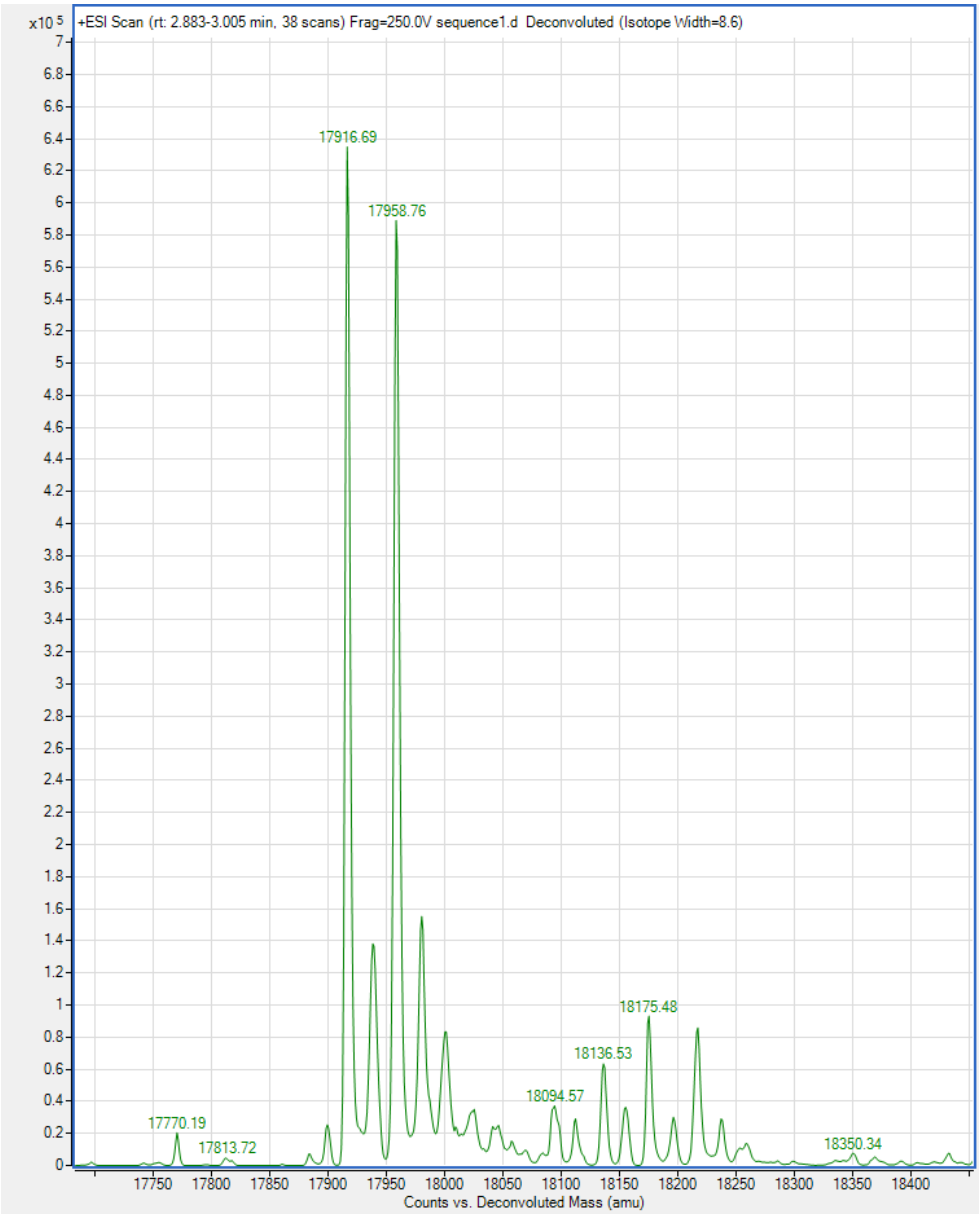


Figure A27: Deconvoluted LC/MS spectrum for DAB1A-p004. Expected mass was 18047.9 Da. Observed masses were 17916.69 Da (consistent with N-terminal methionine loss) and 17958.76 Da.

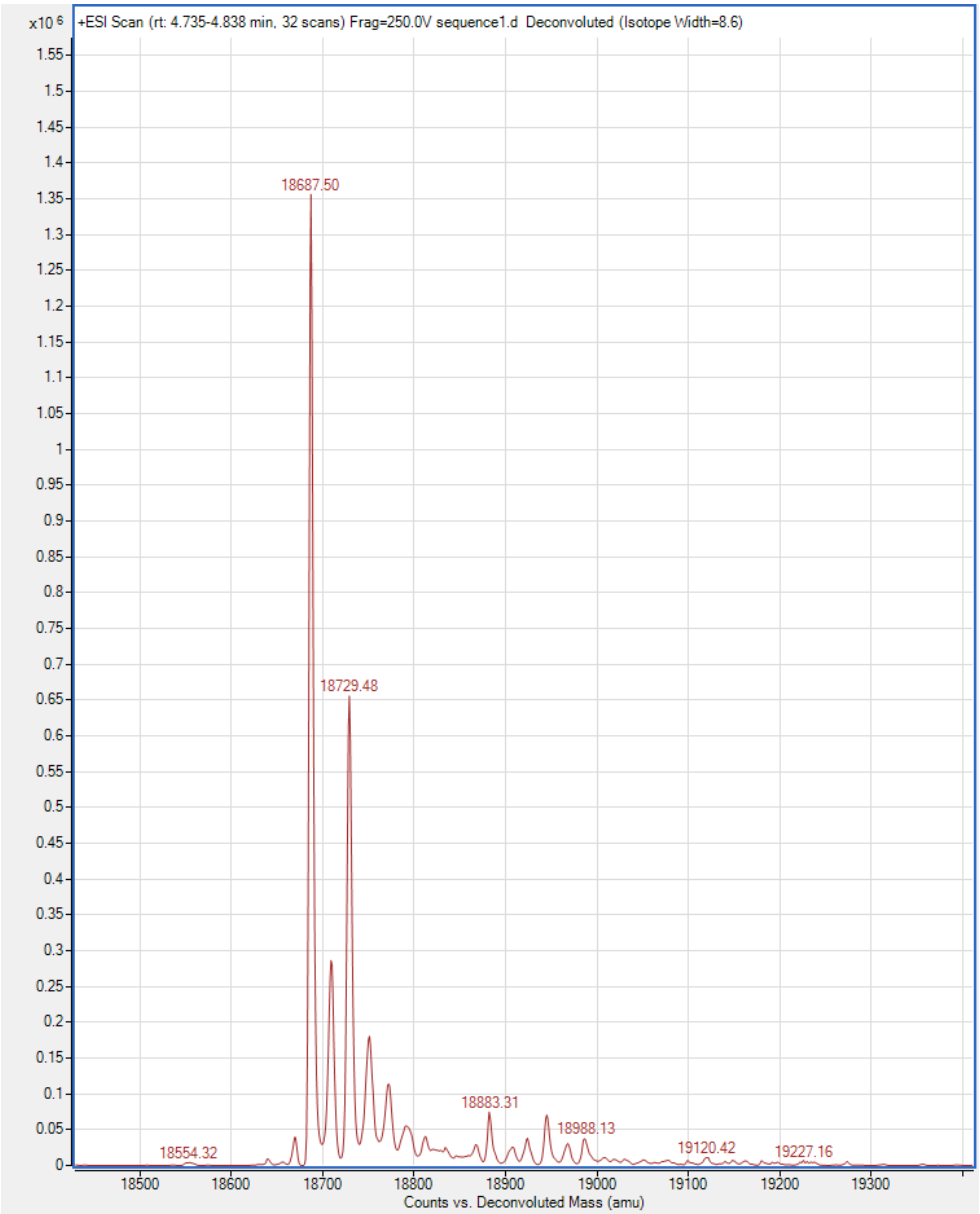


Figure A28: Deconvoluted LC/MS spectrum for DAB2A-p003. Expected masses were 18687.4 Da. Observed mass was 18687.5 Da and 18729.48 Da.

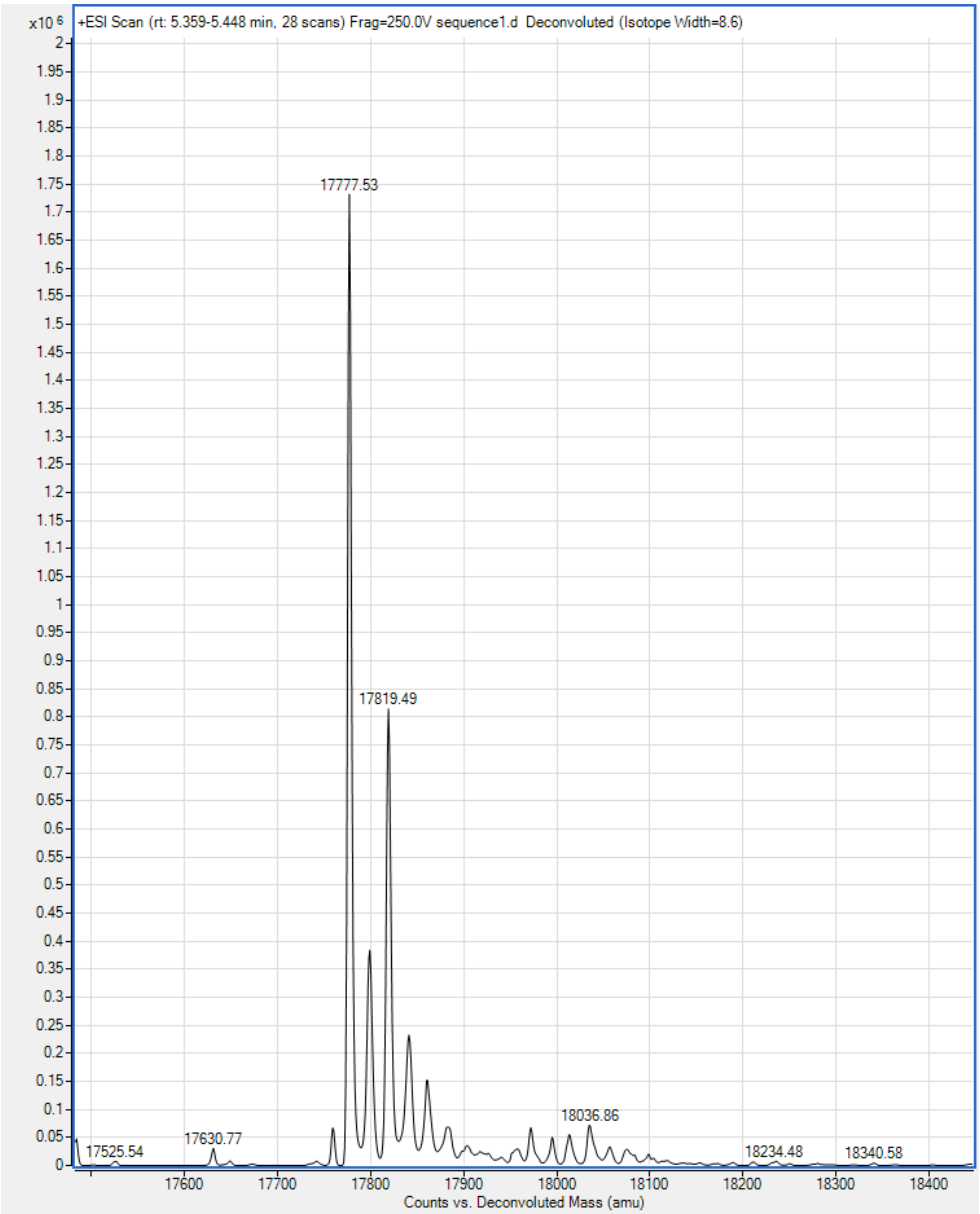


Figure A29: Deconvoluted LC/MS spectrum for DAB2A-p004. Expected mass was 17777.5 Da. Observed masses were 17777.53 Da and 17819.49 Da.

Phospho-mimetic MYLIP mutants

Site directed mutagenesis primers:

5'-ctatactgcatcatcacggc gtc ggt cac atcgtcacacctgtagaatgcgtg-3'
MYLIP1-T279D-S282D-as *

5'-cacgcattctacaggtgtgac gat gtg acc gac gccgtgatgatgcagtatag-3'
MYLIP1-T279D-S282D-s *

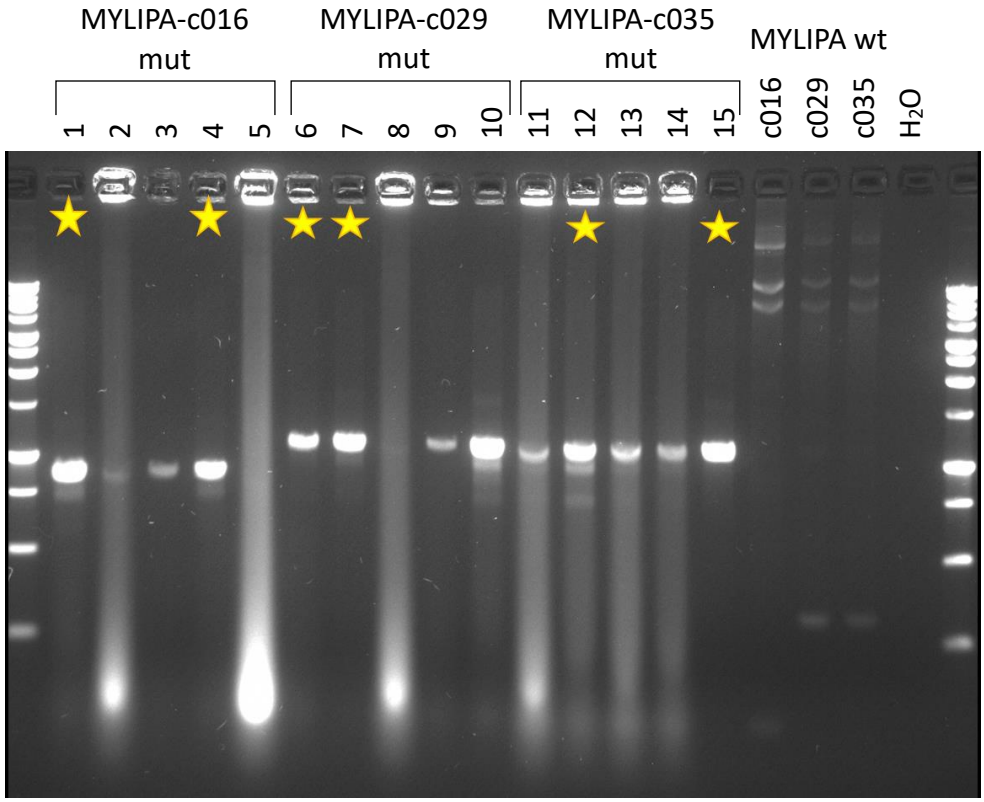


Figure A30: Agarose gel of MYLIP SDM PCRs. Stars denote products selected for sequencing.

Sequencing results:

>MYLIP c016 mut 1

NAGGAAGANNGTATATTTGACGGTCACCAAGGAATCTGGGAACAGCATCGTGCTCTTGTT
TAAAATGATCAGCACCAGGGCGGCCAGCGGGCTCTACCGAGCGATAACAGAGACGCACGC
ATTCTACAGGTGTGACGATGTGACCGACGCCGTGATGATGCAGTATAGCCGTGACTTGAA
GGGCCACTTGGCATCTGCAGAGAACCTCTACTTCCAATCGCACCATCATCACCACCATTG
ATACAAGGATGACGACGATAAGTGAGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGG
CCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGG
AAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTA
AACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGA

CGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGC
TACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCAC
GTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAG
TGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTNNTGGGAC
ATCGCCCNGATAGACN

>MYLIP c016 mut 4

GAAGANNGTATATTTGACGGTCACCAAGGAATCTGGGAACAGCATCGTGCTCTTGTTTAA
AATGATCAGCACCAGGGCGGCCAGCGGGCTCTACCGAGCGATAACAGAGACGCACGCATT
CTACAGGTGTGACGATGTGACCGACGCCGTGATGATGCAGTATAGCCGTGACTTGAAGGG
CCACTTGGCATCTGCAGAGAACCTCTACTTCCAATCGCACCATCATCACCACCATTGATA
CAAGGATGACGACGATAAGTGAGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCG
CACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAG
CTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAAC
GGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGC
GCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTAC
ACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTT
CGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGC
TTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATC
GCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACT
CTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGG
GATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGC
GAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGC
GGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAA
TTCTTANAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATC
AATACCATATTTTTGAAAAAGCCGTTTNNGTAATGAAGGANAAAACTCACCGAGGCAGTT
CCANAGGANGGCAAGATCCTGGNNNCGGN

>MYLIP c029 mut 6

NCCNGTCAGGAAGAATGTATATTTGACGGTCACCAAGGAATCTGGGAACAGCATCGTGCT
CTTGTTTAAATGATCAGCACCAGGGCGGCCAGCGGGCTCTACCGAGCGATAACAGAGACG
CACGCATTCTACAGGTGTGACGATGTGACCGACGCCGTGATGATGCAGTATAGCCGTGAC
TTGAAGGGCCACTTGGCATCTCTGTTTTGACAGTAAAGGTGGATACGGATCCGAATTCGA
GCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGG
CTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAG
CATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAANNN

>MYLIP c029 mut 7

NGTNAGGAAGAANGTATATTTGACGGTCACCAAGGAATCTGGGAACAGCATCGTGCTCTT
GTTTAAATGATCAGCACCAGGGCGGCCAGCGGGCTCTACCGAGCGATAACAGAGACGCAC
GCATTCTACAGGTGTGACGATGTGACCGACGCCGTGATGATGCAGTATAGCCGTGACTTG
AAGGGCCACTTGGCATCTCTGTTTTGACAGTAAAGGTGGATACGGATCCGAATTCGAGCT
CCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTG
CTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCAT

AACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATAT
CCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGT
TACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTT
CCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCC
TTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGA
TGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTC
CACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGT
CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCT
GATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGC
ACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT
ATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCGAGCATCAAATGAAACTGCAA
TTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGG
ANAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCC
GACTCGTCCACATCAANACAACCTATTAATTTCCCCTCGTCAAAANAAGGTATCAGNGNN

>MYLIP c035 mut 12

GCCNGTNGGAAGAANGTATATTTGACGGTCACCAAGGAATCTGGGAACAGCATCGTGCTC
TTGTTTAAAATGATCAGCACCAGGGCGGCCAGCGGGCTCTACCGAGCGATAACAGAGACG
CACGCATTCTACAGGTGTGACGATGTGACCGACGCCGTGATGATGCAGTATAGCCGTGAC
TTGAAGGGCCACTTGGCATCTCTGTTTCTGAATGAAAACATTAACCTTGGCAAGAAATAT
GTCTTTGATATTAAAAGAACATCAAAGGAGGTGTATGACCATGCCAGGAGGGCTCTGTAC
AATGCTGGCGTTTGACAGTAAAGGTGGATACGGATCCGAATTCGAGCTCCGTCGACAAGC
TTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCC
GAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGG
CCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGA
ATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGT
GACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCT
CGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGNTCCG
ATTTAGTGCTTTACGGCACCTCGACCCCNNNNN

>MYLIP c035 mut 15

CNGNNAGGAAGANTGTATATTTGACGGTCACCAAGGAATCTGGGAACAGCATCGTGCTCT
TGTTTAAAATGATCAGCACCAGGGCGGCCAGCGGGCTCTACCGAGCGATAACAGAGACGC
ACGCATTCTACAGGTGTGACGATGTGACCGACGCCGTGATGATGCAGTATAGCCGTGACT
TGAAGGGCCACTTGGCATCTCTGTTTCTGAATGAAAACATTAACCTTGGCAAGAAATATG
TCTTTGATATTAAAAGAACATCAAAGGAGGTGTATGACCATGCCAGGAGGGCTCTGTACA
ATGCTGGCGTTTGACAGTAAAGGTGGATACGGATCCGAATTCGAGCTCCGTCGACAAGCT
TGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCG
AAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGC
CTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGAA
TGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTG
ACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTC
GCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGA
GTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGG
GNNCCATCGCCATGATAGACGGTTTTTNNNCCA

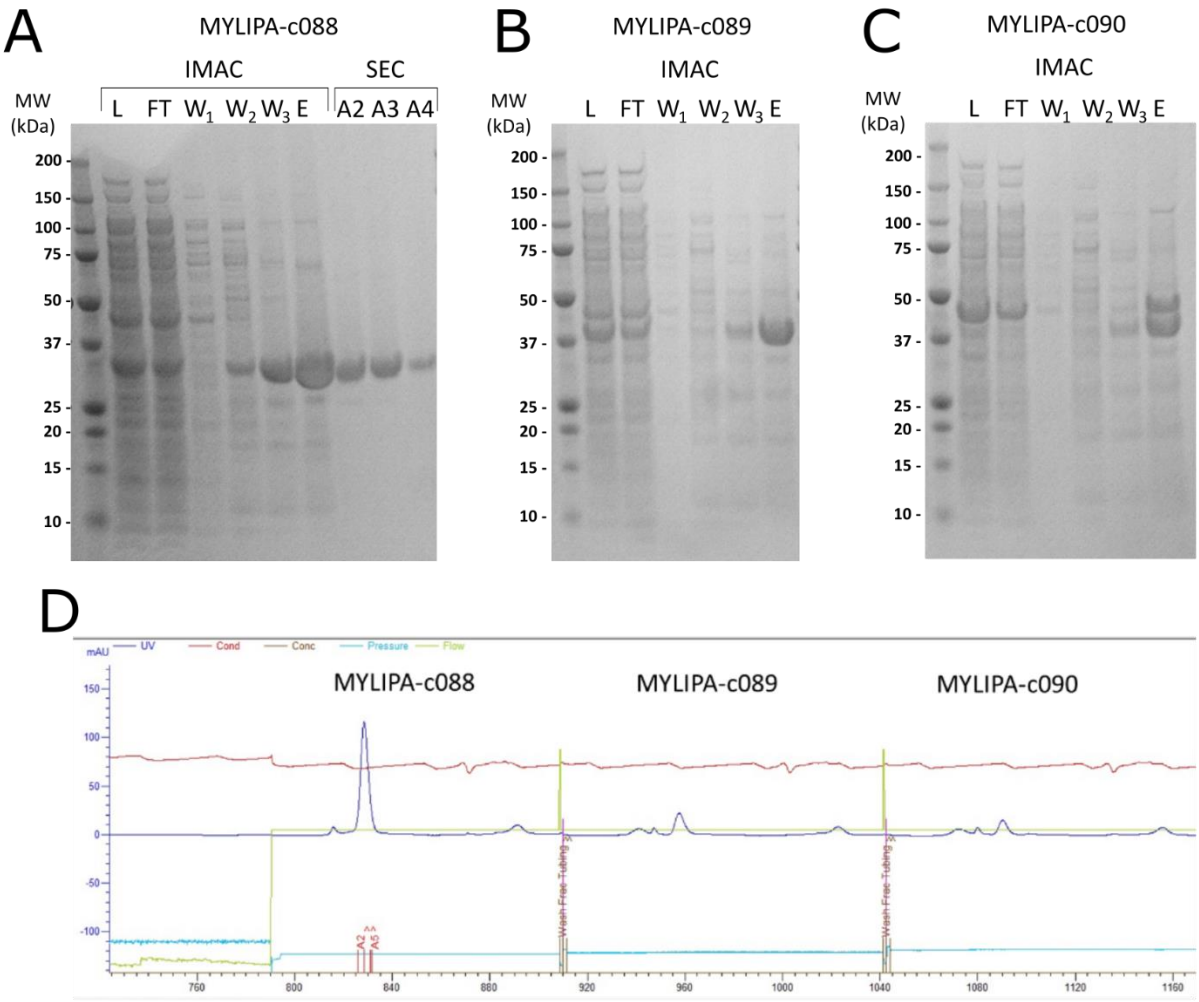


Figure A31: Purification details for phospho-mimetic mutants of MYLIP. **A-C:** SDS-PAGE after chromatography steps for MYLIPA-c088/p015, MYLIPA-c089/p016 and MYLIPA-c090/p017, respectively.: Lysate, FT: Flow through, W1-W3: washes. E: Elution. **D:** SEC chromatograms for MYLIP phospho-mimetics.

Murine E1 construct

The sequence of murine E1 was translated from the Addgene website using the Expasy translate tool (https://web.expasy.org/translate/).

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MSSSPLSKKR RVSGPDPKPG SNCSPAQSAL SEVSSVPTNG

        70         80         90        100        110        120
MAKNGSEADI DESLYSRQLY VLGHEAMKML QTSSVLVSGL RGLGVEIAKN IILGGVKAVT

       130        140        150        160        170        180
LHDQGTTQWA DLSSQFYLRE EDIGKNRAEV SQPRLAELNS YVPVTAYTGP LVEDFLSSFQ

       190        200        210        220        230        240
VVVLTNSPLE AQLRVGEFCH SRGIKLVVAD TRGLFGQLFC DFGEEMVLTD SNGEQPLSAM

       250        260        270        280        290        300
VSMVTKDNPG VVTCLDEARH GFETGDFVSF SEVQGMIQLN GCQPMEIKVL GPYTFSICDT

       310        320        330        340        350        360
SNFSDYIRGG IVSQVKVPKK ISFKSLPASL VEPDFVMTDF AKYSRPAQLH IGFQALHQFC

       370        380        390        400        410        420
ALHNQPPRPR NEEDATELVG LAQAVNARSP PSVKQNSLDE DLIRKLAYVA AGDLAPINAF

       430        440        450        460        470        480
IGGLAAQEVM KACSGKFMPI MQWLYFDALE CLPEDKEALT EEKCLPRQNR YDGQVAVFGS

       490        500        510        520        530        540
DFQEKLSKQK YFLVGAGAIG CELLKNFAMI GLGCGEGGEV VVTDMDTIEK SNLNRQFLFR

       550        560        570        580        590        600
PWDVTKLKSD TAAAAVRQMN PYIQVTSHQN RVGPDTERIY DDDFFQNLDG VANALDNIDA

       610        620        630        640        650        660
RMYMDRRCVY YRKPLLESGT LGTKGNVQVV IPFLTESYSS SQDPPEKSIP ICTLKNFPNA

       670        680        690        700        710        720
IEHTLQWARD EFEGLFKQPA ENVNQYLTDS KFVERTLRLA GTQPLEVLEA VQRSLVLQRP

       730        740        750        760        770        780
QTWGDCVTWA CHHWHTQYCN NIRQLLHNFP PDQLTSSGAP FWSGPKRCPH PLTFDVNNTL

       790        800        810        820        830        840
HLDYVMAAAN LFAQTYGLTG SQDRAAVASL LQSVQVPEFT PKSGVKIHVS DQELQSANAS

       850        860        870        880        890        900
VDDSRLEELK ATLPSPDKLP GFKMYPIDFE KDDDSNFHMD FIVAASNLRA ENYDISPADR

       910        920        930        940        950        960
HKSKLIAGKI IPAIATTTAA VVGLVCLELY KVVQGHQQLD SYKNGFLNLA LPFFGFSEPL

       970        980        990       1000       1010       1020
AAPRHQYYNQ EWTLWDRFEV QGLQPNGEEM TLKQFLDYFK TEHKLEITML SQGVSMLYSF

      1030       1040       1050       1060       1070
FMPAAKLKER LDQPMTEIVS RVSKRKLGRH VRALVLELCC NDESGEDVEV PYVRYTIR
```

Figure A32: Sequence of murine E1 construct.

Protein-peptide titrations in AlphaScreen, HTRF and fluorescence polarisation assays

HTRF – MYLIP$_{(1\text{-}276;\ C\text{-}His6)}$ vs LDLR$_{(811\text{-}834)}$-K-biotin

MYLIPA-p002 (nM) \ LDLR$_{(811-834)}$-K-biotin (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	0.4883	0.2441	-
4000	131	135	129	131	129	124	127	125	128	131	125	131	128	128	130	119
2000	135	129	132	127	127	123	127	126	119	123	122	125	127	123	135	128
1000	141	134	124	134	126	125	131	123	133	129	120	125	129	129	127	124
500	148	131	135	131	140	122	143	131	130	139	124	129	131	126	133	121
250	157	140	145	135	140	123	147	136	126	134	131	131	130	125	129	124
125	158	147	149	151	155	140	140	136	134	134	129	131	131	128	132	130
62.5	168	150	150	139	155	143	153	141	135	138	130	130	135	128	130	125
31.25	157	152	147	155	159	141	153	144	132	133	129	141	131	127	131	130
15.625	152	148	150	149	151	137	151	146	148	135	137	136	134	132	132	129
7.8125	150	149	142	147	153	124	146	139	141	142	127	129	125	131	131	130
3.90625	145	134	142	138	142	132	146	136	131	134	127	135	133	130	131	124
1.953125	140	141	135	142	140	129	138	129	133	128	128	126	127	132	130	124
0.976563	130	141	133	136	136	131	134	132	128	136	132	130	130	127	134	132
0.488281	135	141	135	138	138	132	136	128	131	136	127	131	127	124	131	128
0.244141	135	135	131	130	137	129	133	123	127	131	128	134	138	124	133	124
-	145	137	142	136	138	131	136	127	128	135	127	134	128	134	132	120

Figure A33: HTRF titration of MYLIPA$_{(1\text{-}276;\ C\text{-}His6)}$ against LDLR$_{(811\text{-}834)}$-K-biotin.

HTRF – MYLIPA$_{(1\text{-}273;\ N\text{-}His6\ SUMO)}$ vs LDLR$_{(811\text{-}834)}$-K-biotin

MYLIPA-p008 (nM) \ LDLR$_{(811-834)}$-K-biotin (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	0.4883	0.2441	-
4000	127	124	131	130	128	133	138	180	373	460	499	466	485	427	430	454
2000	128	126	132	126	124	133	137	170	517	701	670	643	651	604	597	599
1000	132	128	128	138	131	132	146	187	818	808	768	737	688	686	675	648
500	132	125	131	128	127	135	143	186	877	810	740	702	672	690	674	624
250	125	132	129	131	134	135	141	183	855	739	689	679	652	656	641	618
125	132	133	134	128	131	133	142	180	725	649	614	615	578	611	590	572
62.5	132	128	130	132	133	135	141	161	590	537	503	529	496	523	511	493
31.25	129	129	129	134	132	133	138	154	429	401	397	414	378	411	402	366
15.625	129	124	130	130	131	130	137	138	302	291	281	298	280	214	291	274
7.8125	132	128	131	130	133	131	135	136	230	216	208	221	209	219	217	203
3.90625	132	128	134	132	129	136	133	133	179	172	170	173	169	175	178	173
1.953125	129	130	128	130	132	132	130	133	159	151	146	151	147	156	151	151
0.976563	135	130	129	128	132	132	135	128	141	136	136	138	137	139	136	135
0.488281	131	132	134	132	136	136	132	132	139	133	129	132	130	133	135	130
0.244141	129	130	132	135	131	130	129	129	133	134	130	131	132	135	130	130
-	136	136	130	131	130	127	132	132	130	129	130	128	132	128	129	129

Figure A34: HTRF titration of MYLIPA$_{(1\text{-}273;\ N\text{-}His6\text{-}SUMO)}$ against LDLR$_{(811\text{-}834)}$-K-biotin in sodium phosphate buffer (20 mM sodium phosphate, pH 7.5, 100 mM NaCl, 0.05% TWEEN-20, 0.1 mg/ml BSA).

HTRF assay –different buffers

A

MYLIPA-p011 (nM) \ VLDLR$_{(820-842)}$-biotin (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
O/N	151	164	165	169	146	158	151	156	149	153	156	149	154	155	157	151
16000	146	153	160	148	163	146	153	153	149	149	145	152	156	152	152	150
8000	155	159	158	156	158	157	161	169	145	159	188	155	161	162	141	160
4000	146	149	166	149	156	149	163	161	157	160	181	165	165	145	168	149
2000	157	159	153	166	165	156	166	164	162	153	165	160	150	163	154	148
1000	154	164	149	164	160	157	149	152	165	172	161	138	151	174	157	146
500	158	153	152	169	151	162	161	151	145	168	156	159	158	147	159	162
250	159	143	165	163	155	154	163	159	153	160	157	153	150	156	147	153
125	148	150	179	153	160	170	157	157	150	143	151	150	150	157	144	158
62.5	161	164	139	156	156	146	179	147	153	151	179	157	158	149	167	164
31.25	143	171	159	149	158	161	164	164	155	163	147	164	165	162	144	156
15.625	169	145	172	151	164	170	160	156	163	157	162	148	153	154	173	164
7.8125	151	163	163	153	161	146	183	163	162	156	182	162	165	148	165	154
3.90625	172	181	176	149	162	158	157	160	157	155	153	165	144	156	160	156
1.953125	159	159	161	157	159	168	173	173	157	165	153	167	154	156	158	155
0.976563	147	149	161	152	154	151	170	157	174	150	141	155	155	155	151	153
-																

B

MYLIPA-p011 (nM) \ ApoER2$_{(894-871)}$-biotin (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
O/N	154	161	142	164	151	149	153	165	143	152	154	150	158	147	150	145
16000	148	158	161	151	161	169	148	160	160	155	157	148	160	147	152	155
8000	143	161	195	166	152	172	150	155	150	155	161	150	153	154	149	152
4000	156	174	154	153	154	153	153	152	156	172	167	164	185	141	149	161
2000	172	186	151	149	161	152	158	171	155	154	159	156	163	146	150	145
1000	147	142	156	166	164	153	195	186	144	159	153	164	163	152	151	149
500	143	164	155	162	176	172	162	160	163	156	173	161	156	145	140	158
250	156	173	166	171	190	173	230	195	182	213	305	189	160	163	145	155
125	161	171	185	195	203	295	256	252	271	374	477	254	165	152	146	153
62.5	161	156	157	238	261	374	377	361	434	535	590	357	249	150	151	150
31.25	168	167	250	284	372	544	491	517	605	655	784	548	314	197	162	152
15.625	164	175	180	320	507	488	538	546	649	646	787	570	338	198	158	154
7.8125	156	165	190	337	457	562	571	536	627	652	723	591	399	238	158	152
3.90625	164	194	225	389	473	567	555	576	603	556	674	556	360	239	158	148
1.953125	161	161	208	388	623	463	469	453	506	459	503	429	340	201	186	165
0.976563	140	163	155	154	163	149	161	147	141	177	148	161	148	147	151	161
-																

C

MYLIPA-p011 (nM) \ APP$_{(745-767)}$-biotin (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
O/N	220	259	217	217	229	251	233	223	221	220	192	179	157	163	182	151
16000	224	220	239	250	242	260	263	271	244	224	196	165	167	151	161	150
8000	147	160	165	155	233	223	232	229	226	252	197	162	144	178	153	159
4000	176	200	199	191	207	198	212	195	189	198	179	146	151	154	155	170
2000	165	173	170	170	193	186	192	180	179	160	157	170	148	143	151	154
1000	143	160	167	172	172	191	153	156	185	155	155	150	177	148	154	159
500	149	158	156	166	163	159	151	157	159	161	161	145	143	170	156	148
250	164	206	181	146	177	166	147	156	158	166	154	165	155	156	154	162
125	151	161	148	164	155	147	149	150	156	158	152	150	153	148	143	151
62.5	156	170	152	154	162	176	152	153	163	156	149	166	157	159	156	159
31.25	156	176	150	149	150	178	162	151	156	160	149	175	170	151	170	147
15.625	155	177	165	159	171	167	177	157	166	148	161	160	150	163	163	146
7.8125	164	153	151	162	153	160	165	157	156	167	148	150	150	166	149	159
3.90625	165	172	155	152	148	168	162	157	151	169	182	157	161	152	157	151
1.953125	147	162	175	168	158	162	168	158	155	164	177	159	161	154	175	162
0.976563	166	146	153	172	157	146	164	156	154	149	150	148	150	162	149	148
-																

Figure A35: HTRF titration in phosphate buffer (20 mM sodium phosphate, 100 mM NaCl, 0.05% (v/v) polysorbate-20, 1 mg/ml BSA, pH 7.4) of MYLIPA$_{(1\text{-}276\ C\text{-}His6)}$ against. **A:** VLDLR$_{(820\text{-}842)}$-biotin. **B:** ApoER2$_{(849\text{-}871)}$-biotin. **C:** APP$_{(145\text{-}767)}$-biotin.

A

VLDLR$_{(820-842)}$-biotin (nM)

MYLIPA-p012 (nM) \ O/N	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
16000	144	153	155	163	173	155	170	152	149	139	164	158	163	160	169	167
8000	145	152	148	153	149	147	151	150	150	154	139	155	162	148	149	173
4000	150	147	161	162	157	147	146	156	142	161	151	150	150	156	172	149
2000	145	154	148	153	160	155	156	164	156	151	145	150	158	156	175	162
1000	146	157	150	156	156	161	145	156	169	177	149	156	155	158	148	157
500	157	143	146	156	156	165	160	153	161	152	146	155	148	167	159	156
250	148	155	156	146	149	140	163	159	154	162	165	151	144	159	152	150
125	148	152	156	156	152	151	150	149	156	165	149	165	152	147	147	142
62.5	141	181	161	159	161	153	149	161	163	154	161	153	151	150	154	160
31.25	146	159	141	169	176	159	153	153	165	150	160	155	158	156	143	167
15.625	146	155	159	150	177	148	147	165	153	159	169	163	157	145	171	151
7.8125	161	149	158	178	172	164	168	165	161	158	176	164	149	175	164	161
3.90625	160	158	156	156	161	157	156	156	165	158	166	152	151	166	161	155
1.953125	150	155	150	147	165	162	160	153	154	171	171	166	155	155	163	162
0.976563	157	154	154	150	158	161	162	158	158	151	155	155	159	151	150	150
-	145	151	148	156	143	155	150	151	152	147	139	145	153	161	146	143

B

ApoER2$_{(894-871)}$-biotin (nM)

MYLIPA-p012 (nM) \ O/N	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
16000	166	151	164	155	148	155	155	167	148	154	156	149	159	159	154	173
8000	160	152	148	141	163	138	154	143	142	164	153	137	155	151	145	153
4000	144	167	190	252	284	313	152	153	151	152	151	150	160	146	145	144
2000	147	153	158	156	155	153	151	156	152	148	143	156	144	160	158	156
1000	149	145	145	151	155	150	151	147	155	157	154	147	156	147	146	142
500	156	151	157	142	150	158	157	160	166	171	160	141	148	145	152	143
250	161	147	164	148	159	185	173	154	174	167	159	149	161	154	149	153
125	142	158	157	170	194	185	188	183	200	229	178	160	148	153	156	150
62.5	159	162	165	175	174	211	249	240	263	329	235	168	156	156	152	145
31.25	159	160	175	180	236	231	308	402	363	477	380	227	165	153	148	151
15.625	181	193	190	237	268	341	486	509	577	662	506	260	212	159	144	154
7.8125	164	175	211	201	300	396	528	658	674	698	577	324	218	157	157	149
3.90625	177	166	176	233	415	448	644	637	676	704	720	420	228	165	164	169
1.953125	172	217	229	283	382	474	562	669	627	674	552	411	259	177	167	179
0.976563	169	162	250	324	474	500	525	562	547	573	628	456	316	181	155	161
-	142	157	151	154	151	167	154	161	148	149	152	144	142	151	143	148

C

APP$_{(745-767)}$-biotin (nM)

MYLIPA-p012 (nM) \ O/N	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.625	7.8125	3.9063	1.9531	0.9766	-
16000	184	163	164	188	153	162	155	159	162	165	153	156	149	161	152	159
8000	159	158	171	185	162	159	152	157	169	173	149	146	147	149	144	158
4000	152	164	151	150	171	148	172	145	148	148	150	153	148	143	148	149
2000	150	153	163	165	158	168	161	150	144	155	157	152	151	150	142	144
1000	149	156	158	165	169	159	150	162	157	152	149	149	148	152	150	153
500	143	147	165	179	156	160	150	155	152	161	169	156	159	149	148	160
250	159	178	161	155	160	160	153	162	159	173	154	159	148	149	161	167
125	152	182	152	167	156	164	156	156	151	162	155	159	152	160	153	160
62.5	152	165	164	152	160	152	161	161	166	153	157	161	155	147	148	149
31.25	159	162	161	169	177	169	151	151	159	149	157	154	154	162	157	147
15.625	160	165	156	159	147	149	168	165	150	162	168	154	159	171	157	156
7.8125	153	156	157	159	158	170	155	156	156	160	159	151	168	158	158	157
3.90625	147	178	151	170	154	166	164	165	153	162	150	161	166	147	161	150
1.953125	149	164	157	170	147	160	143	159	169	153	158	152	168	173	156	160
0.976563	143	146	168	147	174	162	142	150	167	157	175	142	147	159	141	173
-	162	154	169	166	168	145	143	168	154	151	160		149	155	148	144

Figure A36: HTRF titration in phosphate buffer (20 mM sodium phosphate, 100 mM NaCl, 0.05% (v/v) polysorbate-20, 1 mg/ml BSA, pH 7.4) of MYLIPA$_{(1-298,\ C\text{-}His6)}$ against receptor tail peptides. **A:** VLDLR$_{(820-842)}$-biotin. **B:** ApoER2$_{(849-871)}$-biotin. **C:** ApoER2$_{(849-871)}$-biotin.

HTRF – Reagent titration MYLIPA-p011, biotin-ApoER2$_{(849-871)}$

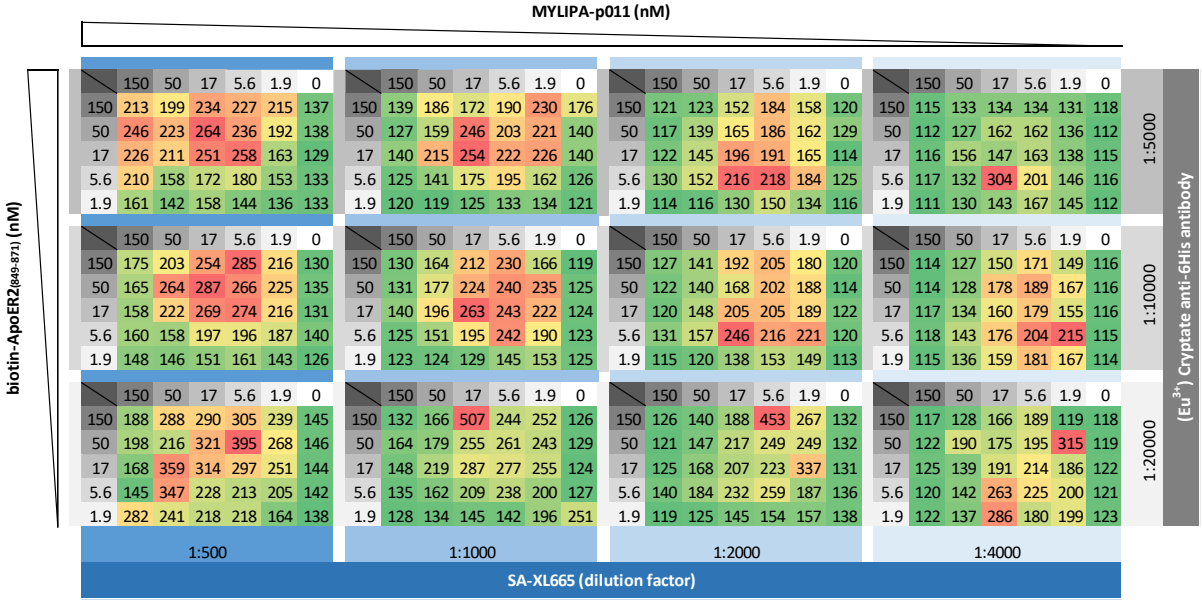


Figure A37: HTRF assay optimisation attempt Concentrations of Eu^{3+} cryptate anti-His_6 antibody, SA-XL665, MYLIPA-p011 and biotin-ApoER2$_{(849-871)}$ were varied to improve signal.

A

2h — LDLR$_{(811-859)}$-K-biotin (nM) — MYLIPA-p012 (nM)

MYLIPA-p012 (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	-
16000	54.5	77	76.6	60.9	61.8	64.2	52	49.8	42.2	38.6	41.6	37.9	36.1	38.7	33.6	21.2
8000	87	108.1	103.9	100.2	94.9	92.3	86.7	76.7	70.6	62.5	50	38.5	26.5	16.5	10.3	5.2
4000	123.3	126.3	139.4	124.8	120	116.9	110.4	100.7	89.9	80.2	63.2	40.4	21.5	10.9	4.8	2.3
2000	111.3	119.3	118.7	119.4	116.1	114.8	115.2	107.8	94.1	80.3	60.6	33.4	20.2	11.1	6.4	2.2
1000	134.2	76.4	86.8	93.1	94.7	94.9	96.4	91.2	82.1	74.3	62	43.2	24.8	14.4	10.3	2.5
500	116.2	85.1	87.2	94.4	94.1	98	98.1	94.8	92.9	83.9	59.4	40.2	27.7	18.6	12.2	2.5
250	205.8	145.9	152.2	155.9	170.1	166.1	168	155.5	135.4	114.7	75.1	57.5	41.5	26.4	15.8	3.2
125	233.7	172.2	192	188.6	203.9	214.5	213.9	200.2	178.3	146.9	103.4	73.3	56.3	35.4	19.1	3.3
62.5	280	196.7	254	299.3	293	317.4	323.8	299.7	257	222.1	146.1	107.1	77.4	50.1	31.4	4
31.25	292.4	244.2	284.4	300.8	347.2	376.4	396	382.9	364.9	308.2	199.1	156.3	108	70.4	41.4	2.3
15.63	253.2	249.5	281.1	309.4	328.2	362.5	385.6	389.9	371.8	318.6	228.8	159.5	116.9	77.8	48.4	3.7
7.81	243.9	237.5	281.8	301.7	331.4	337.6	353.6	346.9	334.3	307.3	217.5	155.4	118.5	72.4	41.9	3.4
3.91	185.3	202.2	249.3	266.8	296.2	288.8	282.4	266.9	257.8	231.9	165.2	115.5	77.9	45	27.2	3.2
1.95	109.5	129.3	148.4	150.4	162.5	183.2	183.7	174.2	159.2	136.1	99.7	68.1	42.9	21.9	13.5	2.8
0.98	58.7	67	68.4	87.1	95.6	78.3	78.6	67.3	66.1	54.6	40	24	18.5	9.3	7.2	2.5
-	9.4	4.9	5	3.7	3.8	3.4	3	2.6	3.3	3.2	3.5	4	3.8	3.4	4.1	3.4

B

2h — biotin-VLDLR$_{(820-842)}$ (nM) — MYLIPA-p012 (nM)

MYLIPA-p012 (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	-
16000	2.5	2.6	2.2	2.4	2.3	3.1	2.4	4.6	3.6	5.4	8.2	11.8	16.3	16.6	21.6	17.5
8000	2.7	2.9	2.6	2.1	2.5	2.5	2.2	2.4	2.7	2.9	3	3.3	4.4	6.2	4.3	4.5
4000	2.2	2.3	2.1	1.8	2.4	2.4	2.6	2.2	2.5	2.7	2.6	3.7	3	2.6	2.4	3.2
2000	2.9	2.6	4.9	2.7	2.3	2.1	2.3	1.9	2	2.4	2.3	1.6	2.5	3.7	2.4	1.9
1000	2.9	2.4	2.5	2.3	2.5	2.2	2.8	2.1	2.5	2.7	2.2	1.7	1.8	1.6	2.6	1.6
500	3.1	2.4	2.9	2.6	3.3	2.7	3	2.7	2.6	3.7	3.1	3	2.2	1.7	2.5	3.4
250	7.3	8.1	9.3	6.4	8.1	9	7.7	6.7	5.8	10.9	7.1	5	3.3	2.7	3.2	2.9
125	12.6	14.6	13.6	14.4	12.6	14.6	15.4	14	14.6	19.8	15.4	7.8	3.4	2.8	2.2	3.1
62.5	15.8	20.6	21.7	25.4	30	31.7	28.1	30	26.6	62.5	50.7	18.1	6.5	3.9	3.4	2.7
31.25	28.2	37.9	51.4	70.5	77.5	90.7	85.1	82.7	80.2	156.4	131.8	39.8	11.3	3.9	3.7	3.1
15.63	30.9	54	76.2	88.5	105.5	114.6	116.4	126.8	120.3	199.3	165.7	77.8	13.9	4.6	3.5	3.7
7.81	25.9	36.2	58.2	65.5	80.8	88	89.7	98.4	96.5	157.4	143.8	62.9	14.3	5	5.1	3.4
3.91	23.3	23.8	36	40.5	49.6	48.9	51.1	53	55.4	78.6	78	38.8	12.2	4.5	3.5	2.4
1.95	13.9	14.4	20	17.3	19.6	22.6	23	23.1	25.3	32.9	29.8	13.8	8.2	4.7	4.3	3.2
0.98	7.8	7	11.3	9	8.1	8.9	9.3	9.3	8.2	8.8	9.7	5.9	3.6	3.8	3.3	2.6
-	4.1	3.4	5	3.6	3.1	2.6	3.7	3.6	3.4	3.6	3.1	4.3	3.3	2.8	3	4

C

2h — biotin-ApoER2$_{(849-871)}$ (nM) — MYLIPA-p012 (nM)

MYLIPA-p012 (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	-
16000	2.2	2.4	2.5	1.7	2.9	3.5	5	13.6	31.7	47	60.4	74.4	88.5	94.9	118	133.1
8000	2.6	2.5	2.3	2.9	2.4	3.3	6	13.9	27.8	51.9	65.2	80.7	92.9	105.6	108.4	125.2
4000	2	2.3	1.9	2.7	2.8	3.3	6.3	9.8	23	44.3	74.3	94.1	106.6	121.2	131.7	137.8
2000	2.3	2.4	2.5	3.1	2.3	3.8	5.1	9.9	21.8	43.4	73.3	102.1	126	113.2	108.2	106.9
1000	2.1	2.1	9.1	4.2	4.4	5.9	7.7	13.2	23.1	45.4	80.6	120	115.2	105.8	88.7	78.2
500	2.1	2.6	2.9	4.7	7.1	7.3	9.4	13.9	22.6	40.2	87.6	129.6	110.9	85.7	68.6	58.2
250	4.9	6.6	6.6	8.1	9	13.1	18.3	21.2	28.3	43	97.1	162.5	113.9	84.1	66.9	41.7
125	2.9	4.9	7.9	10.5	14.4	21.2	27	30.1	42.4	67.5	137.2	216.8	132.7	81.9	56.9	31.9
62.5	25.8	7.4	12.2	12.9	23.3	32.6	42.7	60.2	83.4	129.9	258.3	323.2	192.2	104.7	57.7	24
31.25	7.7	10.5	18.3	23.8	34.2	48.7	66.6	94	137.6	212.4	382.6	440	305.2	147.2	68.6	18.2
15.63	9.4	16.5	28.4	36.9	58.4	64.8	80	119.2	147.7	209.9	410.5	525	344.8	150	68.1	10.6
7.81	10.9	22.6	37.5	57.6	81.4	84.9	89.3	98.2	121.7	175.1	360.9	475.1	267.1	115.8	49.6	9.5
3.91	9.4	16.9	43.6	81.1	105.5	118.7	93.6	84.8	88.7	119.9	211.4	327.9	194.4	71.6	30.3	5.5
1.95	6.8	17.5	36.7	85.6	120.5	112.5	77.1	61.1	56.5	57.7	96.9	143.1	79.8	30.4	15.2	3
0.98	5.7	9.5	25.6	71.7	96.9	83.3	53.6	36.1	27.6	30	36.7	50.6	25.6	11.6	7.7	4.1
-	5.3	2.9	3.1	2.6	3.2	2.4	2.9	3.6	3.4	4	3.6	3.8	4	3.6	3.9	3.4

D

2h — biotin-APP$_{(745-767)}$ (nM) — MYLIPA-p012 (nM)

MYLIPA-p012 (nM)	16000	8000	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	-
16000	16	18.5	16.4	16.9	11.1	12.4	11.5	16.6	18.1	20.4	22.9	25.6	22.5	23.9	18.7	18.8
8000	7.2	8	6.9	5.7	4.9	5.6	6.4	5.9	7.4	8.4	9.6	7.4	5.2	4.4	5	4.1
4000	4.1	5	3.7	3.8	3.2	4	3.4	2.8	3.5	3.2	3.1	3.2	2.9	3.1	2.5	5.2
2000	4	3.4	4.4	3.4	3.2	3.2	2.8	3	3.1	3	3.1	2.1	3.4	2.7	1.9	2.1
1000	3.7	2.6	3.4	2.1	2.6	2.2	2.5	3.3	2.7	2.7	3.3	2.2	1.6	2.5	2.1	2.9
500	2.4	2.2	2.9	2.5	2.4	2.8	2.5	2.9	2.5	2.7	2.4	2.5	3	2.2	2.4	2.5
250	3.3	3.4	3.2	4.1	3.3	2.5	4.2	2.6	3.3	2.6	3.1	2.7	2.9	2.8	3.2	2.7
125	2.7	2.8	3.2	2.8	3.2	3.1	3.1	3.2	3.2	3	3.2	3.2	3.1	3.8	2.4	3.5
62.5	3.4	3	3.7	3.2	2.2	3	3	3.8	3.6	3.5	3.1	2.9	2.9	3.3	2.9	2.8
31.25	3.7	2.8	2.9	2.6	2.2	3.3	4.4	1.8	3	3.3	2.8	2.8	3.4	2.8	3.4	3.1
15.63	2.9	2.9	4	2.3	2.5	2.9	2.9	3.5	3.6	3.2	2.8	3.4	3.5	4.6	3.3	3.3
7.81	3.4	3.2	3.5	2.5	3.5	4.2	3.7	2.8	3	2.7	2.9	3	2.7	3.4	3.5	3.3
3.91	4.4	3.4	3.6	3.1	2.8	3.6	2.8	3	2.8	3.1	3.2	2.6	3	3.2	3.7	2.7
1.95	3.6	3.9	2.9	2.6	2.4	2.7	2.5	2.8	2.6	2.7	2.7	2.7	2.8	3	3.1	7.3
0.98	2.9	2.8	2.4	3.6	2.9	2.8	2.8	2.2	3.5	3.3	2.7	2.7	3.2	2.1	2.6	3.1
-	4.4	3.3	2.4	2.5	3	2.2	3.5	3	4.2	2.9	3	2.4	3.2	3.5	3.4	2.9

Figure A38: AlphaScreen titration of MYLIPA$_{(1\text{-}298\ C\text{-}His6)}$ against receptor tail peptides. **A:** LDLR$_{(811-859)}$-K-biotin. **B:** biotin-VLDLR$_{(820-842)}$. **C:** biotin-ApoER2$_{(849-871)}$. **D:** biotin-APP$_{(745-767)}$. Signal is in thousands.

Co-operativity of MYLIP, PTB domains and receptor tail peptides

LDLR adapter protein PTB domains and LDLR family peptides

biotin-ApoER2$_{(849\text{-}871)}$ (nM)

DAB1A-p001 (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	79.3	93.7	86.5	104.1	143.5	107.5	36.3	15.7	10.1	4.6	4.1	0.6
75	116.9	163.2	174.2	235.5	308.7	205.4	83.2	27.0	17.9	9.7	4.1	0.6
37.5	127.3	207.8	250.9	327.3	398.4	316.3	165.6	56.8	32.5	17.2	6.9	0.5
18.75	108.9	182.1	243.5	330.3	398.4	323.4	185.4	75.7	44.7	20.8	6.8	0.6
9.38	76.4	136.5	191.2	252.1	308.8	271.4	159.8	81.3	45.8	117.5	7.5	0.6
4.69	46.8	73.3	109.9	151.6	189.1	180.6	123.6	77.7	42.0	18.2	5.7	0.7
2.34	25.5	36.8	56.7	73.3	96.5	97.4	82.1	49.8	29.1	10.5	3.6	0.9
-	0.6	0.6	0.6	0.9	0.6	0.4	0.6	0.5	0.4	0.4	0.3	1.3

DAB1A-p003 (nM)												
150	338.3	438.3	461.1	530.0	593.6	570.9	424.1	216.6	117.1	50.0	18.0	0.7
75	336.0	471.7	530.8	639.2	668.0	617.1	444.5	243.0	136.2	65.4	21.9	2.5
37.5	222.9	339.1	427.5	526.2	599.1	550.5	430.1	256.7	153.2	76.7	24.7	0.5
18.75	200.2	326.6	410.9	535.0	611.9	563.4	398.1	239.3	147.9	78.5	32.5	0.6
9.38	141.0	242.6	346.1	458.3	678.6	576.1	397.3	250.7	170.6	98.4	39.1	0.9
4.69	103.5	210.4	328.0	470.5	613.8	571.0	397.6	263.8	176.5	103.4	49.2	0.8
2.34	114.6	200.0	300.4	471.0	587.8	528.1	390.3	261.8	181.4	109.2	62.0	1.2
-	0.8	0.7	0.6	0.6	1.0	0.8	0.6	0.9	1.1	0.5	0.4	1.1

DAB2A-p001 (nM)												
150	147.2	179.2	198.1	237.4	285.1	266.8	144.5	55.0	16.2	6.4	2.0	1.6
75	145.9	190.7	224.4	288.2	326.8	267.6	149.6	57.1	21.2	7.4	2.9	0.4
37.5	122.6	171.1	194.7	251.2	292.2	242.2	135.6	66.7	29.2	10.5	5.6	0.4
18.75	97.4	129.5	156.6	188.3	225.4	198.9	128.7	71.0	35.6	14.0	4.6	0.6
9.38	78.9	99.7	119.7	137.8	163.0	139.2	94.7	62.8	29.1	10.2	2.2	1.0
4.69	43.9	40.9	60.1	68.9	90.9	75.6	57.8	39.7	21.6	8.8	2.2	1.0
2.34	17.0	20.9	23.6	30.6	40.7	36.9	29.5	23.0	12.9	3.6	1.2	0.7
-	0.8	0.6	0.7	0.9	0.5	0.5	1.0	0.7	0.7	0.6	1.0	0.5

DAB2A-p003 (nM)												
150	372.4	485.0	551.7	681.6	746.4	734.0	580.7	356.6	194.3	66.4	11.1	1.0
75	393.9	509.5	588.8	704.0	740.1	708.7	530.0	360.8	196.3	77.2	21.0	0.6
37.5	283.8	397.3	464.7	588.7	655.3	671.4	510.2	336.4	208.0	96.5	26.3	0.6
18.75	287.6	418.5	501.2	618.1	686.8	650.6	493.3	330.7	228.6	95.6	27.1	0.6
9.38	256.3	394.9	473.8	589.5	641.1	625.1	486.2	335.6	196.7	95.7	24.7	0.7
4.69	274.1	385.0	436.2	534.4	604.7	593.8	474.8	339.6	206.8	90.4	22.5	0.9
2.34	233.5	368.6	427.0	511.6	582.6	532.9	455.6	328.0	181.5	78.8	17.7	1.1
-	0.6	0.6	1.6	1.1	0.6	1.6	0.9	0.3	1.3	0.5	0.7	0.5

Figure A39: AlphaScreen titration of biotin-ApoER2$_{(849\text{-}871)}$ against DAB1 and DAB2 PTBs (p001: N-terminally tagged, p003: C-terminally tagged in both cases). AlphaScreen signals are shown in thousands.

	biotin-APP$_{(745-767)}$ (nM)											
	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
DAB1A-p001 (nM)												
150	136.5	149.2	109.8	158.7	115.3	87.4	95.8	63.3	36.4	10.3	2.0	0.7
75	236.7	256.2	189.9	338.2	196.7	149.5	162.9	71.4	43.4	7.1	1.3	0.8
37.5	362.1	401.3	360.4	440.0	360.6	306.7	198.3	118.6	77.2	15.7	1.7	0.9
18.75	440.6	512.3	471.1	563.6	505.8	442.1	246.3	147.7	78.4	18.1	2.1	0.8
9.38	440.8	545.7	540.4	503.2	493.7	414.8	254.0	135.4	59.1	14.1	2.0	0.4
4.69	312.8	369.4	434.9	458.8	471.8	396.9	211.4	98.9	37.4	9.8	2.9	0.5
2.34	181.5	225.4	279.1	121.2	291.9	222.7	109.7	48.7	15.5	2.5	0.8	0.7
-	0.5	0.4	1.0	0.9	0.2	0.8	0.7	0.5	0.6	0.5	0.4	0.6
DAB1A-p003 (nM)												
150	312.2	329.0	328.1	334.0	325.1	375.4	391.0	287.7	194.4	86.3	22.9	0.7
75	366.3	385.8	358.9	370.3	359.3	440.8	412.1	283.5	185.3	81.8	17.4	0.8
37.5	398.1	420.1	398.3	417.0	390.8	454.5	428.5	294.7	184.8	77.3	17.2	0.8
18.75	486.0	550.8	510.0	555.3	519.4	530.5	456.6	288.7	173.9	113.6	18.5	1.0
9.38	604.0	691.0	656.5	748.2	690.4	676.0	509.3	293.9	175.3	68.7	13.3	0.6
4.69	661.6	786.8	825.3	874.7	846.5	762.1	485.1	272.1	136.6	46.6	9.5	0.8
2.34	682.1	766.2	875.2	925.1	910.0	771.0	462.6	232.7	110.5	41.3	9.5	0.9
-	1.3	1.1	1.0	0.5	0.8	0.6	0.5	0.9	0.7	0.6	0.8	0.5
DAB2A-p001 (nM)												
150	444.8	451.6	418.1	462.3	409.1	324.0	143.8	35.2	2.6	0.5	0.4	1.1
75	532.6	550.7	524.2	545.6	482.2	349.0	153.1	34.4	2.5	1.4	0.8	1.0
37.5	494.5	517.9	509.6	514.9	474.5	355.8	147.3	35.1	3.7	0.7	0.7	0.8
18.75	434.6	460.7	454.1	459.0	433.5	318.0	134.1	34.1	3.6	0.9	0.6	0.6
9.38	328.5	347.0	364.2	361.0	332.8	224.9	79.4	18.9	1.6	0.9	0.6	1.3
4.69	165.6	177.9	198.8	188.5	190.5	119.7	44.4	8.8	1.5	1.0	0.7	0.6
2.34	76.4	82.5	92.3	86.4	81.9	49.0	10.8	2.4	0.4	0.5	0.6	0.4
-	0.6	0.9	0.6	0.6	0.9	0.2	0.4	0.3	0.7	0.7	0.6	0.4
DAB2A-p003 (nM)												
150	815.7	898.9	874.2	914.0	854.5	763.0	467.4	183.9	38.5	2.3	1.1	0.9
75	1022.2	1101.0	1038.7	1070.5	1013.7	802.8	434.2	159.4	24.4	2.5	0.7	0.8
37.5	915.6	962.2	915.9	1004.9	907.5	799.0	464.8	172.2	37.1	2.9	1.2	0.4
18.75	949.5	957.4	960.8	967.1	922.0	760.2	372.2	140.0	21.2	1.7	1.1	0.8
9.38	897.4	908.0	874.5	921.6	851.4	685.1	373.1	117.3	20.5	1.8	0.5	0.9
4.69	715.8	763.3	788.5	769.5	742.4	598.1	292.2	98.7	14.7	1.2	1.1	0.9
2.34	621.8	669.5	661.1	651.8	630.8	448.1	193.4	54.2	7.9	1.3	0.6	1.3
-	1.0	0.5	1.0	0.6	0.6	0.7	1.0	1.0	1.0	1.1	0.6	1.1

Figure A40: AlphaScreen titration of biotin-APP$_{(745-767)}$ against DAB1 and DAB2 PTBs (p001: N-terminally tagged, p003: C-terminally tagged in both cases). AlphaScreen signals are shown in thousands.

DAB1A-p001 (nM) \ biotin-VLDLR$_{(820-842)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	405	607	790	911	1024	1032	1141	1120	1091	956	676	97
75	332	585	894	1022	1257	1287	1389	1428	1416	1289	928	99
37.5	242	505	903	1213	1527	1354	1747	1741	1766	1712	1184	101
18.75	193	415	950	1402	1729	1949	2071	2039	2106	1944	1189	102
9.38	139	251	647	1129	1594	1943	2133	2185	2252	2237	1557	100
4.69	141	220	561	1045	1500	1878	2174	2283	2383	2248	1401	103
2.34	119	168	340	668	1102	1593	1919	2078	2221	2130	1598	103
-	98	97	98	97	97	99	98	98	97	95	97	97

DAB1A-p003 (nM) \ biotin-VLDLR$_{(820-842)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	607	814	960	1020	1083	1085	1107	1106	1091	942	650	101
75	516	798	1020	1164	1256	1242	1323	1300	1267	1291	875	98
37.5	507	852	1179	1392	1548	1591	1666	1625	1582	1544	942	99
18.75	390	772	1220	1533	1817	1911	2014	2002	1938	1920	1186	100
9.38	266	561	1150	1575	1921	2127	2302	2328	2328	2292	1592	100
4.69	265	597	1152	1667	2025	2368	2535	2599	2547	2408	1470	102
2.34	179	364	844	1422	1978	2376	2619	2783	2821	2777	1969	101
-	96	96	96	96	98	99	96	97	100	95	96	97

DAB2A-p001 (nM) \ biotin-VLDLR$_{(820-842)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	238	462	846	1017	1197	1359	1557	1502	1421	1377	704	99
75	192	387	790	1160	1370	1510	1776	1761	1827	1637	864	98
37.5	172	323	697	1116	1387	1632	1879	1929	2034	1924	1144	101
18.75	163	278	658	1011	1414	1747	2017	2120	2208	2060	1233	99
9.38	128	194	417	715	1035	1352	1666	1944	2063	1923	1304	100
4.69	124	168	317	513	744	1328	1224	1506	1603	1437	1069	96
2.34	108	126	181	257	350	613	638	843	946	849	753	96
-	97	98	101	97	97	95	94	98	100	98	97	98

DAB2A-p003 (nM) \ biotin-VLDLR$_{(820-842)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	563	874	1091	1218	1331	1254	1705	1463	1434	1250	860	98
75	485	895	1217	1427	1563	1658	1695	1766	1721	1570	1050	100
37.5	463	868	1261	1368	1785	1912	2045	1732	2052	1923	1148	99
18.75	327	689	1140	1628	1852	2154	2313	2309	2346	2261	1395	103
9.38	266	573	1133	1611	2000	2289	2581	2668	2711	2602	1564	102
4.69	248	503	1026	1547	1960	2382	2711	2841	2936	2764	1467	96
2.34	192	387	786	1281	1813	2215	2664	2874	2952	2942	1894	101
-	98	97	99	97	96	94	97	98	98	97	94	96

Figure A41: HTRF titration of biotin-VLDLR$_{(820-842)}$ and against DAB1 and DAB2 PTBs (p001: N-terminally tagged, p003: C-terminally tagged in both cases). AlphaScreen signals are shown in thousands.

DAB1A-p001 (nM) \ biotin-ApoER2$_{(849-871)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	134	147	233	336	443	577	714	764	908	938	717	100
75	115	131	227	301	434	639	808	884	1023	1147	1009	102
37.5	124	132	232	337	496	738	962	1074	1272	1413	1125	100
18.75	115	131	215	330	521	789	1091	1243	1490	1698	1289	102
9.38	129	126	189	289	498	801	1110	1379	1636	1828	1516	105
4.69	114	129	210	346	562	914	1234	1472	187	2012	1370	108
2.34	107	122	174	274	462	744	1055	1425	1624	1923	1766	103
-	102	98	97	97	98	97	96	101	99	98	100	97

DAB1A-p003 (nM) \ biotin-ApoER2$_{(849-871)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	171	264	344	457	539	638	729	778	930	942	696	100
75	131	185	298	436	553	652	833	887	966	1145	948	99
37.5	127	189	329	530	687	881	1009	1096	1190	1377	949	98
18.75	115	150	252	443	654	842	1175	1288	1390	1655	1335	103
9.38	114	151	238	455	712	971	1303	1483	1641	1927	1418	104
4.69	115	154	255	468	758	987	1420	1717	1931	2200	1505	104
2.34	110	140	212	452	684	1240	1463	1803	2034	2407	1859	101
-	97	98	95	96	96	95	96	99	98	95	97	94

DAB2A-p001 (nM) \ biotin-ApoER2$_{(849-871)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	105	124	172	254	360	565	584	679	745	960	631	97
75	103	116	160	232	362	462	625	729	839	1014	939	100
37.5	103	120	194	255	410	528	727	843	978	1177	900	100
18.75	101	116	151	241	371	496	719	880	1009	1217	979	100
9.38	104	115	154	225	365	508	690	843	965	1163	1008	100
4.69	103	114	149	226	333	462	603	748	866	1006	816	98
2.34	103	110	142	187	256	343	433	551	605	704	692	100
-	97	95	97	97	97	97	99	99	98	97	101	94

DAB2A-p003 (nM) \ biotin-ApoER2$_{(849-871)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	116	153	247	366	508	589	740	809	889	1019	847	100
75	116	151	234	367	500	617	774	879	1074	1125	978	101
37.5	112	154	235	396	554	704	926	1027	1160	1358	907	105
18.75	109	131	195	327	517	709	909	1063	1221	1412	1253	103
9.38	112	135	212	387	578	808	1079	1234	1399	1676	1241	102
4.69	113	139	212	389	577	892	1133	1301	1508	1883	1389	99
2.34	118	145	237	398	691	800	1267	1485	1659	1898	2532	103
-	96	95	96	98	96	95	97	96	97	99	97	96

Figure A42: HTRF titration of biotin biotin-ApoER2$_{(849-871)}$ against DAB1 and DAB2 PTBs (p001: N-terminally tagged, p003: C-terminally tagged in both cases)

DAB1A-p001 (nM) \ biotin-APP$_{(745-767)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	815	1036	1092	1143	1270	1216	1251	1220	1130	1113	746	100
75	832	1126	1321	1408	1558	1471	1534	1532	1460	1379	937	102
37.5	759	1135	1447	1655	1834	1786	1853	1849	1716	1677	1113	100
18.75	842	1003	1427	1750	1963	2032	2130	2085	2040	1938	1263	102
9.38	500	912	1339	1739	1962	2123	2202	2227	2190	1968	1295	105
4.69	296	636	1126	1559	1797	2084	2160	2190	2211	2044	1278	108
2.34	212	436	821	889	1482	1867	2020	2077	2149	1869	1147	103
-	97	99	98	99	97	96	97	98	97	96	95	97

DAB1A-p003 (nM) \ biotin-APP$_{(745-767)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	873	1110	1162	1220	1274	1269	1225	1240	1270	1177	779	100
75	986	1271	1374	1442	1436	1472	1518	1473	1447	1339	819	99
37.5	937	1389	1594	1730	1772	1781	1814	1781	1758	1634	1038	98
18.75	912	1415	1774	1979	2072	2148	2175	2081	2075	1864	1290	103
9.38	853	1420	1855	2174	2280	2407	2476	2493	2438	2291	1459	104
4.69	661	1275	1852	2215	2447	2658	2731	2738	2713	2600	1732	104
2.34	598	1267	1855	2270	2570	2826	2949	2921	2970	2797	1685	101
-	96	98	97	94	98	97	97	98	96	97	98	94

DAB2A-p001 (nM) \ biotin-APP$_{(745-767)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	101	302	450	576	674	722	793	845	835	756	774	421
75	98	262	393	544	651	733	784	830	835	815	752	488
37.5	104	262	385	531	639	701	781	798	806	783	685	412
18.75	102	205	309	443	521	586	662	662	708	662	644	435
9.38	101	213	294	405	465	534	583	569	593	549	509	317
4.69	104	166	228	304	344	389	409	390	429	412	390	274
2.34	103	133	156	205	222	241	265	242	274	265	252	191
-	93	97	93	97	98	95	98	94	97	96	94	99

DAB2A-p003 (nM) \ biotin-APP$_{(745-767)}$ (nM)	200	100	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.20	-
150	100	644	821	962	1023	1070	1056	1085	1177	1028	1004	637
75	100	688	886	1046	1149	1218	1231	1261	1187	1157	1064	626
37.5	97	601	862	1040	1168	1230	1245	1291	1264	1228	1135	655
18.75	102	490	762	984	1251	1200	1273	1279	1278	1264	1531	716
9.38	100	491	748	989	1132	1195	1252	1271	1254	1223	1058	558
4.69	105	369	587	813	985	1107	1091	1162	1208	1141	1029	640
2.34	104	371	550	770	924	970	945	1037	1041	1020	865	472
-	97	96	96	97	99	93	94	97	99	97	96	96

Figure A43: HTRF titration of biotin-APP$_{(745-767)}$ against DAB1 and DAB2 PTBs (p001: N-terminally tagged, p003: C-terminally tagged in both cases).

HTRF – MYLIPA$_{(1\text{-}280\ N\text{-}His6\ ZB)}$ vs biotin-VLDLR(820-842) in presence of DAB1A/DAB2A PTBs

A

O/N DAB1A-p002

biotin-VLDLR$_{(820\text{-}842)}$ (nM)

MYLIPA-p013 (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	731	764	866	743	675	749	710	632	666	633	625	570	594	606	535	552
2000	671	696	747	735	598	787	625	566	586	569	547	578	566	588	485	517
1000	517	479	496	514	487	1002	452	400	406	408	424	427	450	410	386	396
500	423	416	414	382	388	360	354	347	348	333	329	382	357	356	373	359
250	368	413	440	435	441	390	388	385	377	333	327	347	363	357	343	350
125	448	483	527	491	508	462	493	419	361	375	371	376	367	374	384	412
62.5	449	573	596	568	596	512	505	446	407	423	411	402	423	396	420	392
31.25	546	615	587	576	679	576	527	480	449	471	441	434	403	409	428	440
15.63	568	561	570	570	516	558	550	477	420	420	422	419	427	424	412	451
7.81	519	520	473	486	510	502	469	421	399	424	415	438	403	442	419	425
3.91	417	460	437	401	450	450	430	430	405	410	390	401	444	423	407	410
1.95	413	422	442	420	435	425	412	412	408	411	373	423	374	363	395	396
0.98	416	396	404	428	416	452	417	399	363	407	387	349	409	381	412	399
0.49	409	399	399	420	401	430	413	404	371	388	371	407	392	400	378	432
0.24	409	400	403	432	414	419	404	384	346	397	371	418	402	402	359	416
-	385	353	373	379	377	401	405	378	358	396	357	369	384	372	388	380

B

O/N DAB2A-p002

biotin-VLDLR$_{(820\text{-}842)}$ (nM)

MYLIPA-p013 (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	653	634	644	659	675		625	683	744	652	675	658	675	708	559	493
2000	560	593	644	490	580	598	603	650	546	486	476	507	544	458	483	475
1000	458	423	456	408	439	420	398	417	406	416	432	407	379	405	334	436
500	357	375	390	364	361	396	345	350	359	313	324	322	342	343	374	364
250	376	421	415	384	417	422	435	368	358	380	358	313	290	340	378	351
125	472	405	425	441	457	473	420	347	409	300	372	332	334	339	321	365
62.5	410	433	505	477	496	451	437	387	439	361	373	359	354	381	411	372
31.25	465	457	445	482	453	440	442	422	431	360	361	390	325	378	336	240
15.63	358	414	448	442	448	437	446	411	377	347	382	362	373	361	389	254
7.81	368	421	446	439	406	447	339	364	367	371	391	344	367	308	365	347
3.91	409	412	389	349	397	397	442	418	388	400	343	327	326	385	358	
1.95	396	356	395	400	375	412	380	363	397	387	347	368	348	368	312	389
0.98	409	407	398	365	422	396	363	365	361	351	344	351	329	327	359	339
0.49	410	311	373	407	405		374	380	342	363	342	350	335	373	369	391
0.24	372	382	350	357	364		384	335	404	390	332	359	354	333	367	359
-	363	344	367	353	493		386	361	351	348	386	372	332	360	374	382

Figure A44: HTRF protein-peptide titration of MYLIPA$_{(1\text{-}280\ N\text{-}His6\ ZB)}$ biotin-VLDLR$_{(820\text{-}842)}$ in presence of A: 20 nM DAB1A-p002. B: 20 nM DAB2A-p002. Wells where air bubbles led to a seeming increase in FRET ratio are greyed out.

HTRF – Phospho-mimetic MYLIP against receptor peptides in presence of PTBs

A

O/N DAB2A-p002

LDLR(811-834)-K-biotin (nM)

MYLIPA(1-298 T279D S282D C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	713	541	631	574	561	598	487	337	216	200	200	195	196	194	190	194
2000	789	876	846	831	816	801	746	532	263	221	204	200	205	203	203	204
1000	1133	1168	1262	1178	1261	1258	1136	719	330	244	228	217	203	206	204	198
500	1367	1397	1501	1443	1328	1431	1301	840	393	285	236	216	207	213	222	206
250	1179	1268	1449	1337	1213	1190	1166	932	514	350	295	240	218	221	213	212
125	1221	1384	718	890	851	801	715	504	346	294	266	229	216	222	228	217
62.5	753	721	814	773	692	630	540	414	296	263	257	231	229	227	230	210
31.25	756	781	770	883	593	565	470	332	269	258	249	237	251	227	240	209
15.63	731	637	755	660	566	593	451	320	253	256	250	235	244	247	242	219
7.81	738	587	714	642	589	570	451	384	269	254	243	246	246	256	260	224
3.91	754	589	962	639	589	568	1017	322	261	249	248	246	248	253	255	223
1.95	688	575	706	626	609	553	444	654	255	231	229	241	245	249	267	221
0.98	648	590	666	614	614	541	466	329	259	241	231	235	242	244	253	224
0.49	645	580	637	586	560	539	427	297	244	231	242	250	251	251	248	228
0.24	292	264	255	632	341	369	403	336	238	241	222	222	244	319	238	228
-	428	216	626	589	491	477	416	275	228	229	223	222	224	221	220	223

B

O/N DAB2A-p002

biotin-VLDLR(820-842) (nM)

MYLIPA(1-298 T279D S282D C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	243	220	214	228	236	235	199	205	194	192	196	195	192	191	196	200
2000	281	269	273	293	325	324	243	217	200	205	203	198	207	201	198	205
1000	451	409	448	492	572	604	318	264	219	216	202	208	213	202	210	197
500	644	576	622	685	790	740	406	320	236	215	204	205	203	206	214	209
250	691	662	766	848	889	917	590	453	298	251	224	216	209	215	202	216
125	450	462	623	374	719	722	460	354	282	246	218	213	212	207	263	273
62.5	393	750	492	511	486	522	405	327	275	247	242	220	212	216	217	204
31.25	418	432	447	463	448	459	378	302	279	250	248	242	227	216	218	216
15.63	443	452	459	566	490	456	562	300	274	256	255	238	235	212	220	225
7.81	730	439	441	462	615	431	364	308	268	255	256	240	236	232	223	223
3.91	429	446	455	445	479	265	332	292	295	271	260	279	251	228	224	222
1.95	391	432	441	475	475	467	359	289	260	656	263	264	256	237	233	226
0.98	373	436	432	478	481	482	348	282	247	258	269	264	247	236	225	229
0.49	345	509	588	436	454	475	323	273	243	245	253	263	241	240	221	225
0.24	329	378	410	473	456	475	355	274	240	239	248	291	240	225	231	224
-	313	307	325	379	345	428	305	232	226	214	211	216	227	217	222	216

C

O/N DAB2A-p002

biotin-ApoER2(894-871) (nM)

MYLIPA(1-298 T279D S282D C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	210	228	208	224	222	196	199	196	190	205	194	200	191	197	196	188
2000	267	320	294	300	328	258	231	199	205	200	198	201	200	201	212	194
1000	467	522	488	510	577	485	311	217	197	209	207	207	199	208	202	201
500	432	517	566	581	562	571	341	226	211	214	211	209	212	211	211	198
250	436	451	490	469	449	487	368	257	224	203	205	214	199	198	210	212
125	223	370	343	322	315	419	268	226	208	219	204	218	229	222	191	209
62.5	254	257	262	264	266	267	259	215	213	217	207	217	206	205	216	215
31.25	273	241	240	237	228	229	235	240	223	220	201	207	205	206	212	215
15.63	231	233	230	231	236	231	234	224	216	229	213	225	215	220	218	205
7.81	236	228	222	236	227	222	230	220	229	227	216	245	212	220	218	220
3.91	634	404	261	254	230	234	266	241	278	267	262	232	228	217	228	216
1.95	246	242	232	240	231	251	255	279	260	353	274	255	228	215	222	225
0.98	246	231	236	235	254	235	243	262	285	278	265	241	227	223	229	213
0.49	232	232	230	224	238	231	223	364	245	248	256	270	222	217	226	233
0.24	223	226	224	239	220	219	220	223	235	250	238	239	229	222	240	225
-	215	208	203	216	216	213	215	217	211	226	206	225	217	217	227	222

D

O/N DAB2A-p002

biotin-APP(745-767) (nM)

MYLIPA(1-298 T279D S282D C-His6) (nM)	4000	2000	1000	500	250	125	62.5	31.25	15.63	7.81	3.91	1.95	0.98	0.49	0.24	-
4000	1860	275	315	362	420	373	677	326	227	215	192	197	196	199	195	194
2000	430	591	536	609	737	638	1138	726	288	246	202	193	192	185	191	202
1000	750	1143	988	1237	1389	1318	2259	1144	402	303	229	197	199	196	198	204
500	1162	1627	1345	1748	1764	1790	2925	1488	500	356	266	216	201	193	204	195
250	1492	1822	1772	1891	1798	2107	2999	2131	732	454	311	245	216	214	206	218
125	352	1048	1494	1607	1533	1624	2662	1333	551	371	307	256	233	221	209	222
62.5	901	1145	1116	1232	1127	318	1944	1656	566	374	312	269	232	219	217	216
31.25	868	963	944	1042	1052	1214	1982	1437	497	354	307	265	242	223	218	208
15.63	802	898	900	976	1059	1292	1858	1049	445	346	303	271	275	243	229	222
7.81	1016	947	910	914	1010	1187	1576	1453	470	342	311	312	284	254	233	210
3.91	1061	887	931	953	1058	1260	1928	957	409	331	293	283	265	237	243	217
1.95	765	810	908	910	937	1204	1574	1321	449	328	285	266	245	243	232	217
0.98	693	776	797	889	913	1149	1542	1297	463	338	290	263	268	250	226	228
0.49	742	761	788	851	958	1174	1942	964	399	287	268	247	237	226	236	214
0.24	570	592	788	798	854	1101	1413	1299	460	303	264	249	235	236	231	236
-	352	433	461	467	710	886	1218	692	341	237	216	218	210	215	212	216

Figure A45: HTRF titration of phospho-mimetic MYLIPA$_{(1\text{-}298\ T279D\ S282D\ C\text{-}His6)}$ against receptor tail peptides in presence of 20 nM DAB2A-p002. **A:** LDLR$_{(811\text{-}834)}$-K-biotin. **B:** biotin-VLDLR$_{(820\text{-}842)}$. **C:** biotin-ApoER2$_{(849\text{-}871)}$. **D:** biotin-APP$_{(745\text{-}767)}$.