

In-vitro expression with unnatural amino acids

– **A tool for site-directed photocrosslinking of membrane proteins**

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Yajing Xiao

aus Henan, China

Jülich, September 2025

aus dem Institut für Physikalische Biologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen
Fakultät der Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Jörg Labahn

Korreferent: Prof. Dr. Henrike Heise

Tag der mündlichen Prüfung: 20.11.2025

Table of Contents

Table of Contents	I
List of Figures	VII
List of Tables.....	IX
1.1 Unnatural amino acids.....	1
1.2 Photo-crosslinking unnatural amino acids.....	4
1.2.1 Application in living cells	4
1.2.2 Orthogonal tRNA/tRNA-synthetase mediated GCE	5
1.2.3 Properties of <i>p</i> -benzoyl-L-phenylalanine.....	8
1.3 Photo-crosslinking unnatural amino acids and membrane proteins	10
1.4 In vitro protein expression system.....	12
1.4.1 The history	13
1.4.2 The application for unnatural amino acid incorporation	15
1.4.2.1 Release factor	15
1.4.2.2 Release factor 1-deficient <i>E. coli</i> S30 lysate	17
1.4.3 Membrane protein expression	19
1.4.3.1 Membrane protein expression with detergents.....	19
1.4.3.2 Membrane protein expression with liposomes	21
1.4.3.2.1 Compositions of liposomes.....	22
1.4.3.2.2 Biophysical properties of liposomes	25
1.4.3.3 Membrane protein expression with nanodiscs	26
1.5 Assembly of PS1 and PEN2.....	28
1.6 The aim of this study	30
2.1 Used instruments and materials.....	32
2.2 Protein expression materials	34
2.3 Buffers.....	36
2.4 Analysis of γ -secretase.....	44
2.5 Cloning.....	45

2.5.1 BpA-RS and AzF-RS expression vector.....	48
2.5.2 tRNA expression vector	48
2.5.3 Site-directed mutagenesis	49
2.6 Protein expression and purification	49
2.6.1 In vivo protein expression system	51
2.6.1.1 Auxiliary proteins.....	51
2.6.1.1.1 TEV, EGFP, BpA-RS, and AzF-RS	51
2.6.1.1.2 T7 RNA polymerase.....	53
2.6.1.1.2.1 T7 RNA Polymerase Lysate Preparation	53
2.6.1.1.2.2 Affi-gel Blue and DEAE Sepharose.....	54
2.6.1.1.2.3 Affi-gel Blue and ammonium sulfate precipitation	54
2.6.1.1.2.4 T7 RNA polymerase activity assay.....	54
2.6.1.2 Auxiliary tRNA.....	55
2.6.1.2.1 Harvest cells	55
2.6.1.2.2 Preparation of acidic phenol	56
2.6.1.2.3 Extraction of total tRNA	56
2.6.1.2.4 Deacylation of tRNA	57
2.6.1.3 Membrane protein.....	58
2.6.1.3.1 Competent cell.....	58
2.6.1.3.2 Expression of PS1 and truncated PS1-N190X.....	58
2.6.1.3.3 Expression of full-length PS1-N190X.....	59
2.6.1.3.4 SDS-PAGE and Western Blot.....	59
2.6.2 In vitro protein expression system	60
2.6.2.1 The pET/pEVOL co-transformed system.....	60
2.6.2.1.1 B95.ΔA and B95.ΔA / pEVOL-pBpF growth curve	60
2.6.2.1.2 Plasmid Extraction	60
2.6.2.1.3 Lysate Preparation.....	61
2.6.2.1.4 In vitro Protein Expression Assay	62

2.6.2.2 The addition of BpA-tRNA and BpA-RS	64
2.6.2.2.1 Lysate preparation	64
2.6.2.2.1.1 B95.ΔA and BL21(DE3) growth curve.....	65
2.6.2.2.1.2 Harvest B95.ΔA or B95.ΔA/pAR1219 cells	65
2.6.2.2.1.3 Crude extract preparation.....	66
2.6.2.2.1.4 Endogenous mRNA degradation.....	66
2.6.2.2.1.5 Extract preparation	66
2.6.2.2.2 Optimization of magnesium and potassium ions.....	66
2.6.2.2.3 The full-length PS1-N190X expression and solubilization	67
2.7 Photo-crosslinking assay.....	67
2.7.1 sjGST-F52X expression.....	67
2.7.2 The method of photo-crosslinking	68
2.7.3 Receivers preparation.....	68
2.7.3.1 Thin-film hydration	69
2.7.3.2 Freeze-thaw cycles and extrusion	69
2.7.3.3 Liposome Titration with CHAPSO.....	70
2.7.3.4 CHAPSO-destabilized liposome preparation.....	70
2.7.3.5 Dynamic light scattering.....	70
2.7.5 Photo-crosslinking of PS1-N190X and PEN2.....	71
2.7.5.1 PS1-N190X and PEN2 expression.....	71
2.7.5.2 Photo-crosslinking assay	72
2.7.5.3 Purification by SEC or Ni-Indigo chromatography	72
2.7.5.4 Purification after photocrosslinking the solubilized PS1-N190X/PEN2	72
2.7.5.5 Mass spectrometry assay	73
3.1 Auxiliary Proteins and tRNA	74
3.1.1 TEV Protease.....	74
3.1.2 BpA-RS and AzF-RS	75
3.1.3 EGFP	80
3.1.4 T7 RNA polymerase	81
3.1.5 Total-tRNA extraction	85

3.2 Protein Constructs	86
3.2.1 Analysis of γ -secretase	86
3.2.2 γ -secretase constructs	87
3.2.3 sjGST constructs	88
3.3 In-vivo incorporation of UAAs	89
3.4 In-vitro incorporation of UAAs	90
3.4.1 The pET/pEVOL system	90
3.4.2 The addition of BpA-RS and BpA-tRNA	92
3.4.2.1 Preparation of B95. Δ A Lysate	92
3.4.2.2 Optimizing T7 RNA polymerase	94
3.4.2.3 The optimization of Magnesium and Potassium ions	95
3.4.2.4 The addition of BpA-RS and BpA-tRNA	96
3.5 Photo-crosslinking test with sjGST-F52X	98
3.6 Photo-crosslinking of PS1-N190X and PEN2	100
3.6.1 Preparation of liposomes and CHAPSO-destabilized liposomes	100
3.6.2 PS1-N190X/PEN2 in liposomes or CHAPSO-destabilized liposomes	103
3.6.3 SEC purification of crosslinked PS1-N190X/PEN2	104
3.6.4 Photocrosslink of solubilized PS1-N190X/PEN2	105
3.6.5 PS1-N190X/PEN2 in Sulfo-DIBMA-DMPC or MSP-2N2-DMPC nanodiscs	107
4.1 Incorporation of unnatural amino acid into PS1 by in vivo assay	113
4.2 Incorporation of unnatural amino acid BpA into PS1-N190X by in vitro assay	114
4.2.1 The 2 x YTPG medium	114
4.2.2 Cell harvesting time	115
4.2.3 Cell disruption	115
4.2.4 T7 RNA polymerase	116
4.2.5 Mg^{2+} and K^{+} concentration	116
4.3 Photo-crosslinking assay	118
4.3.1 PS1 endoproteolysis	118
4.3.2 The incorporation of the PS1-PEN2 subcomplex	119
4.4 Conclusion	121

Summary.....	122
Zusammenfassung	124
Abbreviations	126
References.....	130
Appendix I: DNA and protein sequences	141
Appendix II: Primers	150
Acknowledgements.....	151
Erklärung.....	152

List of Figures

Figure 1.1 Unnatural amino acids for photocrosslinking	5
Figure 1.2 Orthogonal tRNA/tRNA-synthetase pair-mediated GCE in <i>E. coli</i>	8
Figure 1.3 The scheme of photo-crosslinking of benzophenone, adapted from (Glenn D. Prestwich 1997).....	9
Figure 1.4 Cartoon visualization of the elongation and termination phases in protein synthesis.....	16
Figure 1.5 Scheme of the influence of release factor 1 on the in vitro protein expression system to incorporate unnatural amino acids	17
Figure 1.6 The scheme of membrane protein expression by the in vitro expression system, adapted from (Federico Katzen 2009)	21
Figure 1.7 Different lipids form bilayers with different properties, adapted from (Nicola J. Harris 2022).....	24
Figure 3.1 TEV purification by Ni-NTA	74
Figure 3.2 BpA-RS construct and overexpression	75
Figure 3.3 The optimization of BpA-RS purification	77
Figure 3.4 BpA-RS purification by Ni-NTA	77
Figure 3.5 TEV cleavage of the N-terminal His-tag from BpA-RS	79
Figure 3.6 TEV cleavage of the N-terminal His-tag from AzF-RS	80
Figure 3.7 Purification of EGFP	81
Figure 3.8 The activity test of home-made T7 RNA polymerase	82
Figure 3.9 T7 RNAP purification and activity test.....	83
Figure 3.10 T7 RNAP purification and activity test.....	84
Figure 3.11 Denatured SDS-PAGE for extracted total tRNA.....	85
Figure 3.12 The distance between PS1-N190/PS1-N190X and PEN2-L98 in γ -secretase complex	87
Figure 3.13 pET-27b-Flag-PEN2/ApH1-HA/NCT-rho Construct.....	88

Figure 3.14 pET-27b-GST Construct	89
Figure 3.15 The incorporation of BpA to PS1 by the in vivo assay.....	90
Figure 3.16 In-vitro incorporation of BpA to PS1-N190X using the pET/pEVOL system .	91
Figure 3.17 The growth curve and BSA standard curve	93
Figure 3.18 The optimization of magnesium and potassium ions.....	96
Figure 3.19 Incorporating BpA to PS1-N190X	98
Figure 3.20 The crystal structure of sjGST, adapted from(Jason W. Chin 2002a).....	99
Figure 3.21 Photo-crosslinking test with sjGST-F52X.....	99
Figure 3.22 Liposome titration with CHAPSO	101
Figure 3.23 Dynamic Light Scattering for prepared liposomes	102
Figure 3.24 Photo-crosslinking PS1-N190X/PEN2 in liposomes and CHAPSO-destabilized liposomes	104
Figure 3.25 SEC Purification of PS1-N190X/PEN2 complex in liposomes	105
Figure 3.26 PS1-N190X solubilization.....	106
Figure 3.27 Purification of crosslinked PS1-N190X/PEN2 liposome complex	107
Figure 3.28 The expression of PS1/PS1-N190X and PEN2 subcomplex in Sulfo-DIBMA-DMPC nanodiscs.....	108
Figure 3.29 The supernatant of PS1/PS1-N190X and PEN2 subcomplex in sulfo-DIBMA-DMPC nanodiscs.....	109
Figure 3.30 Purification of crosslinked PS1-N190X/PEN2/sulfo-DIBMA-DMPC	110
Figure 3.31 Peptide identification of purified PS1-N190X/PEN2/Sulfo-DIBMA-DMPC ..	111
Figure 3.32 photo-crosslinking of PS1-N190X/PEN2/MSP-2N2-DMPC.....	112

List of Tables

Table 2.1.1 Instruments used.....	32
Table 2.1.2 Materials used.....	33
Table 2.2.1 E. coli strains used	34
Table 2.2.2 Vectors	35
Table 2.2.3 Media for bacterial culture.....	35
Table 2.3.1 Buffers for gel electrophoresis	36
Table 2.3.2 Buffers for tRNA extract	37
Table 2.3.3 Buffers for Competent Cell Preparation.....	38
Table 2.3.4 Buffers for E. coli Extract Preparation.....	38
Table 2.3.5 SDS-PAGE	38
Table 2.3.6 3.5% Denaturing PAGE Gel.....	39
Table 2.3.7 TEV Purification Buffers	39
Table 2.3.8 BpA-RS and AzF-RS Purification Buffers	40
Table 2.3.9 EGFP Purification Buffers	41
Table 2.3.10 Hydration Buffer	41
Table 2.3.11 T7 RNA Polymerase Purification Buffers.....	42
Table 2.3.12 T7 RNA Polymerase Activity Test Buffers	43
Table 2.3.13 sjGST Purification Buffers	43
Table 2.3.14 PS1-N190X and PEN2 Purification Buffers	43
Table 2.3.15 Buffers for mass spectrometry	44
Table 2.5.1 PCR reaction.....	46
Table 2.5.2 PCR program	47
Table 2.5.3 Double enzyme digestion.....	47
Table 2.5.4 Ligation reaction	47

Table 2.5.5 DpnI digestion	48
Table 2.6.1 Transcription Experiment	55
Table 2.6.2 Feeding mixture buffer	63
Table 2.6.3 Reaction master mixture	63
Table 2.6.4 Reaction mixture	64
Table 2.6.5 Feeding mixture.....	64
Table 3.1 Factors on lysate preparation.....	94
Table 3.2 The yield of EGFP in various in vitro systems	95
Table 3.3 DLS results after individual measurements	103
Table II.1 Primers used for cloning	150

1.1 Unnatural amino acids

Unnatural amino acids (UAAs), also referred to as non-proteinogenic or non-canonical amino acids (ncAA), are chemically synthesized or modified forms of natural amino acids, distinct from proteinogenic and non-protein amino acids. UAAs neither participate in the classical central dogma of molecular biology, which maintains the flow of genetic information through ribosomal translation of RNA from DNA, nor in the posttranslational modification of residues such as phosphorylation, glycosylation, acylation, methylation, sulfation, oxidation, nitration, hydroxylation, prenylation, and crosslinking between amino acids (Barrett 1985). Additionally, UAAs are not the products of secondary metabolism (Blaskovich 2016), meaning they do not occur in bacteria, fungi, plants, animals, or marine organisms. Due to their differing biological, chemical, and physical properties compared to natural amino acids, UAAs are essential biological tools for analyzing protein interactions, functions, structures, and drug design. Generally, they are categorized into two classes (direct probe and "click" ncAA) based on whether introduced UAAs require further modification during various studies (Chiara De Faveri 2024). Below, UAAs will be categorized into 4 groups according to their applications.

1. Fluorescent AAs (FIAAs) for light spectroscopy measurements. For example, Dansylalanine (2-amino-3-(5-(dimethylamino) naphththalene-1-sulfonamido) propanoic acid), which possesses dansyl fluorescence, was incorporated into the voltage-sensitive domain (VSD) of CiVSP to report the conformational changes during membrane depolarization (Bin Shen 2011). Coumarine lysins were used to investigate protein structure, function, and dynamics, and to localize proteins in live cells (Ji Luo 2014). Aladan (Bruce E. Cohen 2002), an alanine derivative of 6-dimethylamino-2-acylnaphthalene, was introduced into GB1 at four sites to characterize its dynamics. Acridonylalanine (Acd) (Chloe M. Jones 2021) was incorporated into two fusion proteins, the insulin receptor fused to GFP (IR-K676δ-GFP) and the hyperpolarization-activated cyclic

Introduction

nucleotide-gated ion channel fused to yellow fluorescent protein (spHCN-W355δ-YFP), in mammalian cells, by using an engineered PylRS/tRNACUA pair from *Methanosarcina bakeri* to perform fluorescence lifetime imaging microscopy (FLIM).

2. Biophysical probes to label proteins for nuclear magnetic resonance spectroscopy (NMR), Fourier-transform infrared (FTIR), and electron paramagnetic resonance (EPR) studies. For instance, ¹⁵N-labeled *p*-methoxyphenylalanine (Alexander Deiters 2005), an isotopically labeled UAA, was site-specifically introduced into sperm whale myoglobin in *E. coli*, yielding approximately 0.5 mg of protein for NMR spectroscopy. 2-amino-3-(4-(trifluoromethoxy) phenyl) propanoic acid (OCF₃Phe), ¹³C/¹⁵N-labeled *p*-methoxyphenylalanine (OMePhe), and ¹⁵N-labeled *o*-nitrobenzyl-tyrosine (oNBTyr, a UV-photocaged unnatural amino acid) (Susan E. Cellitti 2008) were introduced individually into human fatty acid synthase (FAS-TE) at specific sites in *E. coli* to develop NMR methodologies for large proteins and a new method for site-specific labeling of proteins without altering their sequences. The UAA *p*-azido-L-phenylalanine (Shixin Ye 2010) was integrated into rhodopsin to monitor significant helix rearrangement via FTIR. UAA *p*-acetyl-L-phenylalanine (p-AcPhe) (Mark R. Fleissner 2009) was site-specifically inserted into T4 lysozyme (T4L) in *E. coli* to react with a hydroxylamine reagent, generating a nitroxide side chain for EPR spectroscopy analysis.

3. Crosslinkers used to investigate protein interactions, including three variable groups. The first one is photo-crosslinkers such as benzoylphenylalanine (BpA, Figure 1.1 a) (Jason W. Chin 2002a), based on photoactivatable groups benzophenone, *p*-azidophenylalanine (AzF, Figure 1.1 b) (Jason W. Chin 2002b), derived from aryl azide, and trifluoromethylphenyl (TmfdPhe, Figure 1.1 c) (Eric M. Tippmann 2007), based on diazirine, were site-specifically incorporated into target proteins synthesized in *E. coli* using the amber suppressor tRNA/aminoacyl-tRNA synthetase (aaRS) pairs. The second one is

chemical crosslinkers. *p*-2'-fluoroacetylphenylalanine 1 (Ffact) (Zheng Xiang 2013) was incorporated into a ZSPA affibody to form a covalent bond with a cysteine from the Z-protein due to proximity-enhanced bioreactivity. The last one is the bifunctional crosslinkers. N ϵ -3-(3-methyl-3H-diazirine-3-yl)-propaminocarbonyl- γ -seleno-L-lysine (Shixian Lin 2014), a crosslinker with photo- and chemical properties, was introduced to the *E. coli* chaperone HdeA to pull down interactive proteins (prey proteins) after illumination at 365 nm, allowing for subsequent separation through H₂O₂-mediated oxidative cleavage and the further analysis of prey proteins via mass spectrometry.

4. Photocaged and photoswitchable AAs to control protein function spatiotemporally. For instance, the photocaged selenocysteine (Adarshi Welegedara 2018) was integrated into *E. coli* peptidyl-prolyl cis-trans isomerase B (PpiB) and *Zika virus* NS2B-NS3 protease (ZiPro) in *E. coli* using the PCC2RS/tRNA^{CUA} pair (a mutant pyrrolysyl-tRNA synthetase from *Methanosarcina mazei*) to study ligand binding through NMR spectroscopy after UV irradiation.

Among these, photo-crosslinking UAAs are powerful tools for studying protein interactions due to their remarkable light-controlled functionality and applicability in living cells. In addition, in identifying protein interactions, photo-crosslinking unnatural amino acids can capture weakly or transiently interactive proteins and prevent false-positive interactions that arise from the loss of spatial organization during the lysis procedure in classical affinity-based methods, such as co-immunoprecipitation (Monika Suchanek 2005). Besides, compared to chemical-crosslinking UAAs, photo-crosslinking UAAs exhibit superior specificity and react with any amino acids of the interactive partner, rather than just nucleophilic amino acids. In the following chapter, I will briefly overview previous work and explain the mechanistic aspects of the photo-crosslinking unnatural amino acid *p*-benzoyl-L-phenylalanine (BpA).

1.2 Photo-crosslinking unnatural amino acids

Photo-crosslinking unnatural amino acids are derivatives of photoactive moieties, such as aryl azide, benzophenone, and diazirine (Yasmin Aydin 2023a). Theoretically, they are chemically inert to other molecules in the physiological milieu and can incorporate into proteins using either in vivo or in vitro reaction systems. However, they convert into bioactive groups once UV light is irradiated at a specific wavelength. The illuminated crosslinker covalently interacts with non-specific biomolecules at a distance of a few angstroms through particular chemical groups to facilitate protein interaction studies in superior specificity (Coin 2018).

1.2.1 Application in living cells

To apply the photo-crosslinking unnatural amino acids in living cells, on the one hand, the photo-crosslinking moieties replace the particular residue of canonical amino acids to generate isosteric UAAs for recognition of the endogenous tRNA/tRNA-synthetases pair. In this case, photo-methionine (Figure 1.1 f) and photo-leucine (Figure 1.1 e) are incorporated into PGRMC1 (a progesterone-binding membrane protein of the endoplasmic reticulum) and Insig-1/ SCAP co-expressed in COS7 cells to identify the interaction between PRMC1 and Insig-1 for the first time (Monika Suchanek 2005). Photo-lysine (Figure 1.1 g) is exploited to capture interactive proteins in lysine post-translational modifications (Tangpo Yang 2016).

On the other hand, photo-crosslinking UAAs are incorporated into interested proteins through orthogonal tRNA/tRNA-synthetase pair-mediated genetic code expansion (GCE) at a specified position, such as BpA, AzF, and DizPK (Figure 1.1 d). Compared to the other method, the significant advantage of orthogonal tRNA/tRNA-synthetase pair-mediated genetic code expansion is that photo-crosslinking UAAs can site-specifically locate to the protein of interest. Combined with the principle of photoreactive moieties and mass

spectrometry, this so-called “targeted photo-crosslinking” method provides more information on protein-ligand interactions or intra-/inter-protein interactions. Therefore, this method is one of the considerations of this work. To better understand the history of discovery and application of photo-crosslinking UAAs through orthogonal tRNA/tRNA-synthetase pair-mediated GCE, I will pay more attention to *p*-benzoyl-L-phenylalanine (BpA).

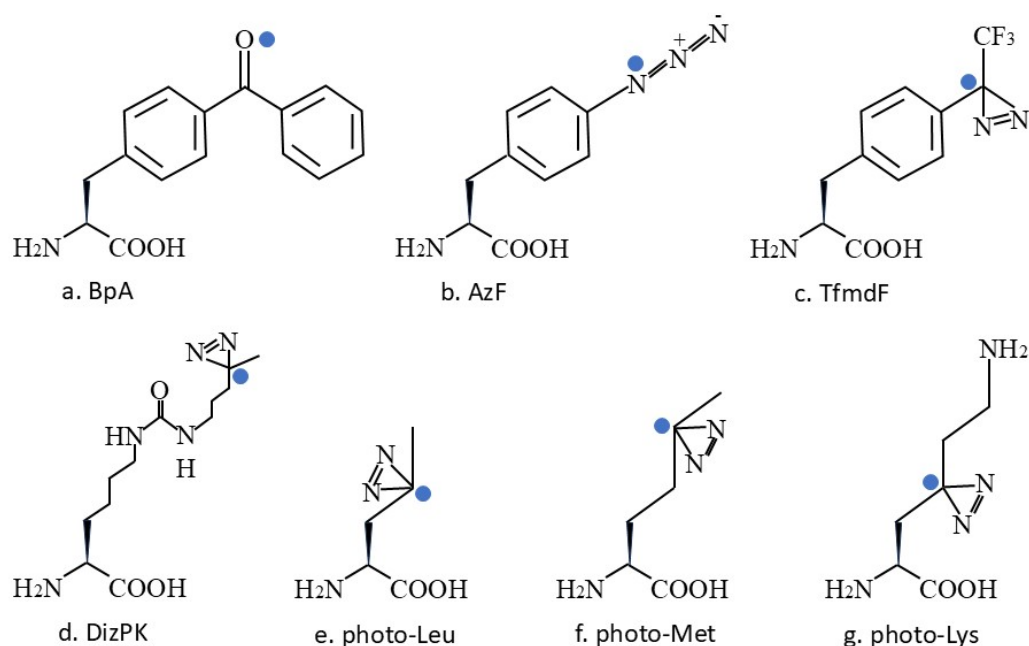


Figure 1.1 Unnatural amino acids for photocrosslinking

p-benzoyl-L-phenylalanine (a. BpA); *p*-azidophenylalanine (b. AzF); *p*-trifluoromethyl-diazirinyll-L-phenylalanine (c. TfmdF); Nε-(3-(3-methyl-3H-diazirine-3-yl)-propamino-carbonyl)-lysine (d. DizPK); photo-leucine (e. photo-Leu); photo-methionine (f. photo-Met); photo-lysine (g. photo-Lys). The photo-crosslinking position is highlighted with a blue dot.

1.2.2 Orthogonal tRNA/tRNA-synthetase mediated GCE

Since 1973, R. E. Galardy et al. have demonstrated that aromatic ketones in benzophenone, which are chemically stable before photolysis, chemically inert with water, and do not damage proteins during photolysis at long wavelengths, can serve as a new photochemical probe to analyze ligand-receptor or protein-protein interactions (R. E. Galardy 1973). Consequently, scientists increased their efforts to study the synthesis and application of benzoyl amino acids. In

Introduction

1986, James C. Kauer and his colleagues utilized solid-phase technology to successfully synthesize and incorporate the photo-crosslinking UAAs *p*-benzoyl-L-phenylalanine (BpA) into a calmodulin-binding peptide, achieving a crosslinking efficiency of 70% (James C. Kauer 1986). However, the chemical introduction of a photoactivatable group into BpA initially limits its application in small peptides that are amenable to total chemical synthesis. To expand the application of BpA on biological macromolecules, Christopher J. Noren et al. developed a biosynthetic method to incorporate synthesized amino acids into proteins in 1989. They acylated chemically a suppressor tRNA to recognize the amber codon (TAG) and substitute unnatural amino acids for the F66 in β -lactamase using an in-vitro system (Christopher J. Noren 1989). Simultaneously, J. D. Bain and coworkers applied a similar method to biosynthesize an acylated tRNA that incorporates the unnatural amino acid iodotyrosine site-specifically into a 16-residue polypeptide (J. D. Bain 1989). But all these efforts could not open the gate of photo-crosslinking UAAs incorporation in living cells, until three years later, J.D. Bain's team developed a new tRNA, with its 65th codon/anti-codon pair incorporating the non-standard nucleosides iso-C and iso-dG, to integrate unnatural amino acids into 18-residue polypeptides using a reaction system based on ribosome translation from rabbit reticulocyte lysate (J. D. Bain 1992). This finding of ribosome-mediated unnatural amino acid incorporation finally provides a pathway for investigating the site-specific substitution of photo-crosslinking unnatural amino acids into proteins in both prokaryotic and eukaryotic organisms.

Theoretically, the site-specific incorporation of unnatural amino acids into proteins in vivo requires several conditions to execute the biosynthetic program. Taking *E. coli* as an example, the new tRNA and aminoacyl-tRNA synthetase must be orthogonal. This means that first, the orthogonal tRNA should not be recognized by the endogenous *E. coli* aminoacyl-tRNA synthetase; second, the orthogonal tRNA must only be charged with non-natural amino acids when

responding to nonsense (Ochre codon UAA, Amber codon UAG, and Opal codon UGA) or four-base codons instead of encoding any of the 20 common amino acids; third, the orthogonal aminoacyl-tRNA synthetase must be capable of aminoacylating the orthogonal tRNA rather than the endogenous tRNAs; fourth, the orthogonal aminoacyl-tRNA synthetase can only aminoacylate the orthogonal tRNA with unnatural amino acids, instead of with common amino acids (Lei Wang 2001).

Considering these requirements, the Peter G. Schultz group not only introduced a mutated tyrosyl tRNA/tRNA-synthetase pair from *Methanococcus jannaschii* into *E. coli* to incorporate the synthetic amino acid O-methyl-L-tyrosine into the desired protein at an amber nonsense codon (Lei Wang 2001), but also introduced the tyrosyl tRNA/tRNA-synthetase pair from *E. coli* to *Saccharomyces cerevisiae* to incorporate five non-amino acids into interesting proteins (Jason W. Chin 2003). Since the establishment of orthogonal tRNA/tRNA synthetase-mediated GCE (Figure 1.2), it has been possible to incorporate unnatural amino acids into proteins through an in vivo expression system. The evolution of *M. jannaschii* tyrosyl tRNA/tRNA-synthetase pair provides the opportunity to introduce photoactive amino acid *p*-benzoyl-L-phenylalanine into dimeric protein glutathione S-transferase (sjGST) in response to amber codon TAG in *E.coli*, resulting in the mutant sjGST being crosslinked with high efficiency after illumination for 1 minute with a 360-nm lamp (Jason W. Chin 2002a). From the same group, another photoactive amino acid, *p*-azido-L-phenylalanine, is introduced into sjGST by the *E. coli* expression system with the help of the mutant *M. jannaschii* tyrosyl tRNA/tRNA-synthetase pair, and the crosslinking efficiency is around 30% after illumination for 5 minutes with a 254-nm UV lamp (Jason W. Chin 2002b). To date, pEvol-plasmids have been exploited and commercialized to support the incorporation of various unnatural amino acids into target proteins by *E. coli*. As the classical

photo-crosslinking UAAs, the properties of *p*-benzoyl-L-phenylalanine will be described in the following section.

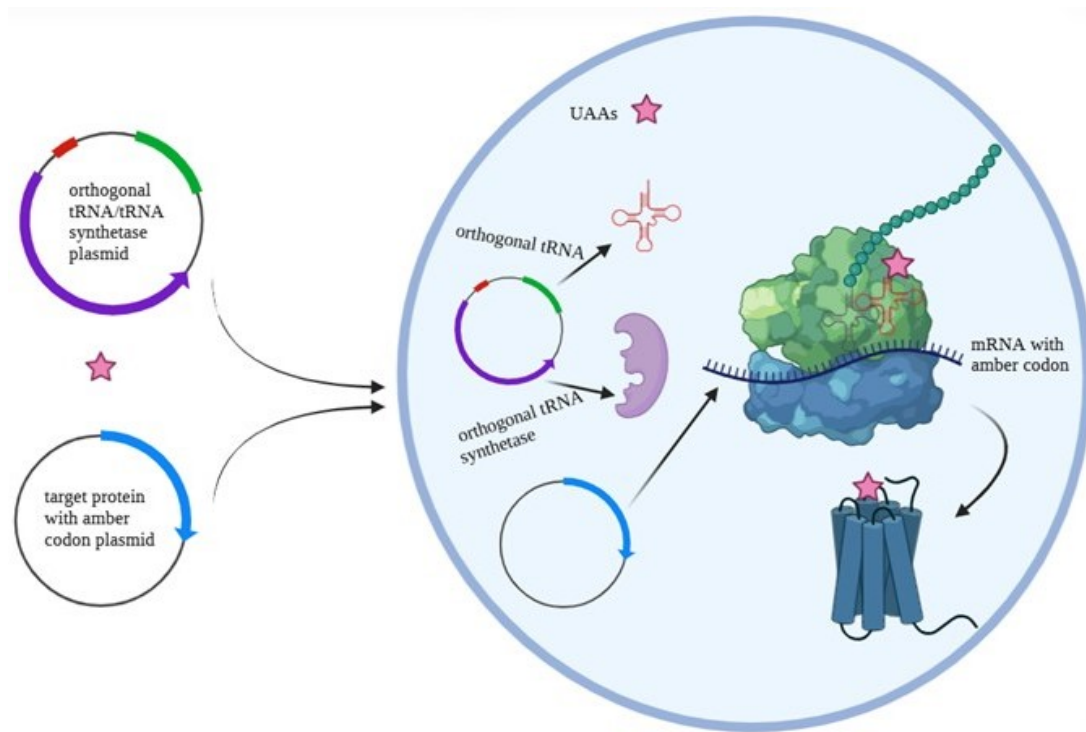


Figure 1.2 Orthogonal tRNA/tRNA-synthetase pair-mediated GCE in *E. coli*

The plasmid containing the orthogonal tRNA/tRNA-synthetase pair (pEvol-) and the plasmid involving the gene with the amber codon for the target protein are co-transformed into *E. coli* competent cells. Unnatural amino acids are added to the media. The orthogonal tRNA, charged with its UAA by the orthogonal tRNA synthetase, recognizes the amber codon (TAG), incorporating UAA into the nascent polypeptide chain.

1.2.3 Properties of *p*-benzoyl-L-phenylalanine

p-benzoyl-L-phenylalanine (BpA), an aryl ketone derivative of phenylalanine, contains a benzophenone moiety, which is a photoactivatable moiety that produces covalent crosslinking (Glenn D. Prestwich 1997). The photolabeling process involves three fundamental steps (Figure 1.3): 1. the excitation of diradicaloid triplet state, where an electron from the n-orbital of the oxygen in the benzophenone moiety is promoted to the π^* -orbital of the carbonyl group by the absorption of a photon at approximately 360 nm, resulting in electron deficiency; 2. hydrogen abstraction, during which the resulting electron-

deficient oxygen n-orbital becomes electrophilic and tends to interact with weak C-H bonds from nearby molecules to fill the n-orbital, leading to the formation of radicals in keto and alkyl; 3. radical recombination, in which the resulting radicals recombine to form a new C-C bond, facilitating photo-crosslinking (Gyorgy Dorman 1994).

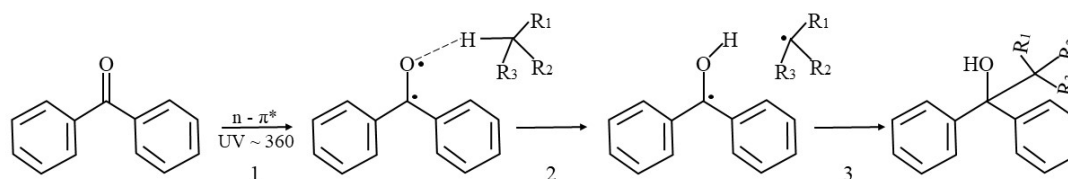


Figure 1.3 The scheme of photo-crosslinking of benzophenone, adapted from (Glenn D. Prestwich 1997)

This specific photoactive method of benzophenone provides BpA with appealing advantages. First, the experiment can operate easily due to the stability of the benzophenone moiety in ambient light. Second, the protein incorporated with BpA will not be damaged by light from the reactive wavelength of benzophenone at approximately 360 nm. Third, the photo-crosslinking efficiency of the incorporated protein will remain unaffected by solvent water or nucleophilic molecules, as benzophenone preferentially reacts with unreactive C-H bonds. Fourth, the photo-crosslinking ability of the incorporated protein persists until the formation of the C-C covalent bond because the triplet state of benzophenone is reversible, resulting in a high crosslinking yield. Fifth, compared to other photo-crosslinking groups, such as aryl azides, diazoesters, and diazaarenes, benzophenone is chemically and physically more stable, not only due to its inertness to acidic or basic environments and resistance to chemical or enzymatic cleavage but also because it does not undergo photo-dissociation. Additionally, BpA prefers to incorporate into the hydrophobic regions of proteins, which facilitates studies of membrane proteins (Jason W. Chin 2002a; Glenn D. Prestwich 1997).

1.3 Photo-crosslinking unnatural amino acids and membrane proteins

Membrane proteins bind to one leaflet of or are integral to a phospholipid bilayer, encoded by ~25% of protein-encoding genes in all organisms (Keenan 2022; Heijne 1996). In comparison to water-soluble proteins, although membrane proteins have only two architectures- α -helix and β -barrel (Heijne 1996)- they perform essential roles in cell activities, including transporting ions, metabolites, and larger molecules; facilitating the propagation of electrical impulses and chemical signal transmission; attaching to the extracellular matrix or cells; maintaining organelle and cell shapes; regulating intracellular vesicular transport; controlling lipid composition; determining cell locations and so on.

However, less than 1% of all resolved protein structures were membrane proteins due to limitations in research techniques and methods before 2007, when X-ray crystallography was the main technique to determine the 3D structure of proteins (Heijne 2007). Although the development of cryogenic electron microscopy (cryo-EM) can determine the structure of membrane proteins at high resolution, researchers can only obtain the static conformation of these proteins rather than other metastable conformations, implying that our understanding of membrane proteins is far from comprehensive. Therefore, the structural information of flexible regions in membrane proteins can only be supplemented by either high-performance computational molecular dynamics (HPC-MD) or biochemical methodologies, such as NMR, EPR, and signal molecular techniques (Chiara De Faveri 2024). In general, these techniques are based on detergent-solubilized membrane protein, which is still challenging in maintaining the function of membrane proteins. However, the emergence of orthogonal tRNA/tRNA-synthetase pair-mediated genetic code expansion (GCE) provides the chance to avoid detergent solubilization and catch conformation changes of variable membrane proteins in native membranes or artificial membranes such as liposomes, nanodiscs, cycloalkane-modified

Introduction

amphipols or copolymers, resulting in the rapid growth of membrane protein studies. The sections below show examples of how photo-crosslinking UAAs facilitate membrane protein studies through orthogonal tRNA/tRNA-synthetase pair-mediated GCE in living cells.

The pioneer of photo-crosslinking UAAs incorporation into membrane protein in its natural environment through orthogonal tRNA/tRNA synthetase pair-mediated GCE happened in 2009, when one UAA *p*-benzoyl-L-phenylalanine was incorporated into one prototypical yeast G protein-coupled receptor (GPCR) Step2 in yeast cells (George Umanah 2009). Since then, researchers have had a new tool to explore membrane protein structure and function changes and provide information inaccessible to classical structure methods. The UAA *p*-azido-L-phenylalanine was introduced into the G protein-coupled receptor rhodopsin in HEK cells (Shixin Ye 2009). After combining this photo-crosslinking UAAs incorporation technique with Fourier-transform infrared (FTIR) difference spectroscopy, they discovered the formation of a new charged cluster at the H5/H6 interface of rhodopsin, which was not observed by fluorescence spectroscopy, EPR spectroscopy, or NMR. Therefore, this photo-crosslinking UAAs incorporation technique developed a novel strategy to monitor how the electrostatic environments of specific interhelical networks change during GPCR activation. The incorporation of UAA *p*-benzoyl-L-phenylalanine (BpA) into arrestin-2 to gain insight into the interaction with the secretin-like parathyroid hormone 1 receptor PTH1R in live HEK293T cells. With this method, they unveiled the unacknowledged interactive network in the flexible domain of the arr2-PTH1R complex from the previously resolved structure (Yasmin Aydin 2023b). In addition, this photo-crosslinking UAAs incorporation technique gives insights into controversial opinions on previous research. Incorporating photo-crosslinking UAA BpA into Der1-Myc in living cells strongly supports the evidence that Der1 directly extracts aberrant proteins from the endoplasmic reticulum (ER) lumen (Martin Mehnert 2013). Besides,

this method also opens the door to studying the structure and function relationship of membrane proteins heterologously expressed in *Xenopus* oocytes with high tolerance to UV treatment. For example, GluN2A, the subunit of N-methyl-d-aspartate receptors (NMDARs), is confirmed to have low sensitivity to Zn^{2+} concentration after the tyrosine at position 281 was replaced by AzF, facilitating ligand-gated ion channels(LGICs) related research (Shixin Ye 2013). AzF-containing GluN1 NTD interacts more tightly with the GluN2B NTD than with the GluN2A NTD, giving impetus to study structures and mechanisms of neurotransmitter receptors (Shujia Zhua 2014).

Another powerful tool for expanding the application of this technique in membrane protein studies is the cell-free protein expression system. Compared to living cells, using cell-free protein expression systems in this technique overcomes limitations such as low transfection efficiency and high costs in mammalian cells, the unexpected formation of inclusion bodies, or cytotoxic effects due to the overexpression of membrane proteins in *E. coli* (Schultz 2016; David Garenne 2021), as well as the time-consuming purification process and the risk of protein instability during solubilization with detergents (Zachary A. Manzer 2023). Therefore, the next section will explore the development and application of the in vitro protein expression system.

1.4 In vitro protein expression system

The in vitro protein expression system is utilized based on the transcription-translation mechanism of living cells to facilitate protein expression in an open and accessible environment. This capability enables researchers not only to manipulate the biochemical environment, thereby satisfying the requirements for different protein modifications, particularly the incorporation of unnatural amino acids, but also to simplify the protein purification procedure. Contemporary in vitro protein expression systems enable the expression of proteins in diverse reaction volumes and containers. Consequently, this technique is appealing from the laboratory to the factory.

1.4.1 The history

The in vitro protein expression system, cell-free protein synthesis (CFPS), is initially used to decipher the translation process of how amino acids are incorporated into proteins. The rat liver extract was used to illustrate the role of ribonucleoprotein particles (currently known as ribosomes) in incorporating amino acids into proteins (John W. Littlefield 1955), whose essential role was also proved in the *E. coli*-mediated in vitro protein expression system (A. Tissières 1960). As the establishment of the rat liver-derived in vitro system, the process of protein translation was elucidated by the observation that ATP activated the amino acid to form a high-energy complex referred to as an enzyme-bound amino acyl-AMP compound (Mahlon B. Hoagland 1957), subsequently followed by the transfer of the activated amino acid to soluble ribonucleic acid (Mahlon B. Hoagland 1958). In the same year, the utilization of pigeon pancreas extract helped to discover that polynucleotides played an intermediary role in the amino acid incorporation process (Samuel B. Weiss 1958), and the *Tetrahymena pyriformis* extract helped to demonstrate the reversible step of activated amino acid to soluble ribonucleic acid in the amino acid incorporation process (Lipmann 1958).

With a comprehensive understanding of the translation process, various in vitro protein expression systems were initially utilized to employ mRNA as the template for protein synthesis. These systems relied on extracts from diverse eukaryotic cells, including rabbit reticulocyte (Língrel 1969), mouse Krebs II ascites cells (Michael B. Mathews 1971), rat and mouse liver (Jeffrey Sampson 1972), HeLa cells, mouse L cells, Chinese hamster ovary (CHO) cells (Maxson J. McDowell 1972), and wheat germ (Paterson 1973), or prokaryotic cells, specifically *E. coli* (Nirenberg 1961). Among these, the most successful and effective in vitro protein expression systems are based on rabbit reticulocyte (William R. Woodward 1974), wheat germ (Carl W. Anderson 1983), and *E. coli* (Nirenberg 1961).

Introduction

To improve the expression of exogenous mRNA, several methodologies have been exploited in the last decades. Considering the influence of endogenous mRNA, the *E. coli* extract was dialyzed (Nirenberg 1961), and the rabbit reticulocyte extract was treated with micrococcal ribonuclease to degenerate endogenous mRNA and run-off ribosomes (Jackson 1976). The risk of unstable mRNA was reduced by introducing the coupled transcription-translation system into an in vitro protein expression system. In this system, linearized or non-linearized DNA fragments replaced mRNA as templates (Zubay 1983; Wieslaw Kudlicki 1992), and phage promoters and RNA polymerases (SP6/T7), which are specific and exclusive for the expression of target genes, were employed to enhance the operational efficiency of the system (Pratt 1991; Wieslaw Kudlicki 1992). Furthermore, the continuous cell-free translation system was established utilizing extracts from *E. coli*, wheat germ, and rabbit reticulocyte to enhance the efficiency of protein expression (Alexander S. Spirin 1988; L. A. Ryabova 1989). This enhancement is achieved by providing the essential small molecules required for protein translation, including ATP, GTP, and amino acids, while simultaneously facilitating the removal of inhibitory by-products, such as phosphate, from the reaction chamber (Chong 2014). Besides, genetically deficient *E. coli* strains, such as MRE600 (RNase I⁻), K-12 strain CF 300 ($\Delta recB21$), and A19 ($\Delta speA$, $\Delta tnaA$, $\Delta sdaA$, $\Delta sdaB$), were developed to avoid the degradation of mRNA, linear DNA, and amino acids, respectively (Gesteland 1966; Huey-Lang Yang 1980; Takanori Kigawa 2004). Endosperm-free wheat germ extract was prepared to eliminate the inhibition of proteins to ribosomes and other translational proteins (Kairat Madin 2002). Eukaryotic translation initiation factors (eIF2, eIF2B, eIF4G, and p97)-supplemented HeLa cell extract was exploited to augment the yield of synthesized proteins (Satoshi Mikami 2006). Except for these, the PURE cell-free system was developed by using purified components from *E. coli* to eliminate the influence of non-participating factors in the translation reaction, for example, nucleases degrading substrates, proteases cleaving target proteins, and release factors

reducing unnatural amino acid incorporation efficiency (Yoshihiro Shimizu 2001; Yoshihiro Shimizu 2005).

The application of an in vitro protein expression system refers to a broad range, including Nucleic Acid Programmable Protein Array (NAPPA) for high-throughput analysis of protein structures, functions, and interactions (LaBaer 2011), site-specific incorporation of unnatural amino acids into target proteins (Christopher J. Noren 1989), and membrane protein expression (Yutetsu Kuruma 2015).

1.4.2 The application for unnatural amino acid incorporation

The orthogonal tRNA/tRNA-synthetase pair-mediated Genetic Code Expansion (GCE) provides a streamlined and user-friendly methodology for the site-specific incorporation of unnatural amino acids into desired proteins. In contrast to in vivo systems, the in vitro protein expression system optimally leverages the advantages of this technology. The open nature of the non-living in vitro system facilitates, firstly, the utilization of unnatural amino acids (UAAs), which can pose challenges for living cells due to their toxicity or poor cellular uptake; secondly, it reduces the limitations in accessing orthogonal tRNAs and aminoacyl-tRNA synthetases (aaRSs), which can be toxic to living cells; and thirdly, it lowers the costs associated with obtaining target proteins, resulting from a reduced required quantity of UAAs (Eiko Seki 2018). When compared to other cell-free protein synthesis systems (CFPSs), the *E. coli* extract-based CFPS exhibits high efficiency in UAA incorporation (Anne Zemella 2015), particularly following the deletion of release factor 1 (RF1) in the extract (Seok Hoon Hong 2014).

1.4.2.1 Release factor

During the elongation of the amino acid sequence, cognate aminoacyl tRNA inserts into the A site of the ribosome to decode triplet codons on mRNA,

Introduction

accompanied by a GTP-elongation factor complex. Subsequently, it forms a peptide bond with the polypeptide carried by peptidyl-tRNA in the P site. With the translocation of the ribosome, the deacylated tRNA exits from the E site, and the peptidyl-tRNA transitions from the A site to the P site, thereby leaving the empty A site available for a new aminoacyl-tRNA (GM 2000). The elongation persists until a stop codon (UAA, UAG, or UGA) occupies the A site, which serves as the termination signal recognized by the release factor, consequently halting protein synthesis (Figure 1.4). In prokaryotic cells, release factors consist of two proteins: release factor 1 (RF1), which recognizes UAG and UAA, and release factor 2 (RF2), which recognizes UGA and UAA (E. Scolnick 1968). Conversely, in eukaryotic cells, there is only one universal release factor (eRF-1) that recognizes all three stop codons (David S. Konecki 1977).

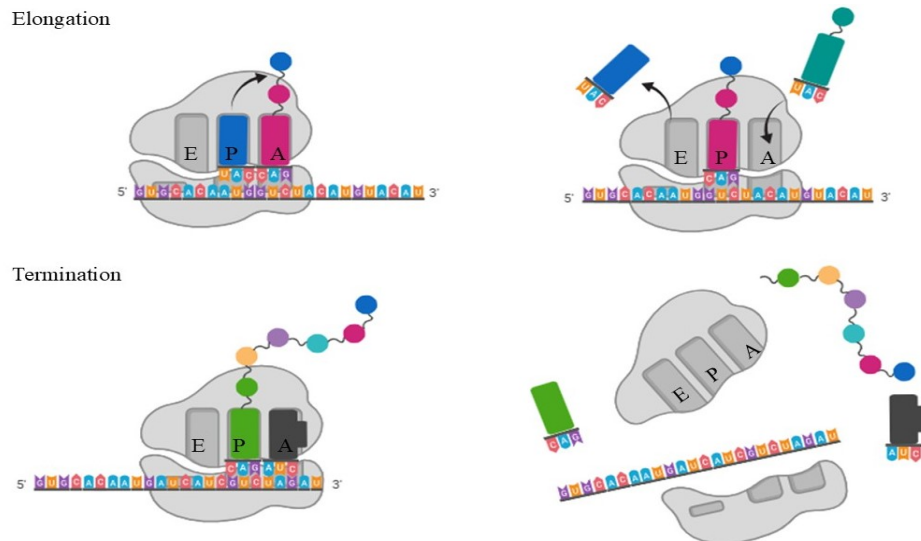


Figure 1.4 Cartoon visualization of the elongation and termination phases in protein synthesis

Letter A denotes the A site of the ribosome, known as the aminoacyl site, which is occupied by the incoming aminoacyl-tRNA except for the initiator methionyl tRNA. Letter P signifies the P site of the ribosome occupied by the peptidyl-tRNA or the initiator methionyl-tRNA, holding the growing polypeptide chain. Letter E indicates the E site of the ribosome, the exit site occupied by deacylated tRNA.

1.4.2.2 Release factor 1-deficient *E. coli* S30 lysate

In the orthogonal tRNA/tRNA-synthetase pair-mediated GCE technique, the amber codon (UAG) is frequently repurposed for the incorporation of unnatural amino acids during translation, serving as a target codon. However, it also functions as a termination signal recognized by release factor 1 (RF1) in *Escherichia coli*. Consequently, the competition between orthogonal tRNA and RF1 leads to the expression of truncated proteins in the in vitro system (Figure 1.5).

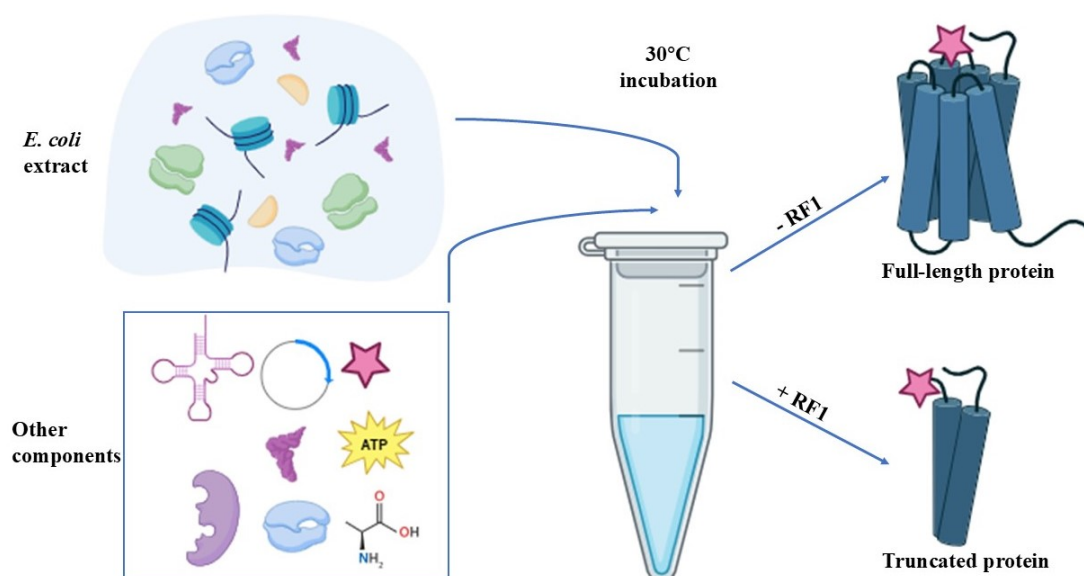


Figure 1.5 Scheme of the influence of release factor 1 on the in vitro protein expression system to incorporate unnatural amino acids

In the standard in vitro protein expression system, the extra addition of orthogonal tRNA/tRNA synthetase and unnatural amino acids is responsible for incorporating corresponding unnatural amino acids. At the same time, release factor 1 (RF1) from the prepared extract affects its efficiency, producing truncated proteins. “- RF1” means RF1-deficient extract. “+ RF1” signifies the extract possesses RF1.

To enhance the efficiency of unnatural amino acid incorporation, one proposed solution is to mitigate the impact of RF1, which is underpinned by the overlapping recognition of RF1 and RF2 at the UAA stop codon. To address

this question, *E. coli* S30 extracts derived from a temperature-sensitive RF1 variant US477 and XAC-RF were heated to partly deactivate RF1, enhancing unnatural amino acid incorporation into β -galactosidase (Lance E. Steward 1997) and *E. coli* dihydrofolate reductase (DHFR) and HIV-1 protease (PR) (Glenn F. Short 1999), respectively. Thereby, scientists are committed to studying the RF1-deficient in vitro system.

The PURE system omitting RF1 demonstrated the feasibility of UAAs incorporation by obtaining the full-length *E. coli* DHFR with the addition of chemically synthesized Val-tRNA^{sup} (Yoshihiro Shimizu 2001). The utility of Anti-RF1 antibodies in *E. coli*-based CFPS significantly improved the full-length esterase expression when adding the appropriate amount of suppressor tRNA^{Ser(CUA)} (Dmitry E. Agafonov 2005). This enhancement was observed in the in vitro expression of *gfp*^{amberX} when an RNA aptamer was used to inhibit the activity of *E. coli* RF1 (Shinsuke Sando 2007). RF1-free *E. coli* S30 extracts were also prepared by removing genetically labeled wild-type RF1 with chitin-binding domain (CBD) through a chitin column (Karin V. Loscha 2012).

In addition to the aforementioned techniques, establishing an RF1-deficient *E. coli* strain has initiated another pathway to develop the in vitro system. Given that 300 open reading frames (ORFs) conclude with an amber codon in *E. coli*, the knockout of *prfA* (which encodes RF1) significantly impacts the strain's growth and can be lethal (Takahito Mukai 2010), thereby greatly reducing the yield of full-length protein. To circumvent these disadvantages, it is prudent to substitute the amber codon TAG in essential genes with the synonymous TAA codon prior to the knockout of RF1 in *E. coli*. This strategy is corroborated by the *rEc.13.ΔprfA* strain, RFzero-iy strain, and the B95.ΔA strain, derived from BL21(DE3). The extract from the *rEc.13.ΔprfA* strain has demonstrated the ability to incorporate unnatural amino acids into superfolder green fluorescent protein (sfGFP) at multiple sites with the addition of transzyme (Seok Hoon Hong 2014). In the in vitro system utilizing the extract from B95.ΔA, when

supplemented with orthogonal tRNA and tRNA synthetase, the yield of full-length GFPS increased by 3-fold, 6-fold, and 18-fold for one, two, and three amber codons, respectively, compared to that based on BL21 (Eiko Seki 2018). This high incorporation efficiency was also observed in the in vitro system based on RFzoro-iy strain (Jiro Adachi 2019).

1.4.3 Membrane protein expression

Utilizing the in vitro protein expression system for membrane protein synthesis offers a platform conducive to investigating a broad spectrum of membrane proteins. This system's inherent flexibility facilitates the introduction of diverse reagents that replicate a hydrophobic environment, including detergents, liposomes, and nanodiscs (Zachary A. Manzer 2023). This configuration allows synthesized membrane proteins to integrate into artificial membranes after translation. The application of these membrane mimetics should give specific considerations due to their advantages and limitations, which will be discussed in the following.

1.4.3.1 Membrane protein expression with detergents

Detergents represent the primary alternative to membranes and exhibit partial compatibility with the in vitro protein expression system. Consequently, one strategy for membrane protein expression involves adding an appropriate detergent into the reaction mixture of Cell-Free Protein Synthesis (CFPS) to yield soluble protein-detergent micelle complexes directly. Alternatively, it may indirectly result in the formation of proteoliposomes through the application of a detergent-lipid mixture and the removal of detergent technique post-expression (Figure 1.6 A). Given this idea, an α -helical multidrug transporter found in *Escherichia coli*, EmrE, was solubilized effectively by 0.08% Dodecyl Maltoside (DDM) or 0.4% N-octyl- β -D-glucopyranoside after translation in the in vitro system (Yael Elbaz 2004). Three distinct topological membrane proteins, EmrE, Tsx, and G protein-coupled receptors (GPCRs), were high-efficiently

Introduction

solubilized by nonionic detergents such as Brij 35, Brij 58, Brij 78, and Brij 98 (Christian Klammt 2005).

Another approach involves the utilization of detergents or a detergent-lipid mixture to solubilize the expressed and precipitated membrane proteins (Figure 1.6 B). By this approach, precipitated EmrE, SugE, and TehA were solubilized using 1% DDM and subsequently incorporated into lipid membranes to generate proteoliposomes, following the removal of detergents using biobeads SM-2, which yielded generation efficiencies of approximately 80%, 80%, and 10%, respectively (Christian Klammt 2004).

Moreover, membrane proteins can be expressed through cell-free protein synthesis (CFPS) in conjunction with a mixture of detergents and lipids. This methodology was exemplified by the synthesis of three GPCRs: the β 2 adrenergic receptor (2AR), the muscarinic acetylcholine receptor M2 (M2), and the neurotensin receptor (NTR) within the in vitro system, which was augmented with 0.2% digitonin or 0.2% Brij35 and phospholipids (Goshi Ishihara 2005). Additionally, the synthesis of *Acetabularia* rhodopsin (ARII) was achieved in the in vitro system supplemented with phosphatidylcholine (PC) and 0.4% digitonin (Takashi Wada 2011).

The application of detergents is advantageous for solubilization and simplifies the purification of membrane proteins. However, there is no single universal detergent that solubilizes all membrane proteins; detergent screening remains an essential and time-consuming step in this approach. Additionally, to form the micelles, the amount of detergent must reach the Critical Micelle Concentration (CMC) in the in vitro system, whereas a high CMC adversely affects the yield of CFPS, resulting from its inhibition of transcription and translation in CFPS (Christian Klammt 2005). Except for these, the structurally monolayer of detergent-micelle is different from the bilayer structure of native liposomes and unable to replicate the complex membrane environment, which may affect the structure and function of membrane protein (Zachary A. Manzer 2023).

1.4.3.2 Membrane protein expression with liposomes

Another alternative to the membrane is the liposome, which is characterized as a spherical lipid bilayer containing an interior cavity separated from the external environment (Zachary A. Manzer 2023). This structure provides tunability in terms of composition, which affects the fluidity, rigidity, and elasticity of the liposome, and biophysical properties of liposomes such as size, shape, curvature, and the quantity of lamellae, thereby enabling the mimicry of both prokaryotic and eukaryotic cellular membranes (Nicholas S. Kruyer 2021). Liposomes are introduced directly into cell-free protein synthesis (CFPS) (Figure 1.6 C) in a one-step membrane protein synthesis process. Subsequently, certain expressed membrane proteins are spontaneously integrated into the liposomes, influenced by the liposome's composition and the hydrophobicity of the protein's secondary structure, such as the alpha-helix (Nicola J. Harris 2020). Alternatively, some proteins can be incorporated into the liposomes with additional reaction components, including the signal recognition particle (SRP) and translocon apparatus (Rapoport 2007).

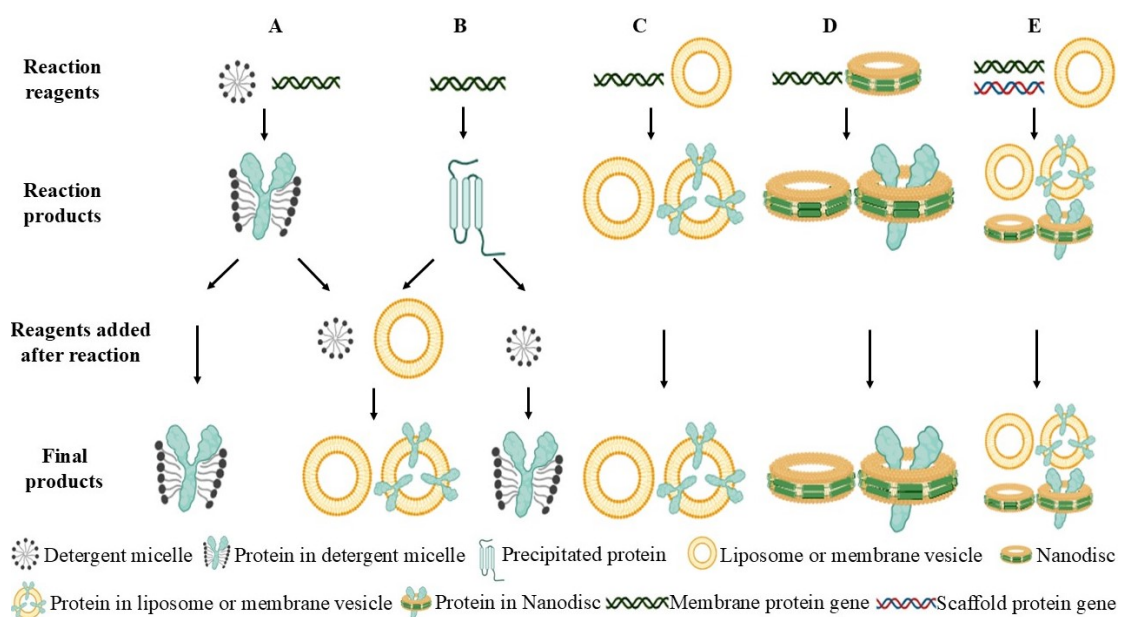


Figure 1.6 The scheme of membrane protein expression by the in vitro expression system, adapted from (Federico Katzen 2009)

Introduction

This scheme shows the main reagents added into and products involved in each of the approaches, except for the common components for the in vitro protein expression system. In all approaches, the gene encoding the membrane protein is added, while other reagents change depending on the specific strategy. Strategy A shows the membrane protein expression in the presence of detergents, where the translated membrane protein is solubilized by detergent micelles. When the additional reagent (liposome and detergent mixture) is added after the reaction, proteins can be reconstituted into liposomes to generate proteoliposomes from the detergent-protein micelle. Strategy B signifies the expression of membrane proteins as a precipitate. The precipitated membrane protein can be solubilized in detergent micelles by the addition of detergents or reconstituted into liposomes with the co-addition of detergents and liposomes after the reaction. Strategy C implies the membrane protein expression in the presence of liposomes or vesicles. Here, the translated protein is inserted into liposomes directly. Strategies D and E demonstrate the membrane protein expression in the presence of nano-discs. In strategy D, nano-discs are pre-fabricated, whereas both target and scaffold proteins are expressed in vitro to form protein-nanodisc complexes in situ in strategy E.

In vivo, a lipid matrix solvates membrane proteins, forming an intrinsic microenvironment. These compositionally and biophysically complex lipids significantly influence the interaction, localization, conformation, and functionality of membrane proteins (Lyman 2023). These influences are manifested in the in vitro protein expression system by intrinsic effects of compositions and physical properties of liposomes on membrane proteins' incorporation, folding, stability, and function.

1.4.3.2.1 Compositions of liposomes

Liposomes are composed of lipids. Each lipid has individual properties of headgroup charge, chain saturation, and chain length; therefore, the extruded liposomes possess special chemical and physical properties. Common lipids used in CFPS include 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) to form a neutral and unsaturated liposome, 1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DOPG) to form a negatively charged and unsaturated liposome, 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) or 1,2-Dimyristoyl-sn-glycero-3-phosphoglycerol (DMPG) to form a neutral and saturated liposome, and 1,2-dioleoyl-snglycero-3-phosphoethanolamine (DOPE), neutral and unsaturated lipid to form liposome with other lipids. When

different lipids mix, the extruded liposome could have different lateral pressure profiles and phase behavior (Figure 1.7).

The primary issue of membrane protein expression in the presence of liposomes by CFPS is how membrane proteins insert into the lipid membrane and whether they can perform insertion without any translocon apparatus, which various experiments answer. In the research on in vitro expression of bacterial α -helical membrane proteins, GlpG and DsbB (Nicola J. Harris 2017), GlpG and DsbB were inserted into liposomes without a translocon, confirming the proposed model of membrane protein incorporation based on thermodynamics, where TM helices are driven by the transfer free energy between aqueous and lipidic phases to insert into the membrane. Meanwhile, this incorporation was influenced by the composition of liposomes, because in five tested liposomes, GlpG preferentially was inserted into liposomes composed of DOPG and COPE in a molar ratio of 1:1, while DsbB reconstituted into DMPC liposomes with the highest efficiency. Therefore, the composition of the liposome is a significant factor in the spontaneous incorporation of membrane into liposomes.

The preference demonstrated by membrane proteins for lipids is corroborated by additional research. Cell-free expression of OprF utilizing six distinct types of liposomes revealed that when the liposome composition included cholesterol, DOPC, DOPE, and DMPA in a molar ratio of 2-4-2-2, OprF was able to reconstitute into liposomes in the largest quantities (Géraldine Mayeux 2021). The incorporation efficiency can be significantly diminished by the presence of differently charged lipids or those with varying phase transition temperatures. Specifically, the reconstitution efficiency of the net negatively charged Connexin-43 (Cx43) in liposomes composed of zwitterionic DOPC decreased as the proportion of negatively charged DOPG increased. This decline is attributed to DOPG reducing the interaction between the liposomes and Cx43. Furthermore, the efficiency also diminished with an increase in the content of

Introduction

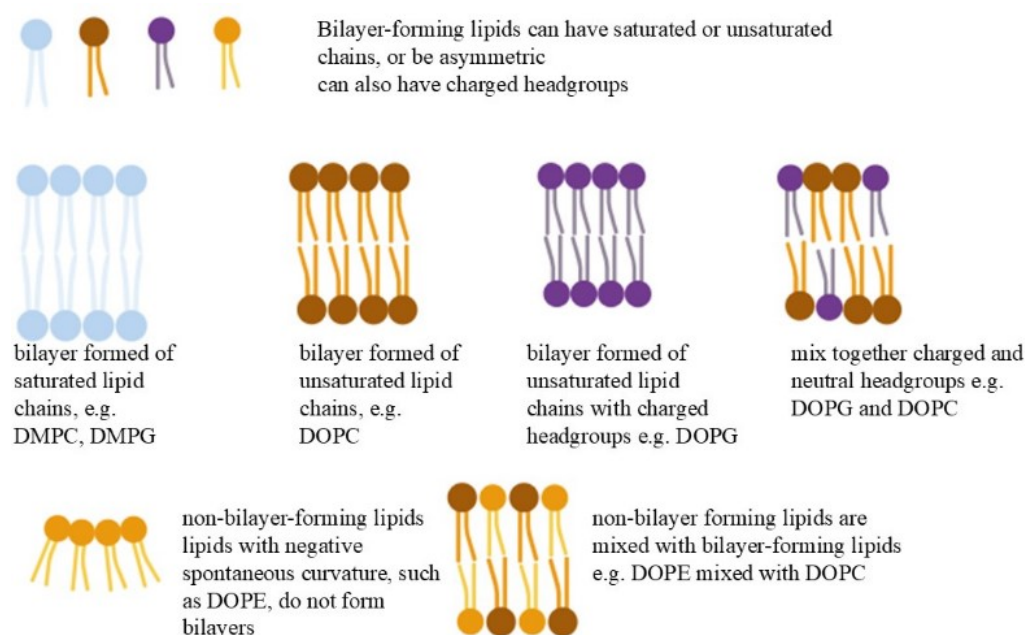


Figure 1.7 Different lipids form bilayers with different properties, adapted from (Nicola J. Harris 2022)

Lipids can be bilayer-forming (e.g. DMPC, DOPC, DOPG) or non-bilayer-forming (e.g. DOPE). They can have different headgroups, charged or neutral, and chains, saturated, unsaturated, or branched. Mixing different types of lipids produces bilayers with various chemical and physical properties.

high phase transition temperature zwitterionic DPPC, due to the reduced fluidity of DPPC within liposomes at room temperature (Yuki Moritani 2010). However, the bacterial leucine transporter (LeuT), characterized by a knotted structure, demonstrated no significant influence from the changes in lipids within liposomes. Nonetheless, the insertion efficiency experienced a substantial increase following the introduction of cardiolipin (CL) and translocon SecYEG in cell-free protein synthesis (CFPS). (Laura R. Blackholly 2022).

The second challenge regarding membrane protein expression in the presence of liposomes during CFPS pertains to the accurate folding of membrane proteins within liposomes, a phenomenon that remains inadequately understood. However, it is recognized that the folding process is highly contingent upon the surrounding lipid environment (Nicola J. Harris 2022). The in vitro folding research of reconstituting LacY into liposomes composed of tri-component lipids (DOPC, DOPE, and DOPG) showed that the recovery

efficiency of the unfolded helical structure decreased with the increase of DOPE content in DOPC/DOPG liposomes, but DOPE was required to attain maximum folding of LacY. Meanwhile, the correct folding is higher in the DOPG liposomes than in DOPC. Besides, the proper folding of GalP in DOPC bilayers also depended on the presence of DOPE (Booth 2017).

1.4.3.2.2 Biophysical properties of liposomes

Various methods may result in distinct biophysical properties of liposomes during their preparation. The alterations in diameter, shape, curvature, or quantity of lamellae within liposomes constitute significant factors in the investigation of membrane protein incorporation, folding, and functionality (Nicholas S. Kruyer 2021). The influence of the surface-area-to-volume ratio in liposomes was first examined twelve years ago. EmrE, a multidrug transporter derived from *E. coli*, was synthesized within giant unilamellar vesicles (GUVs) composed of POPC, with volumes ranging from 1 femtoliter (fL) to 1000 fL. This study demonstrated that the integration efficiency of EmrE increased as the volume of the GUVs decreased, resulting in a heightened surface-area-to-volume ratio in liposomes (Haruka Soga 2014). Furthermore, the morphology of liposomes emerged as a crucial factor in the assembly of membrane proteins. Following the synthesis of MreB, a bacterial analog of eukaryotic actin, within liposomes composed of egg phosphatidylcholine, the incorporation of phosphatidyl-ethanolamine-polyethylene glycol (PE-PEG) transformed the spherical liposomes into rod-like structures, thereby enhancing their polymerization (David Garenne 2020).

Overall, on one hand, liposomes present the advantages of enabling the expression of membrane proteins within a controllably hydrophobic lipid environment analogous to that of the cell membrane, facilitating the investigation of membrane proteins concerning their incorporation, folding, assembly, and interaction under native-like conditions. On the other hand, there are several drawbacks to consider when expressing membrane proteins in the

liposomes. These include the colloidal instability in biological fluids, the high cost associated with scaling up the in vitro membrane protein expression (Kiselev 2022), challenges related to protein purification due to their insolubility in aqueous solution, and non-specific interaction with other components in the in vitro system (Kamaruzaman 2023).

1.4.3.3 Membrane protein expression with nanodiscs

Nanodiscs are nanoscale, discoidal structures in which amphipathic molecules, designed lipid-like peptide detergents, and synthetic polymers are utilized to stabilize both artificial and native phospholipids. Two methods are available for incorporating membrane proteins into nanodiscs by the in vitro protein expression system. One method resembles the functionality of liposomes, wherein prepared nanodiscs are directly added to the in vitro expression system (Figure 1.6D), facilitating the production of functional membrane proteins (Ralf-Bernhardt Rues 2018). The other method involves supplementing the gene for amphipathic molecules and the target protein into the in vitro protein expression system (Figure 1.6E), thereby simultaneously co-translating amphipathic molecules and membrane protein to form the inserted membrane protein into the lipid bilayer (Roman Levin 2022).

Amphipathic molecules encompass amphipathic proteins that possess lipid-binding properties, including apomyoglobin (Kim 1992), zebrafish apolipoprotein A-I (zap 1) (Sourabh Banerjee 2008), membrane scaffold proteins (MSP) derived from human apolipoprotein A-I (apo A-I) (Timothy H. Bayburt 2002), a member of the saposin-like protein family, namely saposin A (Jens Frauenfeld 2016), and α -synuclein (α -Syn) (Cédric Eichmann 2016). Besides, amphipathic peptides are an alternative for forming nanodiscs, such as short filamentous amphipathic β -strand peptides (Houchao Tao 2013), three 18A peptides: rigid beltide-1, limited flexible beltide-2, and flexible beltide-3 (Andreas N. Larsen 2016), along with short amphipathic bi-helical peptides (NSPr) (Michael Luke Carlson 2018).

Introduction

Lipid-like peptide detergents, also referred to as peptergents, were developed by Shuguang Zhang's laboratory (Sylvain Vauthey 2002) to stabilize membrane proteins. The A₆D and V₆K mixture successfully stabilized glycerol-3-phosphate dehydrogenase (GlpD) and facilitated the solubilization of NADH peroxidase (Npx) (Joanne I. Yeh 2005). The A₆K variant provided stabilization for photosystem I (PS-I) at room temperature for at least three weeks (Patrick Kiley 2005). The variants I₆K₂, A₆K, V₆K₂, and V₆R₂ contributed to the stabilization of the multidomain protein complex photosystem I (PS-I) with varying efficiencies (Kazuya Matsumoto 2009). Additionally, ten peptide detergents were incorporated into the in vitro system to solubilize twelve olfactory receptors, exemplified by Brij-35 (Karolina Corin 2011).

Synthetic polymers are considered analogous to amphipathic molecules, generating controllable conformation after the balance between charge repulsion and hydrophobic interaction due to sensitivity to pH changes (S.R. Tonge 2001). They could extract membrane proteins from the native cell membrane, but also provide a native-like membrane platform to integrate expressed membrane proteins from an in vitro system. Styrene maleic acid (SMA), styrene as the hydrophobic residue and maleic acid as the hydrophilic residue, was the first amphipathic polymer to form nanodiscs containing dimyristoyl phosphatidylcholine (PC), successfully solubilizing bacteriorhodopsin (bR) and PagP (Timothy J. Knowles 2009). Compared with SMA, diisobutylene/maleic acid co-polymer (DIBMA), styrene-free and aliphatic, has a mild influence on lipid acyl-chain order, no effect on optical spectroscopy in the far-UV range, and no precipitation under the condition of low millimolar concentrations of divalent cations Mg²⁺ or Ca²⁺ (Abraham Olusegun Oluwale 2017). In contrast to SMA, poly (acrylic acid-co-styrene) (AASTY) copolymers have a controllable molecular weight and dispersity, contributing to the structural determination of membrane proteins in a native-like lipid-bilayer environment using single-particle cryo-EM (Anton A.A. Smith 2020). In

comparison to negatively charged SMA and DIBMA, the electroneutral Sulfo-SMA and Sulfo-DIBMA have no unspecific interactions with inserted membrane proteins or encapsulated lipids, but also no limitation of the application of charge-sensitive techniques to analyze membrane proteins (David Glueck 2022).

Compared with liposomes, the application of nanodiscs in CFPS provides advantages, including no requirement for unique devices to prepare nanodiscs, unlimited membrane protein incorporation due to access to both faces of the lipid bilayer, and simplified protein purification. Despite such merits, the disadvantages of nanodiscs should be considered when using them in CFPS. For example, the application of MSP is challenged by its UV absorbance, especially at 280 nm, due to the existence of aromatic amino acids tryptophan and tyrosine (Mark A. McLean 2020) and the need for detergents in nanodiscs preparation (Timothy H. Bayburt 2002). Disadvantages of SMA are instability in the presence of divalent cations or at low pH (Stefan Scheidelaar 2016), low affinity to Ni-NTA or Strep-Tactin matrices (Anton A.A. Smith 2020), and the interaction between lipids or membrane proteins, which is also the shortcoming of DIBMA and AASTY copolymers (Loretta Eggenreich 2023).

1.5 Assembly of PS1 and PEN2

γ -Secretase is an intramembrane aspartyl protease responsible for the intramembranous cleavage of various type 1 integral membrane proteins, including the amyloid β -protein precursor (APP) (Michael S. Wolfe 1999) and Notch receptor (Bart De Strooper 1999). During the sequential cleavage of APP, the amyloidogenic process refers to two sequential cleavages. The first cleavage is performed by β -site APP-cleaving enzyme-1 (BACE1), also known as β -secretase, generating the extracellular soluble APP β fragment (sAPP β) and the membrane-inserted C99 fragment (APPC99). Subsequently, the second cleavage occurs when APPC99 is cleaved by γ -secretase, resulting in the release of various short-length A β fragments into the extracellular space

(Wong 2011). Among these, the longer peptide A β 42 is considered toxic and is highly prone to aggregation, forming senile plaques, which are one of the neuropathological hallmarks of Alzheimer's disease (AD) (Takeshi Iwatsubo 1994). Hence, γ -secretase has drawn considerable attention for potential treatment strategies for AD.

After decades of study, it has been established that γ -secretase consists of 20 transmembrane domains (TMDs) (Linfeng Sun 2015) and is composed of four subunits, namely the catalytic component presenilin 1 (PS1) (Michael S. Wolfe 1999), the substrate receptor nicastrin (NCT) (Sanjiv Shah 2005), the complex stabilizer anterior pharynx 1 (Aph1), and presenilin enhancer 2 (PEN2) (Ross Francis 2002). The assembly of γ -secretase is intricate and poorly understood, although it is well-known that all four components are co-translationally inserted into the endoplasmic reticulum (ER). The stepwise overexpression of subunits of γ -secretase supports the hypothesis that NCT interacts with Aph1 to form a stable subcomplex, which then binds to the full-length PS1 holoprotein, producing a PS1-NCT-Aph1 trimeric intermediate, thereby decreasing the degradation of PS1 (Iwatsubo 2004; Matthew J. LaVoie 2003). Thereafter, the maturation of γ -secretase is initiated by the final binding of PEN2, facilitating the endoproteolysis of PS1 holoprotein (Nobumasa Takasugi 2003), where PS1 is cleaved between the sixth and seventh domains, yielding N- and C-terminal fragments (27-28 kDa PS-NTF and 16-17 kDa PS-CTF, respectively) that confer γ -secretase activity (Gopal Thinakaran and Samuel E. Gandy 1996). Alternatively, the formation of PS1-PEN2 and NCT-Aph1 subcomplexes may occur prior to the assembly of γ -secretase (Annaert 2008). This is supported by the observation that the detergent dodecyl- β -D-maltoside (DDM) mediates the dissociation of active γ -secretase (Patrick C. Fraering 2004), along with the in vitro endoplasmic reticulum (ER) budding assay based on semi-intact cells (Rosanne Wouters 2021).

Based on these two hypotheses, we assume the conformation of the PS1-PEN2 subcomplex will be different. Since the construction of γ -secretase has been resolved, we will try to analyze the interaction of the PS1-PEN2 subcomplex using an in vitro protein expression system combined with the incorporation of unnatural amino acids.

1.6 The aim of this study

This study aims to analyze the composition of the PS1-PEN2 subcomplex utilizing the photocrosslink assay following the introduction of unnatural amino acids into PS1. To this end, B95. Δ A is employed to conduct the experiments, which are designed using two distinct methodologies.

1. In vivo experiments will be conducted utilizing the pET/pEVOL expression system.
2. In vitro experiments will be performed employing two different approaches. The first approach is based on the application of the pET/pEVOL expression system. The second approach involves the addition of purified tRNA synthetase and extracted tRNA to facilitate the incorporation of unnatural amino acids into PS1.

For the second approach, the in vitro protein expression system must be established initially. Consequently, lysate preparation will be carried out under varying conditions, such as different media, cell harvesting times, cell disruption pressures, and overexpression of T7 RNA polymerase. Subsequently, the incorporation of unnatural amino acids (UAAs) will be executed with the assistance of auxiliary proteins, including TEV protease to remove the affinity tag of BpA-RS/AzF-RS; EGFP to quantify the yield of cell-free protein synthesis (CFPS); T7 RNA polymerase to enhance the translation efficiency of CFPS; and BpA-RS/AzF-RS to attach BpA/AzF to tRNA, thereby increasing the incorporation efficiency of BpA/AzF into PS1. Furthermore, the photocrosslinking conditions will be empirically optimized using the model

Introduction

protein sjGST. Ultimately, membrane proteins PS1 and PEN2 will be expressed through the established CFPS system and subjected to photocrosslinking using the established method, with various receiver molecules to simulate different hydrophobic environments.

2.1 Used instruments and materials

Table 2.1.1 Instruments used

Instruments	Company
Polymax 1040	Heidolph
IKA magnetic stirrer	IKA
250 µl Syringe	Avanti
Mini-extruder	Avanti
Thermomixer	Eppendorf
Avanti J-20 XP High-Speed Refrigerated Floor Centrifuge	Backman coulter
Trans-blot turbo transfer system	Bio-rad
Amersham Biosciences Electrophoresis Power Supply	Amersham
ChemiDoc MP imaging system	Bio-rad
Electrophoresis chamber	Bio-rad
Mini-Sub Cell GT Systems	Bio-rad
Visible spectrophotometer PRIM	Schott
C1000 touch thermal cycler	Bio-rad
Nanophotometer	IMPLEN
Analytical balance	KERN
UV-1800 Spectrophotometer	Shimadzu
Herasafe 2025 Biological Safety Cabinet	ThermoFisher
Incubator shaker	INFORS HT
Incubator	Memmert
Microcentrifuge	Eppendorf
Tube roller	Sunlab
Test tube shakers	Roth
Magbead separator	Cube Biotech
Dounce tissue homogenizer	VWR

Materials and Methods

Microfluidics M-110P	Newlife scientific
Crosslinker CL-1	Herolab
CLARIOstar plate reader	BMG LABTECH
Multiscreen® 96 well plate (MSSBNFX40)	Merck
Transparent NUNC PolySorp 96-well plate	ThermoFisher
Black 96-well Immuno plate	ThermoFisher
MINI INCU-shaker	Benchmark Scientific

Table 2.1.2 Materials used

Materials	Company
Acetyl phosphate lithium potassium salt	Sigma-Aldrich
Amino acids for cell-free expression	Sigma-Aldrich
Complete protease inhibitor cocktail	Roche
Adenosine 5'- triphosphate (ATP)	Roche
Cytidine 5'- triphosphate di-sodium salt (CTP)	Sigma-Aldrich
Guanosine 5'- triphosphate di-sodium salt (GTP)	Sigma-Aldrich
Uridine 5' - triphosphate tri-sodium salt (UTP)	Sigma-Aldrich
Folinic acid calcium salt	Sigma-Aldrich
Phosphoenol pyruvic acid monopotassium salt (PEP)	Sigma-Aldrich
Pyruvate kinase (PK)	Roche
RiboLock RNase inhibitor	Thermo Scientific
tRNA <i>E. coli</i> MRE 600	Roche
T7 RNA Polymerase	Roche
PageRuler Plus Prestained Protein Ladder	Thermo Scientific
FastDigest restriction enzyme	Thermo Scientific
T4 DNA ligase	Thermo Scientific
GeneRuler DNA ladder	Thermo Scientific

Materials and Methods

GeneJET Plasmid Miniprep Kit	Thermo Scientific
GeneJET Plasmid Maxiprep Kit	Thermo Scientific
GeneJET Gel Extraction Kit	Thermo Scientific
CHAPSO	Anatrace
Lipids	Anatrace
PVDF blotting membranes	Sigma-Aldrich
Phusion Hot Start II DNA Polymerase	Thermo Scientific
FastDigest DpnI	Thermo Scientific
GelRed Nucleic Acid Stain	Biotium Inc.
Xpress micro dialyzer MD 100	Scienova
Affinity resin (Ni-NTA resin, Ni-Indigo Magbeads)	Cube Biotech

2.2 Protein expression materials

Table 2.2.1 E. coli strains used

Strain	Genotype	Source	Purpose
Top10	F- mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80lacZ Δ M15 Δ lacX74 nupG recA1 araD139 Δ (ara-leu)7697 galE15 galK16 rpsL (StrR) endA1 λ -	Invitrogen	Plasmid amplification
B95. Δ A	<i>Escherichia coli</i> BL-21(DE3)-based host strain with no specific assignment of the UAG codon	RIKEN	Cell-free extract preparation
BL21 (DE3)	<i>E. coli</i> str. B F- ompT gal dcm lon hsdSB(r _B -m _B -) λ (DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+] K-12(λ S)	Agilent Technologies	Protein expression

Materials and Methods

BL21 Star	<i>E. coli</i> str. B F ⁻ <i>ompT hsdSB</i> (r _B ⁻ , m _B ⁻ ThermoFisher) <i>gal dcm rne131</i> (DE3)	Protein expression
-----------	--	-----------------------

Table 2.2.2 vectors

Vectors	Inducer	Resistance	DNA inserted
pET-15b	IPTG	Ampicillin	Presenilin-1 and its mutants (i) PEN2 and its mutants (ii) APH1 and its mutants (iii) NCT and its mutants
pET-27b	IPTG	Kanamycin	(iv) BpA-RS (v) AzF-RS (vi) sjGST and its mutants
pEVOL-pBpF	Arabinose	Chloramphenicol	BpA synthetase and BpA-tRNA
pEVOL-pAzF	Arabinose	Chloramphenicol	AzF synthetase and AzF-tRNA
pAR1219	IPTG	Ampicillin	T7 RNA polymerase
pET21a	IPTG	Ampicillin	EGFP
pRK793	IPTG	Ampicillin	TEV
pGEX-6P-2	IPTG	Ampicillin	sjGST

Table 2.2.3 Media for bacterial culture

Medium	Compositions
Lysogeny broth (LB)	Yeast extract (5 g/l) Tryptone (10 g/l) Sodium chloride (10 g/l)
Super Optimal Broth (SOB)	2% w/v tryptone 0.5% w/v Yeast extract 10mM NaCl 2.5mM KCl (pH6.8-7.0) add 10mM MgCl ₂ and 10mM MgSO ₄ to 1l of SOB media before use
2 x YT	Tryptone (16 g/l) yeast extract (10 g/l) NaCl (5 g/l)

Materials and Methods

2 x YTPG	Yeast extract (10 g/l)
	Bactotryptone (16 g/l)
	NaCl (5g/l)
	Glucose solution 300 (ml/l)
	Phosphate buffer 100 (ml/l)
	Phosphate buffer: KH_2PO_4 (29.9 g/l)
	K_2HPO_4 (91.3 g/l)
	Glucose solution: Glucose (66 g/l)
	Add 300 ml of Glucose solution and 100 ml of Phosphate buffer into 1l of medium before use.
M9TB	Bacto-Tryptone (10g/l)
	NaCl (5g/l)
	1x M9 buffer
	10x M9 buffer: NH_4Cl (10g/l)
	KH_2PO_4 (30g/l)
	Na_2HPO_4 (60g/l)
	Add 200 ml of 20% Glucose solution and 10 ml of 1M MgSO_4 into 1 L of 10x M9 buffer before use

All media are dissolved in Milli-Q water and autoclaved at 121°C for 20 minutes.

2.3 Buffers

Table 2.3.1 Buffers for gel electrophoresis

Buffers	Compositions
Agarose gel running buffer (50X TAE)	2 M Tris-base
	1 M Acetic acid
	50 mM EDTA (pH 8.0)
Denaturing PAGE gel running buffer (10X TBE)	890 mM Tris-base
	20 mM EDTA (pH 8.0)
	890 mM Boric acid
	95% (v/v) Formamide
2X denaturing loading buffer	18 mM EDTA (pH 8.0)
	0.025% (w/v) SDS
	0.05% (w/v) Bromophenol blue
	0.05% (w/v) Xylene cyanol
Gel fixation solution	1X TBE
	5% Methanol
	5% Ethanol

Materials and Methods

Denaturing PAGE Gel staining buffer	1X TBE 0.5 µg/ml Ethidium bromide 62 mM Tris 2 % (w/v) SDS
SDS-Sample buffer (5X)	5 % (v/v) β-Mercaptoethanol 20%(w/v) Glycerol 0.2 % (w/v) Bromophenol blue 1.5M Tris-HCl (pH 6.8) 8M urea
8M Urea lysis buffer	0.1% Glycerol 1% SDS 1mM DTT 25 mM Tris
SDS-PAGE running buffer	192 mM Glycine 0.1 % (w/v) SDS% 10% Phosphoric acid
Blue silver staining solution	10% Ammonium sulfate 1.2g/l Coomassie G-250 20% (v/v) Methanol 39 mM Glycine
Transfer Buffer for Western blot	48 mM Tris base 20% (v/v) Methanol 20 mM Tris base pH 7.6
TBS buffer	137 mM NaCl
TBS-T buffer	TBS buffer 0.1% (w/v) Tween 20
Blocking buffer	TBS-T buffer 5% (w/v) Non-fat milk

Table 2.3.2 Buffers for tRNA extract

Buffers	Compositions
Resuspend buffer	0.9% NaCl
Extraction buffer	50 mM Sodium acetate 10 mM Magnesium acetate (pH 5.0) Phenol
Acidic phenol solution	0.1% (w/v) 8-hydroxyquinoline 0.1 M Sodium acetate (pH 4.5)

Materials and Methods

	0.2% (v/v) β -mercaptoethanol
Total nucleic acids precipitation buffer	5M NaCl
rRNA precipitation buffer	1M NaCl
DNA precipitation buffer	0.3M NaOAc
Deacylation buffer I	1.5 M Tris (pH 9.0)
Deacylation buffer II	3M NaOAc (pH 5.0)

Table 2.3.3 Buffers for Competent Cell Preparation

Buffers	Compositions
TBF-1	30 mM NaOAc
	50 mM MgCl ₂
	100 mM NaCl
	10 mM CaCl ₂
	15 % (v/v) Glycerol (pH 6.0)
TBF-2	10 mM MOPS
	75 mM CaCl ₂
	10 mM NaCl
	15 % (v/v) Glycerol (pH 7.0)

Table 2.3.4 Buffers for E. coli Extract Preparation

Buffers	Compositions
50 X S30 A buffer	500 mM Tris-acetate (pH 8.2)
	700 mM Magnesium acetate
	30 mM KCl
	1 mM DTT
50 X S30 B buffer	500 mM Tris-acetate (pH 8.2)
	700 mM Magnesium acetate
	30 mM KCl
	1 mM PMSF
50 X S30 C buffer	500 mM Tris-acetate (pH 8.2)
	700 mM Magnesium acetate
	30 mM Potassium acetate
	0.5 mM DTT

Table 2.3.5 SDS-PAGE

Materials and Methods

Components	10% resolving gel (20ml)	12% resolving gel (20ml)	15% resolving gel (20ml)	5% stacking gel (5ml)
H ₂ O	7.9 ml	6.6 ml	4.6 ml	3.4 ml
30% Acrylamide	6.7 ml	8.0 ml	10.0 ml	0.83 ml
1.5 M Tris (pH 8.8)	5.0 ml	5.0 ml	5.0 ml	-
1 M Tris (pH 6.8)	-	-	-	0.63 ml
10 % SDS	0.2 ml	0.2 ml	0.2 ml	0.05 ml
10 % Ammonium Persulfate	0.2 ml	0.2 ml	0.2 ml	0.05 ml
TEMED	8 µl	8 µl	8 µl	5 µl
0.5% Trichloroethanol (TCE) (optional)	0.1 ml	0.1 ml	0.1 ml	-

Table 2.3.6 3.5% Denaturing PAGE Gel

Component	10ml	20ml
10xTBE	1 ml	2 ml
30% Acrylamide/bis-acrylamide (29:1)	1.167 ml	2.334 ml
10M Urea	6.5 ml	13 ml
10% Ammonium persulfate	0.08 ml	0.16 ml
TEMED	4 µl	8 µl
ddH ₂ O	1.249 ml	2.498 µl

Table 2.3.7 TEV Purification Buffers

Buffers	Compositions
	25 mM HEPES (pH 8.0)
	300 mM NaCl
Lysis buffer	10 mM Imidazole

Materials and Methods

	1 mM TCEP
	25 mM HEPES (pH 8.0)
	1 M NaCl
Wash buffer	40 mM Imidazole
	1 mM TCEP
	25 mM HEPES (pH 8.0)
	250 mM NaCl
Elution buffer	400 mM imidazole
	1 mM TCEP
	25 mM HEPES (pH 8.0)
	250 mM NaCl
Dialysis buffer	2 mM DTT
	25 mM HEPES (pH 8.0)
	250 mM NaCl
Storage buffer	2 mM DTT
	50% Glycerol
	500 mM Tris (pH 8.0)
	5 mM EDTA (pH 8.0)
10X reaction buffer	20 mM DTT

Table 2.3.8 BpA-RS and AzF-RS Purification Buffers

Buffers	Compositions
	25 mM HEPES (pH 8.0)
	300 mM NaCl
	10 mM Imidazole
Lysis buffer	5% Glycerol
	1 mM DTT
	Proteinase inhibitors: 1 μ M E64, 1 μ M Pepstatin A, 1 μ M Leupeptin, 1mM AEBSF, 1mM PMSF, 1mM Benzamidine
	25 mM HEPES (pH 8.0)
	500 mM NaCl
Wash buffer I for AzF-RS	20 mM Imidazole
	5% Glycerol
	1 mM DTT
	25 mM HEPES (pH 8.0)
	500 mM NaCl

Materials and Methods

Wash buffer for BpA-RS (Wash buffer II for AzF-RS)	30 mM Imidazole 5% Glycerol 1 mM DTT 25 mM HEPES (pH 8.0) 250 mM NaCl
Elution buffer	500 mM Imidazole 5% Glycerol 1 mM DTT 25 mM HEPES (pH 8.0)
SEC buffer	250 mM NaCl 10% Glycerol 1 mM DTT

Table 2.3.9 EGFP Purification Buffers

Buffers	Compositions
Lysis buffer	1X PBS buffer 10 mM Imidazole Proteinase inhibitors: 1uM E64, 1uM Pepstatin A, 1uM Leupeptin, 1mM AEBSF, 1mM PMSF, 1mM Benzamidine
Wash buffer I	1X PBS buffer 25 mM Imidazole
Wash buffer II	1X PBS buffer 50 mM Imidazole
Elution buffer	1X PBS buffer 250 mM Imidazole 20 mM Tris (pH 7.4)
SEC buffer	5 mM EDTA 100 mM NaCl

Table 2.3.10 Hydration Buffer

Compositions	Concentration (mM)
HEPES (pH 7.0)	20
NaCl	150

Materials and Methods

CaCl ₂	5
MgCl ₂	5

Table 2.3.11 T7 RNA Polymerase Purification Buffers

Buffers	Compositions
Lysis buffer	30mM Tris-Hcl (pH8.0) 10mM EDTA 10mM 2-mercaptoethano 5%(v/v) Glycerol 50mM NaCl 5%(w/v) Streptomycin sulfate
Precipitation buffer	30mM Tris-HCl (pH8.0) 10mM EDTA 1mM DTT 5%(v/v) Glycerol 50mM NaCl
Wash buffer	30mM Tris-HCl (pH=8.0) 10mM EDTA 1mM DTT 5%(v/v) Glycerol 0.4M NaCl
Elution buffer	30mM Tris-HCl (pH=8.0) 10mM EDTA 1mM DTT 5%(v/v) Glycerol 1.5M NaCl
Regeneration buffer	2M Guanidine Hcl
Storage buffer (Affi-gel Blue)	0.02M Sodium phosphate buffer (pH 7.4) 0.05% Sodium azide
Dialysis buffer	30mM Tris-Hcl(pH8.0) 10mM EDTA 50mM NaCl 1mM DTT 5% Glycerol
Gradient elution buffer	30mM Tris-Hcl(pH8.0) 10mM EDTA

Materials and Methods

	500mM NaCl
	1mM DTT
	5% Glycerol
Saturated ammonium sulfate buffer	ammonium sulfate (761 g/L)
	30mM Tris-HCl (pH=8.0)
Resuspension buffer	10mM EDTA
	1mM DTT
	5%(v/v) glycerol

Table 2.3.12 T7 RNA Polymerase Activity Test Buffers

Buffers	Compositions
	200mM Tris-HCl (pH 8.0)
Transcription buffer A	30mM MgCl ₂
	50mM DTT
	10mM Spermidine
	200mM Tris-HCl (pH 8.0)
Transcription buffer B	30mM MgCl ₂
	50mM DTT
	10mM BSA

Table 2.3.13 sjGST Purification Buffers

Buffers	Compositions
	20 mM HEPES (pH 8.0)
Wash buffer	300 mM NaCl
	20 mM Imidazole
	1 mM DTT
	20 mM HEPES (pH 8.0)
Elution buffer	300 mM NaCl
	400 mM Imidazole
	1 mM DTT

Table 2.3.14 PS1-N190X and PEN2 Purification Buffers

Buffers	Compositions
	20 mM HEPES (pH 8.0)
Resuspension buffer	300 mM NaCl

Materials and Methods

Wash buffer	10 mM Imidazole
	1% Fos14
	1 mM DTT
	20 mM HEPES (pH 8.0)
	300 mM NaCl
	20 mM Imidazole
Elution buffer	1% Fos14
	1 mM DTT
	20 mM HEPES (pH 8.0)
	300 mM NaCl
	400 mM Imidazole
	1% Fos14
	1 mM DTT

Table 2.3.15 Buffers for mass spectrometry

Buffers	Compositions
Activated trypsin buffer	100 ng/μl Trypsin in storage solution
	Stored at -20°C for up to 2 months
Digestion buffer	50 mM Ammonium Bicarbonate
	Store at 4°C for up to 2 months
Reducing buffer	100mM DTT
Alkylation buffer	100 mM Iodoacetamide
	Prepare this buffer before use

2.4 Analysis of γ -secretase

The Cryo-EM structure of γ -secretase with PDB entry 5FN5 was downloaded from the Protein Data Bank-RCSB PDB (<https://www.rcsb.org/>) in PDB format. This structure was utilized to analyze γ -secretase orientations in membranes using the PPM 2.0 Web Server from the OPM database (<https://opm.phar.umich.edu/>). Specific amino acids at the interface of subunits were identified as potential mutation sites for further analysis using the PyMOL software (<https://pymol.org/edu/>). The selected amino acids were substituted with BpA labeled with BPF in the protein sequences, employing the structural analysis software WinCoot. Subsequently, the reconstructed structure of γ -

secretase was analyzed further with the educational-use-only PyMOL to confirm the distance between BpA and the theoretically interactive amino acid C-H bonds. A distance of approximately 6 Å (Gyorgy Dorman 1994) was used for photo-crosslinking research.

2.5 Cloning

Polymerase Chain Reaction (PCR), developed by Nobel laureates Mullis and his colleagues in 1983 (Bachman 2013), is a remarkable laboratory technique for amplifying specific nucleic acid segments from various organisms or plasmids using appropriate primers (MT Rahman 2013). In classical cloning experiments, PCR products and the corresponding vector are digested by double restriction endonucleases at specific sites adjacent to or within the recognition sequence to generate sticky or blunt ends (Grossmann 1995). The 3'-hydroxyl and 5'-phosphoryl termini can then be catalyzed by DNA ligase (Lehman 1974) to synthesize phosphodiester bonds, constructing a recombinant plasmid for further research. In this work, all primers were designed using SnapGene Viewer software and synthesized by Integrated DNA Technologies (IDT) (see Appendix I: DNA and protein sequences and Appendix II: Primers). The primers were resuspended in sterile ddH₂O to a final concentration of 100 µM. The PCR assays are detailed in Tables 2.5.1 and 2.5.2. Each PCR program had an annealing temperature set to around 5°C below the T_m of the primers. The extension time for each amplicon was set at 15 seconds per 1 kb, based on the extension speed of Phusion Hot Start II DNA polymerase. Furthermore, PCR products were separated on a 0.7% agarose gel for 40 minutes at a voltage of 90 V, and potential bands were cut and extracted using the GeneJET Gel Extraction Kit from Thermo Fisher. The concentration of DNA was measured using a Nanophotometer from IMPLLEN. Subsequently, the resulting bands and corresponding vectors were digested by the relevant enzymes as mentioned in Table 2.5.3 at 37°C for 1 hour, then purified using the GeneJET Gel Extraction Kit from Thermo Fisher. The purified

Materials and Methods

bands were used to perform a ligation assay. The ligation reaction was set up as indicated in Table 2.5.4 and incubated for 10 minutes at 22°C. Next, transformation and sequencing were performed as described in the following section. The 10 µl of the ligation products were incubated with 100 µl of Top 10 competent cells in an ice bath for 30 minutes and then heat-shocked for 90 seconds at 42°C. After supplementing with 500 µl of LB media, the culture was shaken at 37°C and 500 rpm for 1 hour, followed by centrifugation at 20°C and 5000 rpm with the FA-45-48-11 rotor for 10 minutes to discard 400 µl of supernatant, leaving 200 µl behind. The remaining solution was spread on an LB agar plate containing the appropriate antibiotics and incubated overnight at 37°C. To evaluate the reconstructed plasmid, 3 colonies were picked individually into 4 mL of LB media with corresponding antibiotics and shaken overnight at 37°C and 130 rpm. The bacterial liquid was then centrifuged at 6000 rpm with the FA-45-48-11 rotor, 4°C for 10 minutes to harvest cells, followed by plasmid extraction using the GeneJET Plasmid Miniprep Kit. After double enzyme digestion with the relevant enzymes, the potential plasmid was confirmed through DNA sequencing conducted by Microsynth Seqlab. The following section details the process of constructing three specific expression vectors.

Table 2.5.1 PCR reaction

Components	50 µl reaction	Final concentration
5X Phusion HF buffer	10 µl	1X
10 mM dNTPs	1 µl	200 µM
Forward primer	0.5 µl	0.5 µM
Reverse Primer	0.5 µl	0.5 µM
Template DNA	1 µl	1 ng
DMSO	1.5 µl	6%
Phusion Hot Start II DNA polymerase	0.5 µl	0.02 U/µl

Materials and Methods

ddH₂O 35 µl

Table 2.5.2 PCR program

Steps	Temperature	Duration	No. of cycles
Initial denaturation	98°C	30s	1
Denaturation	98°C	10s	30
Annealing	55-72°C	30s	
Extension	72°C	15s per kb	
Final extension	72 °C	5min	1

Table 2.5.3 Double enzyme digestion

Components	50 µl reaction system
10x Fast Digest Green buffer	5 µl
DNA segment/vector	1 µg
Fast Digest Enzyme I	2 µl
Fast Digest Enzyme II	2 µl
ddH ₂ O	Up to 50 µl

Table 2.5.4 Ligation reaction

Components	20 µl reaction system
T4 DNA ligase buffer	2 µl
Linearized vector	50 ng
DNA segment	200 ng
T4 DNA ligase	0.5 µl
ddH ₂ O	Up to 20 µl

2.5.1 BpA-RS and AzF-RS expression vector

To construct plasmids pET27b-His-TEV-BpA-RS or pET27b-His-TEV-AzF-RS, primers His-BpA-RS-F and His-BpA-RS-R or His-TEV-AzF-RS-F and His-TEV-AzF-RS-R were employed to amplify the sequence of the BpA or AzF synthetase gene from the original plasmids pEVOL-pBpF or pEVOL-AzF, generating the His-BpA-RS or His-AzF-RS fragment. This was followed by NdeI and XhoI digestion and ligation with pET27b to produce plasmid pET27b-His-BpA-RS. After sequencing, the confirmed plasmid pET27b-His-BpA-RS was amplified with primers His-TEV-BpA-F and His-TEV-BpA-R to insert the TEV fragment between His and BpA-RS. DpnI digestion, as shown in Table 2.5.5, was conducted to cleave the template plasmid by incubating for one hour at 37°C. Subsequently, transformation and sequencing were performed.

Table 2.5.5 DpnI digestion

Components	40 µl reaction system
10x reaction buffer	4 µl
DNA segment	30 µl
DpnI	1 µl
ddH ₂ O	Up to 40 µl

2.5.2 tRNA expression vector

To generate an expression vector for tRNA to attach to BpA or AzF covalently, the pEVOL-pBpF was digested with BglII and PstI to remove the BpA-RS gene, followed by T4 DNA polymerase treatment to generate blunt ends and ligation with T4 DNA ligase. The ligation, pEVOL-BpA-tRNA, was transferred into the Top 10 competent cells for plasmid amplification.

2.5.3 Site-directed mutagenesis

To introduce unnatural amino acids into γ -secretase or sjGST, the specified nucleotides of the subunit in γ -secretase or sjGST were mutated to TAG using a site-directed mutagenesis technique. Based on the PCR method, this technique requires specially designed forward and reverse primers that are centered on the desired base changes and overlap completely at the 5' end, or one of the primers is replaced with TAG at the desired base without overlap to amplify the entire plasmid with the inserted mutation. Since the original plasmid grown in *E. coli* is methylated, it can be digested by the restriction endonuclease DpnI (Bachman 2013).

This work designed primers for nucleotides longer than 30 bp, among which at least 15 bp nucleotides matched perfectly at both 5' ends. After PCR, the product was digested by DpnI, as shown in Table 2.5.5, with incubation for 1 hour at 37°C. The resulting product was transferred into the Top 10 competent cells for selection and sequencing.

2.6 Protein expression and purification

Various heterologous expression systems exist to produce recombinant proteins, including *E. coli*, yeast, mammalian cell lines, insect cell lines, and cell-free reaction systems. Among these, *E. coli* is the primary host expression system due to its ease of manipulation, rapid growth, cost-effectiveness, availability of diverse vectors, and high expression levels of target proteins (Bilgimol C Joseph 2015). Although the *E. coli* expression system encounters challenges, such as low expression levels of proteins that may be toxic to *E. coli* and low yields stemming from a low growth rate due to unregulated basal expression of target proteins, these issues are addressed by the widely accepted technique of cell-free reaction systems. In this study, both systems were utilized for protein expression. Protein purification is a vital technique for isolating the desired protein from cell extracts or tissue homogenates,

employing a high-resolution bioseparation method known as chromatography. This includes various types such as affinity chromatography, ion exchange, size exclusion, hydrophobic interaction, and reverse-phase HPLC (Labrou 2014). The most effective and specific method is affinity chromatography, introduced by Cuatrecasas and colleagues in 1968, which plays a crucial role in protein purification (Ana C.A. Roque 2007) due to the highly specific recognition between the target protein molecule and its complementary ligand. NTA (nitrilotriacetic acid), a new quadridentate chelating adsorbent, is covalently attached to oxirane-activated agarose to create a gel that can be charged with metal ions such as Ni^{2+} , Cu^{2+} , or Zn^{2+} (E. Hochuli 1987). Among these, Ni-NTA resin is the most commonly used commercial resin for purifying recombinant proteins with six consecutive histidine residues, either at the amino or carboxyl terminus, due to its high purification efficiency. In the Ni-NTA resin, the two remaining ligand positions in the coordination sphere of Ni^{2+} interact tightly with the 6xHis tag of the target protein, which remains unaffected even when washing away unwanted proteins under stringent conditions (Harwood 1994). In comparison, another commercial resin, Ni-Indigo, can tolerate high concentrations of chelators such as 20 mM DTT or EDTA; this resin is also used to purify proteins treated with TEV in this work. Size exclusion chromatography (SEC) is known by various names, including molecular sieve chromatography, steric-exclusion chromatography, liquid exclusion chromatography, restricted-diffusion chromatography, gel filtration chromatography, and gel permeation chromatography. Since 1995, when Boerje Lindqvist and Torsten Stograde first separated amino acids by SEC using a column packed with starch (Boerje Lindqvist 1995), the SEC technique has been widely adopted as the final step in protein purification, based on the relative size or hydrodynamic volume of proteins, taking into account the pore size of the packing (Howard G. Barth 1994). Theoretically, when protein samples are introduced into a column filled with spherical porous particles, the largest molecules are excluded from the porous medium and eluted first, while smaller molecules are eluted

subsequently in an orderly decreasing size (Szabolcs Fekete 2014). Moreover, a calibrated curve can be generated after running known molecular weight proteins or peptides, which allows for the estimation of the molecular weight of unknown molecules (Edmund R.S. Kunji 2008). This feature enhances the utility of SEC in the separation and quantification of trimers, dimers, oligomers, or aggregates of target proteins, and even facilitates the purification of membrane proteins. Although matrices such as Superdex, Superose, Suphadex, or Sephacryl exhibit minimal interaction with detergent-solubilized membrane proteins, selecting an appropriate separation range for the column is critical and should be based on the molecular mass of the membrane protein (Edmund R.S. Kunji 2008). In this study, Superose 6 and Superdex 200 columns were used to separate the PS1-PEN2 complex from *E. coli* extract and to purify EGFP, respectively.

2.6.1 In vivo protein expression system

To incorporate BpA into the membrane protein PS1, the in vivo protein expression system derived from the B95.ΔA strain was used. Different auxiliary proteins and tRNA were expressed and purified by the in vivo protein expression system based on the BL21(DE3) strain to establish the in vitro expression system.

2.6.1.1 Auxiliary proteins

Auxiliary proteins, TEV, EGFP, BpA-RS, AzF-RS, and T7 RNAP, were expressed and purified from *E. coli*. The method of expression and purification will be described in the following sections

2.6.1.1.1 TEV, EGFP, BpA-RS, and AzF-RS

Soluble proteins, including TEV, EGFP, BpA-RS, and AzF-RS, were expressed and purified from *E. coli*. The confirmed plasmid was transformed into expression-competent BL21 (DE3) cells using the method mentioned in the 2.5

Materials and Methods

cloning section. After incubating overnight at 37°C, a colony was transferred into 5 mL of LB media containing the appropriate antibiotics, shaking for about 8 hours at 37°C and 120 rpm (Infors HT Multitron Pro Triple Incubator Shaker). One percent of fresh bacterial LB medium was added to 50 mL of 2xYT medium supplemented with the appropriate antibiotics and 2% glucose and incubated overnight at 37 °C and 120 rpm (Infors HT Multitron Pro Triple Incubator Shaker). The overnight culture was diluted 100-fold with 200 mL of 2xYT medium containing the appropriate antibiotics, shaking at 120 rpm and 37°C. When the optical density at OD600 reached approximately 0.5, 1mM IPTG was added to the bacterial culture, which was then allowed to grow at 28°C and 120 rpm for about 5 hours. The cells were harvested by centrifugation at 4500 rpm and 4°C for 30 minutes with the Beckman JLM8.1 rotor. The pellet was flash-frozen using liquid nitrogen and stored at -80°C for further experiments. One gram of cell pellets was resuspended in 10 mL of lysis buffer and homogenized with a Dounce tissue homogenizer. The resulting homogeneous solution was supplemented with a protein inhibitor and subjected to microfluidization twice using an M-110P microfluidizer at 15,000 psi. The lysate was centrifuged for 45 minutes at 50,000 x g and 4°C to remove unbroken cells and inclusion bodies. Subsequently, the supernatant was filtered through a 0.22 µm filter and mixed with Ni-NTA resin that had been equilibrated sequentially with 5 column volumes of ddH₂O and 2 column volumes of lysis buffer at 4°C for 4 hours or overnight. Non-binding proteins were washed away using 10 column volumes of wash buffer, and the binding proteins were eluted with 2 column volumes of elution buffer. The eluted fraction was concentrated using a protein concentrator with the appropriate MWCO, performed by centrifugation with a swing-bucket rotor A-4-44 at 4°C and 3000 rpm for 10 minutes at a time until the volume of the eluted fraction decreased to 500 µl. The concentrated samples were then injected into a Superdex 200 column equilibrated with SEC buffer to perform size exclusion chromatography. All fractions from the two-step purification were verified by SDS-PAGE and Western blot assays.

2.6.1.1.2 T7 RNA polymerase

Two strategies, including two steps, were tried to purify T7 RNA polymerase based on the previous method (Parichehre Davanloo 1984). The first step of the two strategies was performed using the same method: Affi-Gel Blue chromatography. Then, DEAE Sepharose chromatography and ammonium sulfate precipitation methods were employed to complete the second step, respectively.

2.6.1.1.2.1 T7 RNA Polymerase Lysate Preparation

Two methods were exploited to purify T7 RNA polymerase. First, using the heat-shock method mentioned in the 2.5 cloning section, pAR1219 was transformed into BL21(DE3) competent cells. One colony on the LB agar plate supplemented with 100 ng/μl ampicillin was picked up into 5 ml of LB media containing 100 ng/μl ampicillin. After overnight incubation at 37°C and 120 rpm with the INFORS HI incubator shaker, 1% of the bacterial solution was inoculated into 5 ml of M9TB media supplemented with 100 ng/μl ampicillin, and continuously incubated for around 6 hours at 37°C and 120 rpm. 1% of the 6h culture was inoculated into 50 ml of M9TB media containing 100 ng/μl ampicillin and incubated overnight under the same conditions. Hereafter, 1% of the overnight culture was inoculated into 3 L of M9TB media containing 100 ng/μl ampicillin, growing at 37°C and 120 rpm to OD₆₀₀ = 0.5 before induction. 1 mM IPTG was added to the culture to overexpress T7 RNA polymerase for approximately 4 hours at 30°C and 100 rpm. Afterward, cells were harvested by centrifugation for 40 mins at 4°C and 4500 rpm with JLM8.1 rotor, and stored at -80°C after flash-frozen by liquid nitrogen. 1g of pellets was resuspended with 30 ml of lysis buffer, disrupted twice by the cell disrupter Microfluidics M-110P at 1.8 kbar, followed by centrifugation for 20 mins at 4°C and 38,000 xg. The supernatant was mixed in the ice bath with precipitation buffer slowly and thoroughly until the final concentration of streptomycin sulfate was 0.2%. After

centrifugation for 15 mins at 4°C and 27000 xg, the supernatant was collected and filtered through a 0.2 µm filter for further purification.

2.6.1.1.2.2 Affi-gel Blue and DEAE Sepharose

The filtered supernatant was loaded into the pre-packaged Affi-gel Blue column after being equilibrated with lysis buffer. After the contaminated proteins were washed with wash buffer, the target protein was eluted with elution buffer. Samples eluted in different peaks from Affi-gel blue were prepared for SDS-PAGE and dialyzed twice at 4°C against a 50-fold volume dialysis buffer, one for 4h and the other overnight. After elution, the Affi-gel blue column was washed with regeneration buffer and stored in the storage buffer. The dialyzed sample was loaded into the packaged DEAE column equilibrated with dialysis buffer and eluted with gradient elution buffer. 5 µl of Samples from different peaks were prepared for SDS-PAGE, while others were flash-frozen in liquid nitrogen and stored at -80°C.

2.6.1.1.2.3 Affi-gel Blue and ammonium sulfate precipitation

The eluted sample from the Affi-gel Blue column was precipitated by slowly adding saturated ammonium sulfate buffer to a final concentration of up to 20%. After centrifugation for 15 mins at 4°C and 10000 rpm with FA-45-6-30 rotor, the precipitate was resuspended in the storage buffer, followed by flash-freezing with liquid nitrogen and stored at -80°C. The resultant supernatant was precipitated by being supplemented with saturated ammonium sulfate buffer to a final concentration of 30%, followed by centrifugation. This procedure was repeated to supplement saturated ammonium sulfate to the final concentrations of 40%, 50%, and 60%. Each precipitated sample was prepared for SDS-PAGE.

2.6.1.1.2.4 T7 RNA polymerase activity assay

pET21b-EGFP was digested with XhoI at 37°C for 2 hours to generate a linearized plasmid. After terminating the reaction at 65°C for 10 minutes, the linearized plasmid was separated by gel electrophoresis and extracted from the agarose gel using the GeneJET Gel Extraction Kit. Subsequently, the

transcription experiment was conducted as shown in Table 2.6.1, completed by incubating for 2 hours at 37°C, followed by termination with the addition of 0.5 µl of 0.1M EDTA solution. The resultant samples were mixed with 2x denaturing loading buffer, heated for 5 minutes at 94°C, and loaded onto the denaturing PAGE, which was pre-run for 1 hour at a voltage of 50 V/cm, with pockets rinsed with 1x TBE buffer. After running for approximately 3 hours under the same voltage, the denaturing PAGE gel was fixed twice in the gel fixation solution, each time for 5-10 minutes, and then incubated in Denaturing PAGE Gel Staining buffer for 15 minutes. Afterwards, the denaturing PAGE gel was imaged using the ChemiDoc MP imaging system.

Table 2.6.1 Transcription Experiment

Components	20 µl reaction system
Linear pET21-EGFP	50ng
RNase inhibitor	0.1U/µl
NTPs(A/U/G/C)	2mM
5x Transcription Buffer	4µl
T7 RNAP	1mg/ml-0.05mg/ml
ddH ₂ O	Up to 20µl

2.6.1.2 Auxiliary tRNA

The confirmed plasmid pEVOL-BpA-tRNA was transformed into BL21(DE3). After overexpression, the total tRNA containing orthogonal tRNA for BpA and AzF was extracted. All buffers used in the following sections refer to Table 2.3.2.

2.6.1.2.1 Harvest of cells

After transformation and overnight incubation at 37°C, the colony was picked into 5 ml of LB medium containing 34 µg/ml chloramphenicol, growing at 37°C and 120 rpm for around 6 hours in the MINI Benchmark INCU-shaker. 1% of 6h

media was added to 50 ml of 2xYT medium supplemented with 34 µg/ml chloramphenicol, incubating overnight under the same conditions, followed by 100 times dilution into 200 ml of 2xYT medium containing 34 µg/ml chloramphenicol, incubation at 37°C and 120 rpm with the INFORS HI incubator shaker before the OD600 reached to 0.5. Afterwards, the culture continuously grew under the same conditions for approximately 4h. Cells were harvested at 4500 rpm and 4°C for 30 min with the Beckman JLM 8.1 rotor. Every 5g of cell pellets was resuspended with 30 ml of resuspend buffer, followed by centrifugation at 5000 rpm and 4°C for 30 mins using the FA-45-6-30 rotor. The pellet was snap-frozen with liquid nitrogen and stored at -80°C for further extraction.

2.6.1.2.2 Preparation of acidic phenol

To eliminate the effects of RNase, all used materials were RNase-free, pipettes were wiped with DEPC water, and operators were masked and gloved before doing experiments. Crystal phenol was placed at room temperature shortly and then melted at 68°C in the water bath. The melted phenol was mixed with 0.1% 8-Hydroxyquinoline, not only a yellow antioxidant to inhibit RNase partially, but also a weak chelator of metal ions. The resultant phenol was extracted 3 times with 1.0 M sodium acetate at pH 4.5, in the same volume as the melted phenol, followed by 3 times extraction with 0.1 M sodium acetate at pH 4.5 and 0.2% beta-mercaptoethanol solution, until the pH of the aqueous phase is around 4. The acidic phenol solution was stored at 4 °C for up to 1 month.

2.6.1.2.3 Extraction of total tRNA

The method was adapted from the previous method (Irem Avcilar-Kucukgoze 2020). 10g of pellets were resuspended with 18 ml of extraction buffer and shaken with 17.2 ml of acidic phenol solution. The emulsion was incubated for 30 mins at 37°C, and centrifuged for 15 mins at 4°C and 5000 rpm using the FA-45-6-30 rotor to separate different phases. The aqueous phase was pipetted into

a new tube gently and carefully to avoid the aspiration of the interface; meanwhile, the phenol phase was extracted again with 14 ml of extraction buffer. After centrifugation, the second-extracted aqueous phase was collected into the same tube. The collected aqueous phase was mixed with the total nucleic acids precipitation buffer to a final concentration of 0.2 M NaCl and the same volume of isopropanol, stored at room temperature for 10 minutes to precipitate all nucleic acids. The resultant solution was centrifuged for 15 mins at 14,500 xg and 20°C. The pellet was washed with cold 70% ethanol and air-dried for a few minutes, followed by resuspension with 15 ml of cold rRNA precipitation buffer and centrifugation for 20 minutes at 9,500 xg and 4°C to precipitate rRNA from the nucleic acids. The resultant supernatant was cooled with 30 ml of cold 100% ethanol and stored at -20°C for 30 mins. Hereafter, centrifugation was followed at 4°C and 14500 xg for 5 mins to precipitate remaining nucleic acids, including DNA and tRNA. Afterwards, the pellet was washed with cold 70% ethanol and air-dried for 5 mins, followed by resuspension with 6 ml of DNA precipitation buffer overnight at room temperature. The overnight nucleic acid solution was mixed with 3.4 ml of isopropanol and incubated at 20 °C for 10 mins, followed by centrifugation for 5 mins at 14,500 xg and 20°C to precipitate the DNA. The resultant supernatant was collected and mixed with 0.95 volume of isopropanol, incubated at -20°C for 30 mins. After centrifugation for 15 mins at 14,500 xg and 4°C, the white pellets were washed with cold 70% ethanol, air-dried for 5 mins, and resuspended with 200 µl of DEPC-treated water. The resultant tRNA solution was deacylated immediately using the following method.

2.6.1.2.4 Deacylation of tRNA

The tRNA solution was incubated with 14µl of deacylation buffer I at 37 °C for 45 mins, followed by being supplemented with 21.4µl of deacylation buffer II and 636 µl of cold 100% ethanol progressively. After incubation at -80°C for 30 mins, the solution was centrifuged for 25 mins at 16100 xg and 4°C. Then the

pellet was washed with cold 70% ethanol, air-dried for a few minutes, followed by being dissolved into 20µl of sterilized Milli-Q water and stored at -80°C for further experiments. The Nanophotometer was employed to determine the concentration of tRNA, and the denaturing PAGE gel was used to evaluate the purity of tRNA.

2.6.1.3 Membrane protein

To site-directly incorporate BpA into the membrane protein PS1 by the B95.ΔA strain, competent cells were prepared, followed by transformation and expression.

2.6.1.3.1 Competent cell

Membrane proteins, PS1 and PS1-N190X, were expressed using B95.ΔA. To prepare competent B95.ΔA cells, the frozen strain was streaked on an LB agar plate. After an overnight incubation at 37°C, a colony was picked and transferred into SOB media to incubate at 30°C and 120 rpm overnight. 1% of the overnight preculture was inoculated into fresh SOB media under the same incubation conditions until the OD600 reached 0.4. Subsequently, 50 mL of the bacterial solution was transferred into a Falcon tube and incubated in an ice bath for 15 minutes, followed by centrifugation at 4°C and 3500 rpm for 15 minutes with the Beckman JLM 8.1 rotor. After discarding the supernatant, the pellet was resuspended in 10 mL of ice-cold TBF-1 buffer and incubated on ice for 15 minutes, which was repeated once. The pellet was then resuspended in 2 mL of ice-cold TBF-2 buffer and incubated on ice for 15 minutes, followed by transferring an aliquot of 100 µl of cells into pre-cooled Eppendorf tubes. After being flash-frozen with liquid nitrogen, all cells were stored at -80°C.

2.6.1.3.2 Expression of PS1 and truncated PS1-N190X

To express PS1 and PS1-N190X, the confirmed plasmids, pET15b-His-PS1 and pET15b-His-PS1-N190X, were transformed separately into competent B95.ΔA cells using the aforementioned method in the 2.5 cloning section.

Proteins were over-expressed in 2xYT media supplemented with 100 mg/L ampicillin by adding 1 mM IPTG when the optical density of the bacterial solution was approximately 0,5. After incubation at 30°C and 90 rpm for 4 hours with the incubator, cells were harvested by centrifugation at 4500 rpm and 4°C for 40 minutes with the Beckman JLM 8.1 rotor. Cell pellets were then flash-frozen in liquid nitrogen and stored at -80°C.

2.6.1.3.3 Expression of full-length PS1-N190X

To express full-length PS1-N190X, pET15b-His-PS1-N190X was co-transformed with pEVOL-pBpF into competent cell B95.ΔA. After incubation overnight at 37°C, one co-transformed colony was picked up to 5 ml of LB media supplemented with 100 mg/L ampicillin and 50 mg/L chloramphenicol, shaking for around 8 hours at 37°C and 120 rpm. 1% of fresh bacterial LB solution was inoculated into 50 ml of 2xYT medium supplemented with identical antibiotics and 2% glucose, incubating overnight at 37°C and 120 rpm with an INFORS HI incubator shaker. The overnight culture was diluted 100 times with 200 ml of 2xYT medium containing appropriate antibiotics and grown at 120 rpm and 37°C to OD₆₀₀ = 0.6 before induction. Induction was processed by adding 1 mM IPTG, 1 mM BpA, and 0.02% arabinose to the culture, which continued growing at 30°C and 100 rpm for 4-5 h before the culture was harvested at 4500 rpm and 4°C for 40 min with the Beckman JLM 8.1 rotor. Cell pellets were flash-frozen with liquid nitrogen and stored at -80°C for further experiments. The control trial was carried out similarly to the expression of full-length PS1-N190X, except for the co-transformation of pET15b-His-PS1 with pEVOL-pBpF into competent cell B95.ΔA (Jason W. Chin 2002a).

2.6.1.3.4 SDS-PAGE and Western Blot

0.1g of cell pellets from PS1, truncated PS1-N190X, full-length PS1-N190X, and PS1 with pEVOL-pBpF were resuspended in 1 ml of 8 M urea buffer and incubated at room temperature for 1h. All resultant samples were heated at 40°C

for 30 minutes after being mixed with 5xSDS-sample buffer and confirmed by SDS-PAGE and Western blot assays.

2.6.2 In vitro protein expression system

Two strategies were used to incorporate BpA into the membrane protein PS1. The first one was done by combining the pET/pEVOL co-transformed system and the in vitro protein expression system. The second one was completed by the addition of purified BpA-RS and extracted BpA-tRNA into the in vitro expression system.

2.6.2.1 The pET/pEVOL co-transformed system

To utilize the pET/pEVOL co-transformed system, pET/pEVOL series plasmids were extracted, and the S30 lysate was prepared after pEVOL-pBpF was transformed into the B95.ΔA strain.

2.6.2.1.1 B95.ΔA and B95.ΔA / pEVOL-pBpF growth curve

pEVOL-pBpF was transformed into competent cell B95.ΔA. One colony was picked up into 5 ml of 2xYT media supplemented with 50 mg/l chloramphenicol, incubated overnight at 37°C and 120 rpm. 1% of the overnight culture was inoculated into 200 ml of 2xYT media containing 50 mg/l chloramphenicol and grown at 37°C and 120 rpm. OD600 was measured every hour after inoculation. As a control, the growth curve of B95. ΔA was measured.

2.6.2.1.2 Plasmid Extraction

The plasmids used to express protein via an in vitro system were extracted from the *E. coli* Top 10 cells according to the manual from the GeneJET Plasmid Maxiprep Kit. Afterwards, the extracted plasmids were precipitated using ethanol precipitation. Briefly, the extracted plasmid was mixed with 1/10 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of 100% ethanol, frozen at -80°C for at least 20 min. Afterwards, the mixture was centrifuged at 4°C and full speed for 20 min, followed by washing the pellet with 70% and 100% ethanol. The

washed pellet was air-dried for a few minutes, resuspended in 200µl of sterile ddH₂O, and stored at -20°C. The concentration of the plasmid was measured by the nanodrop.

2.6.2.1.3 Lysate Preparation

The frozen B95.ΔA strain containing the pEVOL-pBpF was recovered by plating on an LB agar plate supplemented with 50 mg/L chloramphenicol. After overnight incubation, one colony was picked up into 5 ml of 2xYT media supplemented with 50 mg/L chloramphenicol, incubating for 8 hours at 37°C and 120 rpm with an INFORS HI incubator shaker. 1% of the 8h culture was inoculated into 50 ml of 2xYT media containing 50 mg/l chloramphenicol and grown overnight under the same conditions. 1% of the overnight culture was inoculated into 4 L of 2xYT media containing 50 mg/L chloramphenicol, growing under the same conditions to OD₆₀₀ = 0.5 before induction. The culture was induced by 0.02% arabinose to overexpress BpA-RS, growing under the same conditions. When OD₆₀₀ was around 3.0, cells were harvested with centrifugation at 4°C and 4500 rpm with JLM8.1 rotor for 40 mins. Cell pellets were resuspended with 1x S30 C buffer and centrifuged under the same conditions. After repeating this step, 1g of cell pellet was resuspended with 1 ml of 1x S30 C buffer to open the cells once with a cell disrupter, Microfluidics M-110P, under the pressure of 15,000 psi. Opened cells were centrifuged twice at 4°C and 30,000 xg, each for 30 mins. The supernatant was divided into three crude extracts. The first was performed with the incubation step by incubating with the preincubation buffer for 80 minutes at 37°C and 120 rpm with the MINI Benchmark INCU-shaker, the second was incubated without the preincubation buffer under the same conditions, and the last was performed without the incubation step. After this step, all extracts were dialyzed three times against S30 C buffer with 100 times the volume of samples at 4°C with a dialysis tube at MWCO=15KD, followed by centrifugation for 30 mins at 4°C and 30,000 xg. Afterwards, the supernatant was aliquoted 100 µl to pre-cool Eppendorf tubes,

flash-frozen with liquid nitrogen, and stored at -80°C. As the control, the extract from the B95.ΔA strain was carried out similarly.

2.6.2.1.4 In vitro Protein Expression Assay

Two reaction modes, namely batch mode and dialysis mode, were employed to optimize the concentrations of magnesium and potassium for the enhancement of productivity of the in vitro protein expression system in further research. This methodology is based on the cell-free lysate protein expression protocol provided by Cube Biotech, with all components prepared and stored in small aliquots under the appropriate conditions as stipulated in the protocol. In dialysis mode, the volume ratio between the reaction mixture and the feeding mixture is established at 1:16; consequently, to prepare 1 ml of the reaction mixture (see Table 2.6.4), 16 ml of the feeding mixture must be prepared (refer to Table 2.6.5). Accordingly, 5.9095 ml of the reaction master mixture and 1 ml of the feeding mixture buffer were prepared in advance by combining the components listed in Tables 2.6.2 and 2.6.3, respectively. Subsequently, the reaction master buffer was divided into two portions, with 296.7 µl allocated for the reaction mixture, while 4762.1 µl was combined with 5600 µl of the feeding mixture buffer to complete the feeding mixture. It is noteworthy that no feeding mixture is required for batch mode; therefore, Tables 2.6.3 and 2.6.4 serve as the protocol for the batch mode reaction. For the preparation of 1 ml of the reaction mixture, 296.7 µl of the reaction master buffer was mixed, with all components being scaled down accordingly. The batch mode represents the primary strategy employed in this work. In order to express the full-length PS1-N190X protein, pET15b-PS1-N190X, 1 mM BpA, and various lysates were incorporated into the batch mode reaction. As a control, the *E. coli* lysate provided by Cube Biotech was utilized.

Materials and Methods

Table 2.6.2 Feeding mixture buffer

Component	Stock concentration	Final concentration	Volume(μl)
Tris-acetate (pH 8.2)	1 M	10 mM	100
Magnesium acetate	1 M	14 mM	140
Potassium acetate	10 M	60 mM	60
DTT	500 mM	0.5 mM	10
ddH ₂ O			9690

Table 2.6.3 Reaction master mixture

Component	Stock concentration	Final concentration	Volume(μl)
6 Amino acid mix	16.7mM	1mM	1020
Amino acid mix	25mM	0.5mM	340
Li ⁺ , K ⁺ Acetyl phosphate (AcP)	1M	20mM	340
Phosphoenolpyruvate, K ⁺ (PEP)	1M	20mM	340
75 X NTP mix	90mM ATP; 60mM G/C/U TP	1.2mM ATP; 0.8mM G/C/U TP	226.7
DTT	500mM	2mM	68
Folinic acid (Ca ²⁺)	10mg/ml	0.1mg/ml	170
Complete protease inhibitor	50X	1X	340
HEPES/EDTA buffer	24X	1X	623.3
Magnesium acetate	1M	16mM	274
Potassium acetate	4M	270mM	382.5
PEG 8000	40% (W/V)	2% (W/V)	850
Sodium azide	10% (W/V)	0.05 (W/V)	85
		total	5059.5

Table 2.6.4 Reaction mixture

Component	Stock concentration	Final concentration	Volume(μ l)
Reaction Master Mix			297.6
Pyruvate kinase	10 mg/ml	0.04 mg/ml	4
tRNA (<i>E. coli</i>)	40 mg/ml	0.5 mg/ml	12.5
Ribolock RNase Inhibitor	40 U/ μ l	0.3 U/ μ l	7.5
Plasmid	200 ng/ μ l	20 ng/ μ l	100
<i>E. coli</i> lysate	1X	0.35X	350
T7 RNA polymerase	20 U/ml	0.5 U/ml	25
ddH ₂ O			203.4
		total	1000

Table 2.6.5 Feeding mixture

Component	Stock concentration	Final concentration	Volume(μ l)
Reaction Master Mix			4762.1
Feeding mixture Buffer	1X	0.35X	5600
Amino acid mix	25 mM	0.5 mM	320
ddH ₂ O			2117.9
		total	16000

2.6.2.2 The addition of BpA-tRNA and BpA-RS

To complete the second strategy, the S30 lysate was prepared in various ways, and the amount of different components was optimized.

2.6.2.2.1 Lysate preparation

The S30 lysate was prepared by varying the medium, cell harvesting time, the pressure for cell disruption, the process of endogenous mRNA degradation, and the overexpression of T7 RNAP.

2.6.2.2.1.1 B95.ΔA and BL21(DE3) growth curve

The frozen B95.ΔA was cultured on an LB agar plate and incubated overnight at 37°C. A single colony was selected and transferred into 5 ml of 2 x YTPG media, where it was subsequently grown overnight at 37°C and 120 rpm with the MINI Benchmark INCU-shaker. An aliquot of 1% from the overnight culture was inoculated into 200 ml of 2xYTPG media and incubated at 37°C and 120 rpm with the INFORS HI incubator shaker. The growth at OD600 was recorded hourly following inoculation. For control purposes, the growth curve of BL21(DE3) was assessed utilizing a methodology analogous to the one previously described.

2.6.2.2.1.2 Harvest B95.ΔA or B95.ΔA/pAR1219 cells

The pAR1219 plasmid was introduced into B95.ΔA competent cells. A single colony was isolated from the transformed plate and inoculated into 5 ml of LB media, which was supplemented with 100 ng/μl ampicillin, and incubated for 6 hours at 37°C and 120 rpm with the MINI Benchmark INCU-shaker. An aliquot of 1% from the 6-hour culture was transferred into 50 ml of 2xYTPG media containing 100 ng/μl ampicillin and incubated overnight under the same conditions. Subsequently, 40 ml of the overnight culture was transferred into 4 L of 2xYTPG media supplemented with 100 ng/μl ampicillin and grown at 37°C and 120 rpm with the INFORS HI incubator shaker until OD600 reached 0.4 before induction. To overexpress T7 RNA polymerase, 1 mM IPTG was introduced, and the culture was maintained at 30°C and 120 rpm with the INFORS HI incubator shaker. Once OD600 approximated 2.5, the culture was rapidly cooled using an appropriate quantity of ice, followed by the centrifugation of cells for 40 minutes at 4°C and 4500 rpm using the Beckman JLM 8.1 rotor. The harvested cells were then resuspended in pre-cooled 1xS30A buffer and subjected to centrifugation under the same conditions. This process was repeated twice, and during the final iteration, cell pellets were

flash-frozen in liquid nitrogen and stored at -80°C. As a control, the B95.ΔA strain was cultured and harvested under identical conditions without induction.

2.6.2.2.1.3 Crude extract preparation

The frozen cells were weighed and resuspended with 100% (V/W) of pre-cooled 1x S30B buffer. The thawed cells were homogenized with a Dounce tissue homogenizer and then disrupted once at 700 Bar using Microfluidics M-110P, followed by centrifugation for 30 minutes at 4°C and 30,000 xg using the Beckman JA25.50 rotor. 2/3 of the non-turbid supernatant was transferred to a new centrifugation tube to repeat the centrifugation. Afterwards, 2/3 of the non-turbid supernatant was collected for further experiments.

2.6.2.2.1.4 Endogenous mRNA degradation

The crude extract was adjusted to a final concentration of 400 mM NaCl, wrapped with aluminum foil, and incubated in the dark with a water bath for 45 minutes at 42°C. As controls, one trial was carried out at 37°C, while the other was carried out without this step.

2.6.2.2.1.5 Extract preparation

After degradation of endogenous mRNAs, all samples were dialyzed at 4°C for 4 hours with dialysis tubes at MWCO=15KD against 5 L of 1x S30C buffer. The buffer was exchanged with fresh 5 L of 1x S30C buffer to continue dialysis overnight. Afterwards, the dialyzed extract was centrifuged for 30 minutes at 4°C and 30,000 xg using the Beckman JA25.50 rotor. 2/3 of the non-turbid supernatant was aliquoted to pre-cooled Eppendorf tubes, flash-frozen by liquid nitrogen, and stored at -80°C. 3 ml of the flash-frozen extract from B95.ΔA was lyophilized and dissolved in 1 ml of 1x S30C buffer.

2.6.2.2.2 Optimization of magnesium and potassium ions

To optimize the concentration of Mg²⁺ and K⁺ for the new extract, EGFP was expressed by using the batch mode of the in-vitro protein expression system. After overnight expression, all samples were centrifuged at 5000 rpm and 4°C

for 5 minutes using the FA-45-48-11 rotor to remove the precipitate from the reaction. 50 ml of the supernatant was transferred to the black 96-well Immuno plate to measure the fluorescence with an excitation wavelength of 488nm and an emission wavelength of 507nm using a CLARIOstar plate reader.

2.6.2.2.3 The full-length PS1-N190X expression and solubilization

All components for the in vitro reaction system in batch mode were added to a 1.5 ml microtube. To express the full-length PS1-N190X, 1 mg/ml of 100nm liposomes, which were prepared as mentioned in 2.7.3, 10 μ M BpA-RS, 10 μ M tRNA, and 1 mM unnatural amino acid BpA were supplemented in the reaction mixture. After an overnight reaction at 30°C and 90 rpm with the MINI Benchmark INCU-shaker, the solution was ultracentrifuged at 150,000 xg for 1 hour to collect the membrane fractions. This pellet was solubilized with the corresponding detergents overnight at 4°C. After ultra-centrifugation, the fractions were further analyzed with a Western blot assay.

2.7 Photo-crosslinking assay

After the establishment of the in vitro protein expression system to incorporate BpA into the membrane protein PS1, the method for photocrosslinking was completed with the self-dimer soluble protein sjGST. Subsequently, different receivers for solubilizing expressed PS1 were prepared to help the crosslinking between PS1 and PEN2.

2.7.1 sjGST-F52X expression

The confirmed plasmid pET27b-GST-F52X was prepared as mentioned in 2.6.2.1.2, followed by mixing with other components for the in vitro reaction system in batch mode. To express the full-length sjGST-F52X, 10 μ M BpA-RS, 10 μ M tRNA, and 1 mM unnatural amino acid BpA were supplemented in the reaction mixture. After an overnight reaction at 30°C and 90 rpm with the MINI

Benchmark INCU-shaker, the solution was centrifuged for 5 mins at 4°C and 5000 rpm using the FA-45-48-11 rotor. The supernatant was snap-frozen by liquid nitrogen and stored at -80°C for further experiments.

As the control, pET27b-GST and pET27b-GST-F52X were transformed into BL21(DE3) separately to overexpress sjGST or sjGST-F52X. After harvesting and centrifugation, the cell pellets were lysed with the urea lysis buffer.

2.7.2 The method of photo-crosslinking

Crosslinker CL-1 was employed to perform the photo-crosslinking assay (Janhavi A. Kolhe 2023). 50µl of the resultant supernatant was pipetted to each wall of a transparent NUNC PolySorp 96-well plate, and put in the Styrofoam box filled with ice. To crosslink protein efficiently, the distance between the top of the 96-well plate and the 352 nm UV tubes was kept around 2 cm, and irradiation time was 5 minutes, followed by cooling down for 30 seconds. Subsequently, the irradiated sample was pipetted out of the 96-well plate. The rest samples were repeatedly irradiated and pipetted in the same way until the irradiation time of the last sample reached 90 minutes. All samples were purified using Ni-Indigo MagBeads from Cube Biotech. Hereafter, 6µl of purified samples were heated for 2 mins at 94°C for SDS-PAGE and Western Blot. The remaining purified samples were flash-frozen with liquid nitrogen and stored at -80°C for further experiments. As the control, the sample without irradiation was processed in the same way.

2.7.3 Receivers preparation

Liposomes can be fabricated in several approaches, such as thin film hydration, reverse-phase evaporation, solvent injection, detergent removal, dehydration-rehydration, heating method, pH jumping method, microfluidic channel method, and supercritical fluidic method (Hamdi Nsairat 2022). In this work, the liposome preparation was processed in the following steps.

2.7.3.1 Thin-film hydration

A. D. Bangham and his colleagues first described the thin-film hydration method in 1965 (A. D. Bangham 1965), which primarily produces large, polydisperse multilamellar vesicles (MLVs) ranging from 0.5 to 10 μm . In this work, the procedure is as follows. 40 mg of lipid mixture, 25% (w/w) brain extract total lipids, 65% (w/w) egg phosphatidylcholine (eggPC), and 10% (w/w) Cholesterol (Tao 2022), was dissolved into 1 ml of chloroform using a screw cap storage vial from Avanti. The chloroform solution was deposited in a round-bottom flask to evaporate chloroform at room temperature for 30 minutes with rotary evaporation to create a thin film layer, which was resuspended with 1 ml of hydration buffer. The resuspension solution was vacuumed overnight to remove a trace amount of chloroform using a vacuum apparatus in the dark. Thereafter, the lipids resuspension was aliquoted to 200 μl per Eppendorf microtube with a capacity of 1.5 ml, covered by Aluminum foil, and stored at -20°C .

2.7.3.2 Freeze-thaw cycles and extrusion

Olson and his colleagues proposed a new procedure of liposome preparation by sequential extrusion in 1979, which is facile, reproducible, and highly efficient in generating liposomes with a defined size distribution (F. Olson 1979). Before extrusion, the aliquoted lipid suspension was hydrated for 1 h at 65°C using a heating block, which is above the phase transition temperature of the lipid, followed by 5 cycles of freeze and thaw step to increase the proportion of uni-lamellar vesicles (LUVs) from multilamellar vesicles (MLVs) (Barbara Mui 2003). Briefly, the lipid suspension was frozen in liquid nitrogen for 10 minutes and unfrozen at 65°C for 10 minutes. After 5 times of repetition, the resultant was ready to extrude. The mini-extruder was assembled according to the manual and preheated at 65°C using a heating block. Subsequently, the resultant was loaded into one of the gas-tight glass syringes, pre-wetted with 250 μl of hydration buffer, and gently placed into one end of the Mini-extruder

(Avanti). After the pre-wetted and empty gas-tight syringe was placed into the other end of the extruder, the plunger of the filled syringe was pushed gently to transfer all suspension to the empty one by extruding through a 100 nm filter, repeating more than 10 times until the solution became translucent to generate 100 nm liposomes. Then the resultant 100 nm liposomes were stored in the dark at 4°C for one week for further experiments.

2.7.3.3 Liposome Titration with CHAPSO

The 100 nm liposome was titrated with the hydration buffer containing 10% CHAPSO in different ratios. The resultant was shaken at 21°C and 1000 rpm for 1 hour using a thermomixer (Josep Cladera 1997), then measured optical density at the wavelength ranging from 250 nm to 600 nm with an IMPLAN Nanophotometer to identify an appropriate proportion between the lipid mixture and CHAPSO.

2.7.3.4 CHAPSO-destabilized liposome preparation

20 mg/ml of lipid mixture was mixed with the defined concentration of CHAPSO buffer, incubated at 40°C and 1000 rpm for 30 mins using a thermomixer, and then the mixture underwent 5 cycles of freeze and thaw steps, frozen by liquid nitrogen and unfrozen at 40°C using a heating block, followed by being extruded into 100nm liposomes. Afterwards, the resultant 200 µl liposomes were dialyzed twice at 4°C against 500 mL of the hydration buffer, one for 4 hours and the other overnight. The prepared CHAPSO-destabilized liposome was stored in the dark at 4°C for further experiments.

2.7.3.5 Dynamic light scattering

Dynamic light scattering (DLS) was employed to analyze the size of liposomes or CHAPSO-destabilized liposomes after being diluted 10-fold with hydration buffer and filtered with Ultrafree-MC microcentrifuge filters.

2.7.5 Photo-crosslinking of PS1-N190X and PEN2

To photocrosslink PS1-N190X and PEN2, different receivers were added to the established CFPS. Considering the possible influence of the number of receivers and the extracted plasmids on CFPS, the addition of plasmids and receivers was variable. The resulting samples were photocrosslinked, followed by SEC purification to identify the incorporation efficiency of the PS1-N190X/PEN2 complex into liposomes, and by Ni-indigo purification to obtain pure samples. The purified samples were identified by mass spectrometry.

2.7.5.1 PS1-N190X and PEN2 expression

pET15b-PS1-N190X and pET27b-PEN2, in the final concentration of 500 ng separately, were added to the in vitro reaction system. To identify the appropriate hydrophobic circumstance for PS1-N190X and PEN2 subcomplex expression, 1mg/ml of liposomes was substituted by 1mg/ml of CHAPSO-destabilized liposome, or 0.1mg/ml of home-made MSP2N2-DMPC nanodiscs. As controls, PS1-N190X or PEN2 was separately expressed with liposomes or CHAPSO-destabilized liposomes, and the reaction system containing 1mg/ml of liposomes and without plasmids was treated similarly. After overnight incubation, the sample was snap-frozen and stored at -80°C for further experiments.

For the experiments involving Sulfo-DIBMA-DMPC nanodiscs, PS1/PS1-N190X and PEN2, in the molecular ratio of 1:1, were added to the in vitro reaction system as mentioned above, where 1 mg/ml of liposomes was replaced with 67.2 mM sulfo-DIBMA-DMPC nanodiscs. The expression of PS1-N190X and PEN2 was performed at 30°C overnight. Subsequently, all samples were ultracentrifuged for 1h at 100,000 xg and 20°C. The resultant supernatant and pellet were snap-frozen and stored at -80°C for further experiments.

2.7.5.2 Photo-crosslinking assay

All frozen samples were thawed in an ice bath and irradiated for 90 minutes using the method mentioned above. Subsequently, all 50 μ l of the samples were precipitated using the acetone precipitation method. Briefly, 10 times the volume of cold acetone was added to the samples, followed by storage overnight at -20°C. Subsequently, after centrifugation for 30 minutes at 4°C and 10,000 xg, the supernatant was discarded, and the pellet was air-dried for around 1 minute. Finally, the resultant sample was resuspended with 50 μ l of 1xS30C buffer. They were treated with or without the loading buffer containing 1% Triton X-100 to identify the crosslinked bands by the western blot.

2.7.5.3 Purification by SEC or Ni-Indigo chromatography

To analyze the insertion of PS1-N190X and PEN2 subcomplexes into liposomes, the irradiated samples, PS1-N190X, PEN2, and PS1-N190X-PEN2 subcomplexes, were centrifuged for 5 minutes at 4°C and 4500 rpm using the FA-45-48-11 rotor. Subsequently, the supernatant was purified by SEC with a Superose 6 column. To monitor liposomes in the Superose 6 column, 10mg/ml of liposomes were purified using the same method. To analyze the interaction of PS1/PS1-N190X and PEN2 subcomplexes in the presence of Sulfo-DIBMA-DMPC nanodiscs, the irradiated samples were purified by Ni-Indigo Magbeads. All fractions were snap-frozen and stored at -80°C for further experiments. For Western blot, samples were treated with the loading buffer containing 1% Triton X-100 after acetone precipitation.

2.7.5.4 Purification after photocrosslinking the solubilized PS1-N190X/PEN2

After 4 hours of co-expression or single expression of PS1-N190X and PEN2 with liposomes, all samples were ultracentrifuged for 1 hour at 100,000 xg to collect the membrane fractions and aggregated proteins, which were solubilized overnight at 4°C with resuspension buffer (Table 2.3.14). After

irradiation for 90 mins, the solubilization solution was purified by Ni-Indigo Magbeads, followed by snap-frozen and stored at -80°C for further experiments. For Western blot, samples were treated with the loading buffer containing 1% Triton X-100 after acetone purification.

2.7.5.5 Mass spectrometry assay

The purified sample based on Sulfo-DIBMA-DMC nanodiscs was precipitated by acetone precipitation, and the concentration was measured using the BSA method. 10.5µl of sample containing 0.025-10µg of protein was added into the 1.5ml microcentrifuge tube containing 15µl of digestion buffer and 1.5µl of reducing buffer, incubated for 5 minutes at 95°C, and cooled down to room temperature, followed by supplement with 3µl of Alkylation buffer and incubation for 30 minutes in the dark and room temperature. The resultant was incubated overnight with 2µl of activated trypsin buffer at 30°C and 90 rpm using the mini Benchmark INCU-shaker, and then supplemented with 0.5µl of 100% formic acid to stop the reaction.

The purified sample based on liposomes was cut from the SDS-PAGE and added to a 1.5ml microcentrifuge tube containing 15µl of digestion buffer, and then digested for 3 hours at 37°C. Afterward, the supernatant was transferred carefully to a new 1.5ml microcentrifuge. After adding 1.5µl of reducing buffer, the sample was prepared according to the procedure mentioned above. Subsequently, all prepared samples were sent to my colleagues for completion of mass spectrometry and further data analysis.

3.1 Auxiliary Proteins and tRNA

3.1.1 TEV Protease

TEV protease is a nuclear inclusion protease derived from the tobacco etch virus. It is a stringently specific protease that can cut between Q and S or G in the ENYLFQS/G peptide sequence. It provides a convenient biological tool for releasing peptides or proteins from genetically engineered fusion proteins (T. Dawn Parks 1994). To remove the 6xHis-tag from BpA-RS or AzF-RS, TEV protease was purified using Ni-NTA affinity chromatography. As shown in Figure 3.1 A and B, after centrifugation at 20,000 xg, most of the TEV (26.8 KDa) was in the supernatant, while the misfolded TEV, which remained in the inclusion body, was pelleted down. The flow-through lane showed that all TEV proteins were bound to the Ni-NTA resin. Except for the wash buffer lane showing a small amount of TEV eluted from the column, most of the TEV was eluted by the elution buffer. The elution buffer lane also indicated that the purity of TEV was sufficient to proceed with further experiments. This experiment was repeated twice, and the concentration of TEV was 1.68 mg/ml and 6.68 mg/ml, respectively.

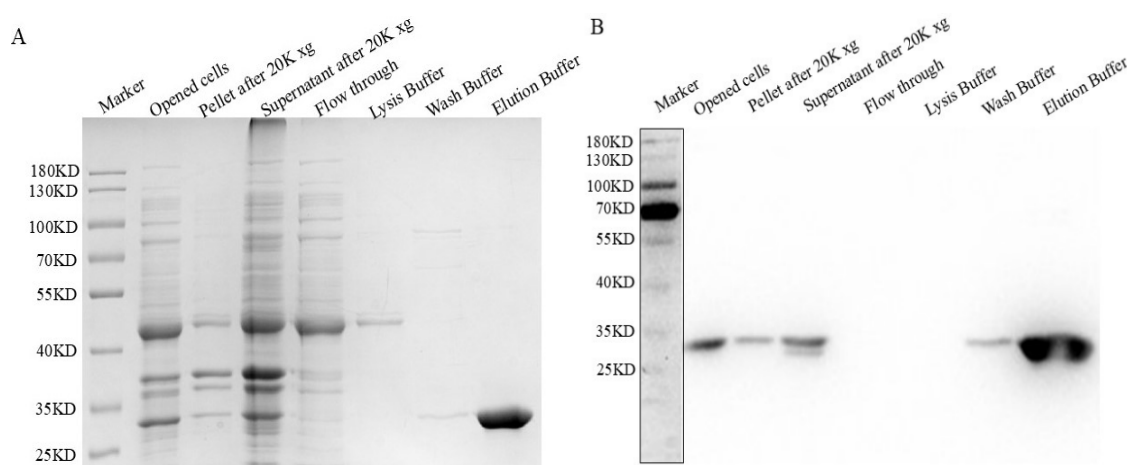


Figure 3.1 TEV purification by Ni-NTA

TEV was purified using a Ni-NTA affinity chromatography assay. The SDS-PAGE and Western Blot analyze different fractions. Panel A shows the stained SDS-PAGE by Coomassie brilliant blue solution, while panel B shows the Western blot result.

3.1.2 BpA-RS and AzF-RS

To enhance the incorporation of UAAs into target proteins, the related RNA synthetases were purified via Ni-NTA affinity chromatography, with the 6xHis tag cleaved by TEV protease. For purification, pET27b-BpA-RS and pET27b-AzF-RS were constructed employing standard enzyme digestion and ligation techniques. As depicted in Figures 3.2 A and 3.10 A, the two bands observed in the lanes containing digested pET27b-BpA-RS or digested pET27b-AzF-RS correspond in size to the bands in the other two lanes, indicating successful reconstitution of BpA-RS or AzF-RS into the pET27b vector, where a 6xHis tag is fused to the N-terminus of the target proteins. Subsequently, using BpA-RS as an illustrative example, the process of protein overexpression was examined. To monitor the bacterial growth pattern in the presence of 1 mM IPTG, OD600 was sampled at half-hour intervals post-inoculation. As illustrated in Figure 3.2 B, the orange line indicates that cellular growth was slow following induction with 1 mM IPTG, in contrast to the blue line, which depicts the development of the uninduced sample. To optimize the IPTG induction duration, induced samples were harvested, centrifuged, and lysed using an 8 M urea lysis buffer, followed by Western blot analysis. Figure 3.2 C displays the BpA-RS yield at various induction times, demonstrating that the yield plateaued after 4.5 hours of induction. Accordingly, cells were harvested after 5 hours of growth. Subsequently, the protein purification method was established for subsequent experiments.

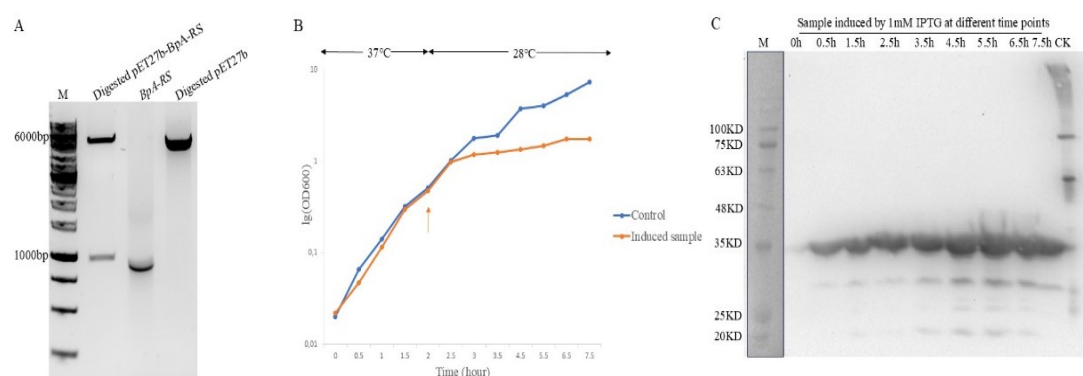


Figure 3.2 BpA-RS construct and overexpression

Results

pET27b-BpA-RS was constructed through classical restriction endonuclease digestion and ligation assays. Subsequently, it was transformed into the BL21 strain to facilitate overexpression of the BpA-RS protein following IPTG induction. Panel A shows the agarose gel electrophoresis of the constructed pET27b-BpA-RS. Panel B shows the growth curve of BL21 cells containing the pET27b-BpA-RS plasmid under varying incubation temperatures. The blue line represents the control group in which 1 mM IPTG was not added to induce BpA-RS overexpression. The orange line depicts the experimental group in which BpA-RS overexpression was induced with 1 mM IPTG. The orange arrow indicates the time point at which IPTG was added. Panel C presents the Western blot analysis of His-tagged BpA-RS samples collected at different time points following induction with 1 mM IPTG.

The buffer containing different concentrations of imidazole was used to purify BpA-RS to determine the optimal imidazole concentration for washing contaminant proteins or eluting target proteins. As shown in Figure 3.3 A and B, the inclusion body, including misfolded BpA-RS, was pelleted down after centrifugation at 10,000 xg, which demonstrates that 28°C is a high temperature for the induction of BpA-RS. A trace of BpA-RS unbound to resin was present in the flow-through fraction; therefore, the amount of Ni-NTA resin should be increased in further purification assays. When the concentration of imidazole was below 30 mM, the contaminant proteins were washed out of the column, accompanied by a trace of BpA-RS, which indicates that 30 mM imidazole is the optimal concentration to separate contaminant proteins from BpA-RS. As shown in Figure 3.3 C and D, BpA-RS eluted from the column when the imidazole concentration was up to 250 mM in the buffer, so 250 mM imidazole is the optimal concentration to elute BpA-RS from the Ni-NTA column. These results were confirmed again in Figure 3.4. Meanwhile, Figure 3.6B shows that 30 mM imidazole washed AzF-RS out of the column in a large amount; therefore, 20 mM imidazole is the optimal concentration to wash unexpected proteins from AzF-RS. The concentrations of BpA-RS and AzF-RS are 0.508 mg/ml and 6.3 mg/ml, respectively, after calculation with the Beer-Lambert Law.

Results

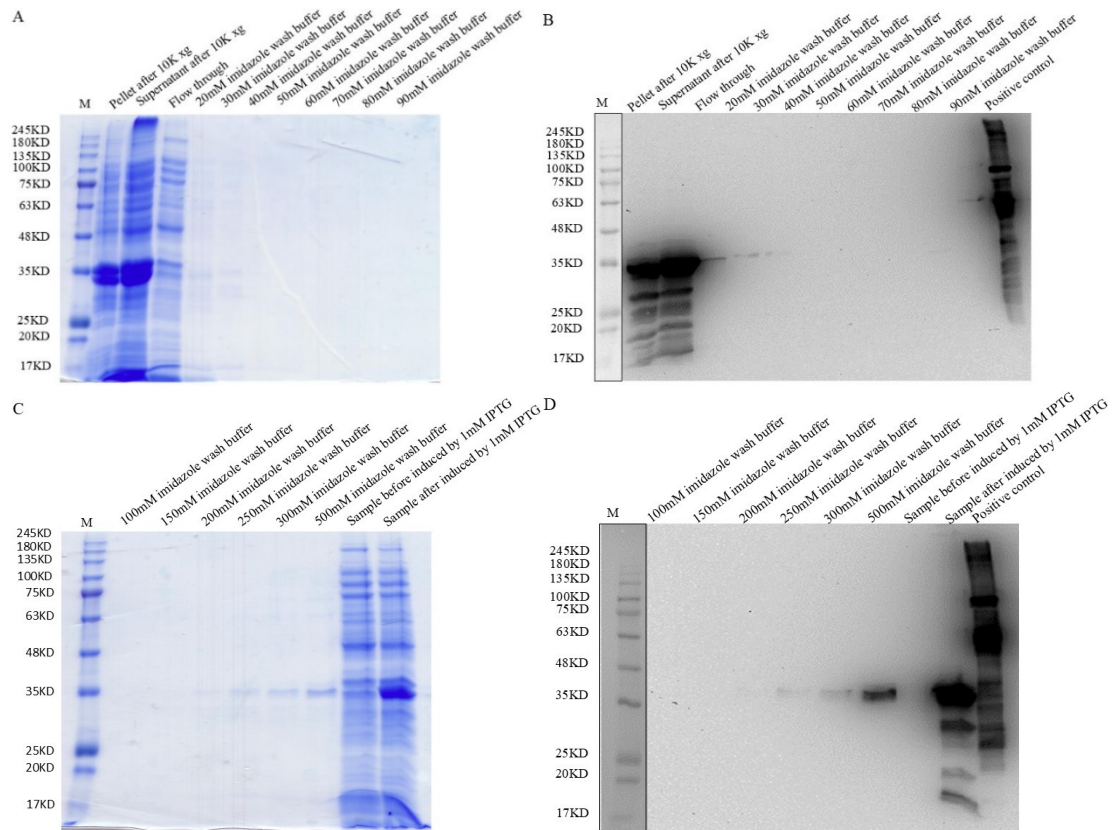


Figure 3.3 The optimization of BpA-RS purification

To establish the method of RNA synthetase purification by Ni-NTA chromatography, the lysate containing BpA-RS was bound to the resin and washed with wash buffers containing imidazole in variable concentrations. Panels A and C show the stained SDS-PAGE by Coomassie brilliant blue solution. Panels B and D show the corresponding Western blot on the His-tag in the N-terminal of BpA-RS. The positive control is an indicator for the Western blot on the His-tag.

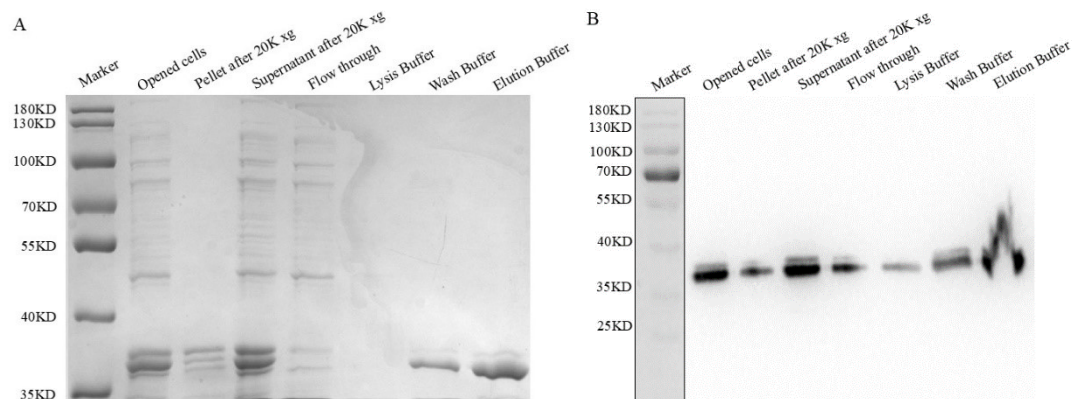


Figure 3.4 BpA-RS purification by Ni-NTA

After establishing the purification method for RNA synthetase, BpA-RS was purified on a large scale with determined buffers. Panel A shows the stained SDS-PAGE gel with a Coomassie Brilliant Blue solution. Panel B shows the corresponding results of the Western blot on the His-tag in the N-terminal of BpA-RS.

Results

The next step is removing the 6xHis tag from purified BpA-RS or AzF-RS. Figure 3.5 A shows that 2 μ l of TEV could completely cut the 6xHis tag from 10 μ l of BpA-RS in the 20 μ l reaction system, and the molar ratio between BpA-RS and TEV is 1.1, while Figure 3.6 C shows that 0.25 μ l of TEV could completely cut the 6xHis tag from 2 μ l of AzF-RS in the 10 μ l reaction system, whose molar ratio is 5.7. Since the molecular ratios between TEV and the target protein BpA-RS or AzF-RS varied, the efficient amount of TEV to release proteins should be tested in each assay. After scaling up this experiment, the resultant protein was purified with Ni-NTA chromatography. As shown in Figure 3.5 B and C or Figure 3.6 D and E, the target proteins without His-tag and a small amount of unbound TEV remained in the flow-through fraction. Uncleaved target proteins and TEV were eluted by the elution buffer, remaining in the elution buffer fraction, indicating that the efficiency of TEV cleavage decreased after scaling up the reaction. This may be caused by the changes in the concentration of monovalent ions, reaction time, and local reaction temperature. Due to the remaining TEV in the cleaved BpA-RS or AzF-RS, the effects of TEV on the in vitro protein expression system should be considered.

Results

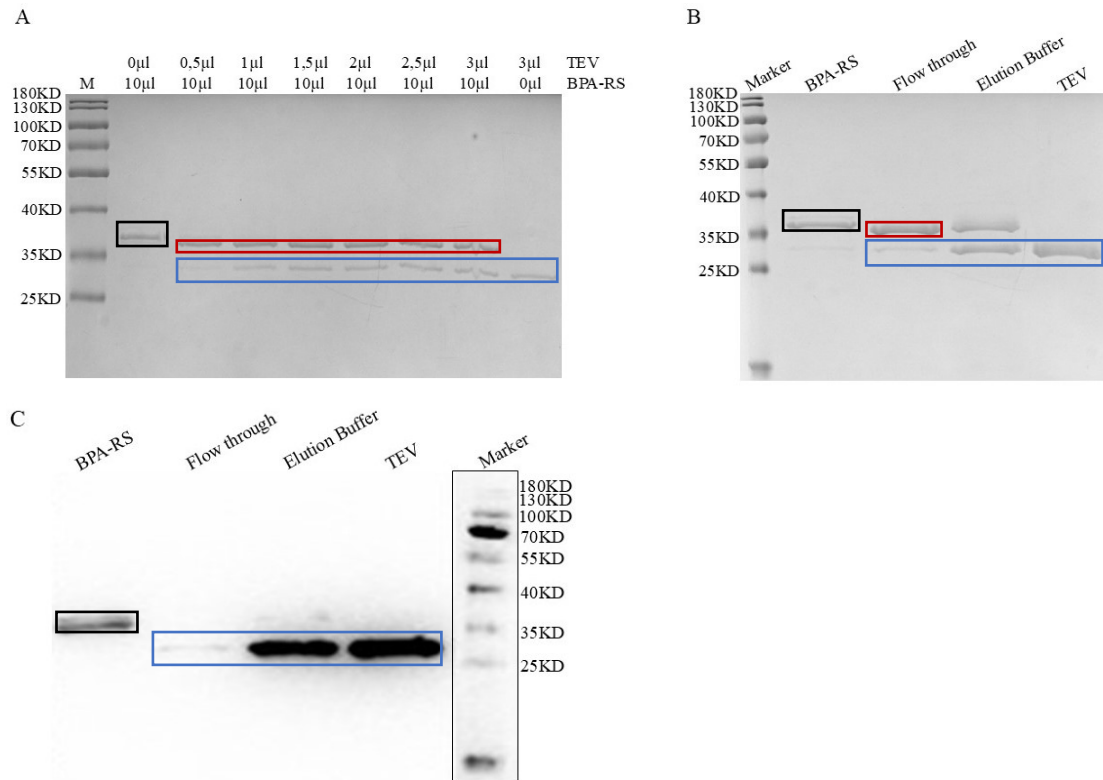


Figure 3.5 TEV cleavage of the N-terminal His-tag from BpA-RS

The optimal molecular ratio between TEV and BpA-RS for His-tag cleavage was initially determined using a testing mode, demonstrated in Panel A, where the volume of BpA-RS was maintained at 10 μ l and the volume of TEV was increased in 0.5 μ l increments. Subsequently, the His-tag was cleaved from BpA-RS by TEV at the established molecular ratio, followed by Ni-NTA purification. The purified fractions are shown in Panels B and C. Panels A and B are the stained SDS-PAGE gel subjected to Coomassie Brilliant Blue staining, while Panel C presents the Western blot for the His-tag. The BpA-RS lacking the His-tag is enclosed within a red frame, TEV within a blue frame, and BpA-RS within a black frame.

Results

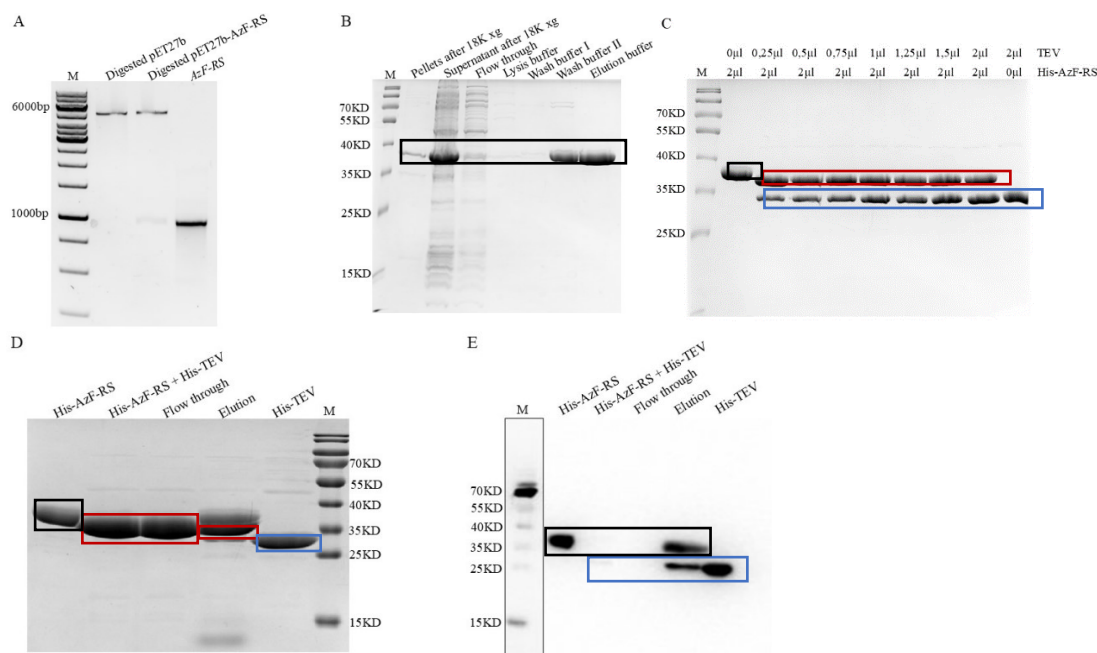


Figure 3.6 TEV cleavage of the N-terminal His-tag from AzF-RS

Plasmid pET27b-AzF-RS was constructed by classical restriction endonuclease digestion and ligation assay (Panel A), which was transformed into the BL21 strain to express AzF-RS protein after induction with 1 mM IPTG. The overexpressed AzF-RS was purified by Ni-NTA chromatography (Panel B). Subsequently, the purified AzF-RS was incubated with TEV in different molecular ratios to identify the optimal one (Panel C). Finally, the Azf-RS without His-tag was purified by Ni-NTA chromatography after the scale-up cleavage assay (Panels D and E). Panel A is the agarose gel electrophoresis of the constructed pET27b-AzF-RS. Panels B, C, and D are SDS-PAGE gel subjected to Coomassie Brilliant Blue staining. Panel E is the Western blot for the His-tag. The AzF-RS lacking the His-tag is enclosed within a red frame, TEV within a blue frame, and Azf-RS within a black frame.

3.1.3 EGFP

Since EGFP is a convenient tool for quantifying the expression efficiency of an in vitro protein expression system, EGFP was purified using Ni-NTA chromatography and SEC. Figure 3.7 A and B show that the target band appears in the wash I and II buffers, indicating that the target protein EGFP did not fully bind to the Ni-NTA resin, likely due to an inadequate amount of Ni-NTA resin. Following SEC purification, the purity of EGFP is approximately 95%, with a concentration of 0.2 mg/mL, suggesting that the purity of the purified EGFP may be a contributing factor to the potential overestimation of the yield in an in vitro protein expression system during subsequent experiments. Figure 3.7 C

Results

presents the standard curve correlating the concentration of EGFP along the X-axis with fluorescence on the Y-axis. The derived equation will be employed to calculate the yield of the in vitro protein expression system in the following experiments.

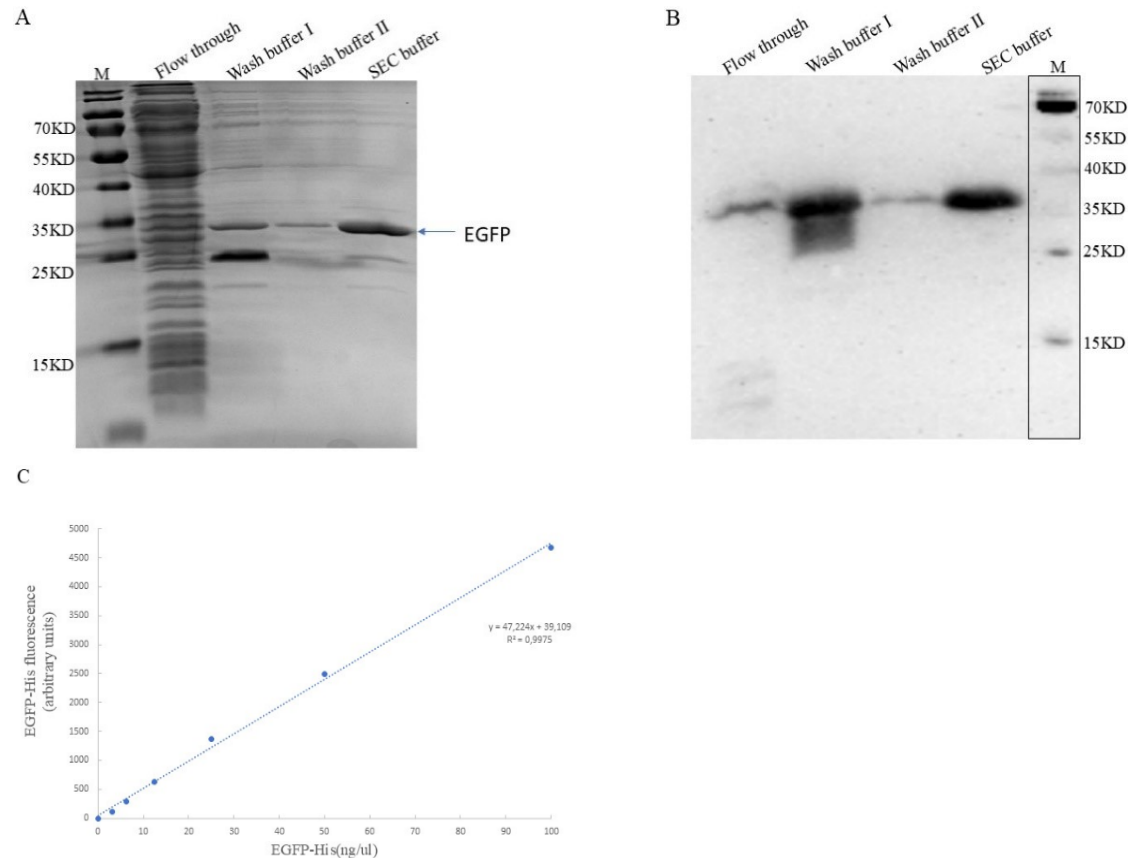


Figure 3.7 Purification of EGFP

EGFP was purified using Ni-NTA and SEC chromatography. Panel A shows the SDS-PAGE using Coomassie Brilliant Blue solution. Panel B shows the result of the corresponding Western blot for the His-tag. Panel C shows the EGFP fluorescence curve, correlating the concentration of EGFP along the X-axis with fluorescence on the Y-axis.

3.1.4 T7 RNA polymerase

To purify the active T7 RNA polymerase, an activity assay was established using a homemade T7 RNA polymerase with two different reaction buffers. As shown in Figure 3.8 A, the EGFP gene transcribes 0.717 kb mRNA in buffer B rather than buffer A, indicating that BSA is the activating factor of T7 RNA polymerase. Therefore, buffer B is a suitable transcription buffer for analyzing

Results

T7 RNA polymerase activity. Figure 3.8 C shows no transcribed band when T7 RNA polymerase was absent from the transcription reaction system. Additionally, the transcription activity of T7 RNA polymerase decreases with lower concentrations and is inhibited at high concentrations. The translation results in Figure 3.8 D show the same trend when the homemade T7 RNA polymerase was added to the in vitro protein expression system, which will be described in detail in the following chapters. These findings suggest that 125 $\mu\text{g}/\text{ml}$ of homemade T7 RNA polymerase can serve as a positive control for testing the activity of purified T7 RNA polymerase in further experiments.

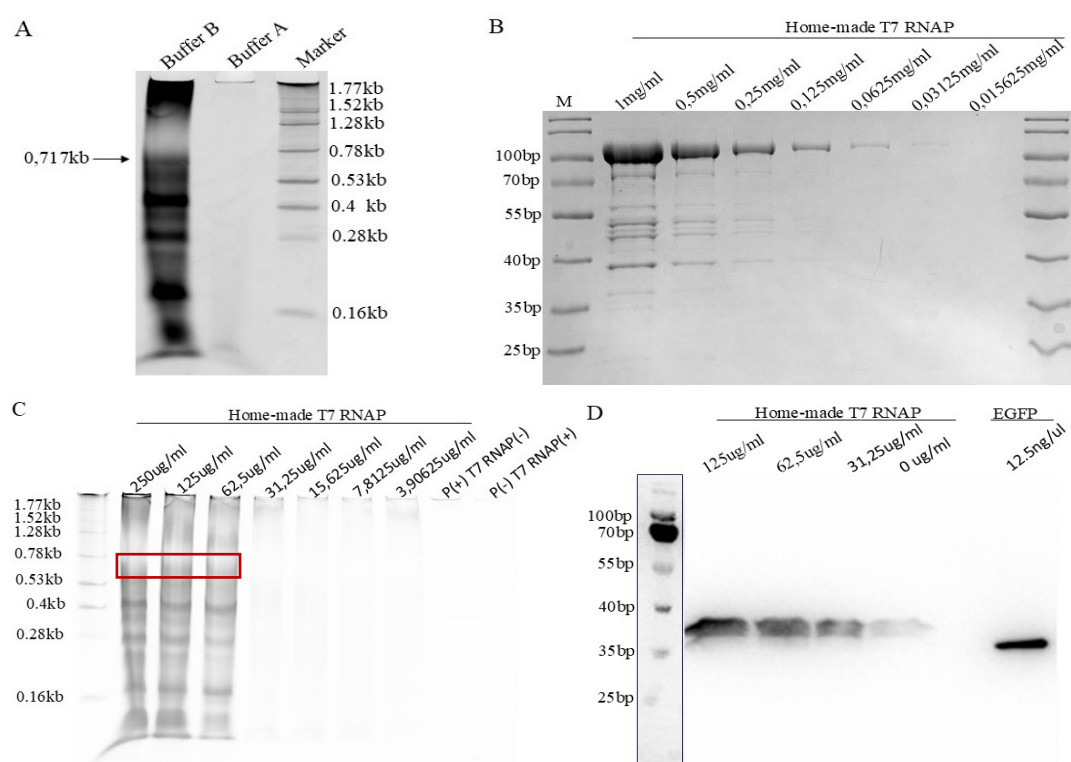


Figure 3.8 The activity test of home-made T7 RNA polymerase

Homemade T7 RNA polymerase was used to establish an appropriate method for analyzing the activity of purified T7 RNA polymerase. Panels A and C show the denatured PAGE. Panel B shows stained SDS-PAGE by Coomassie Brilliant Blue solution. Panel D is the Western blot result of EGFP expressed by the in vitro protein expression system. The target transcribed bands are enclosed in the red frame.

Results

T7 RNA polymerase was purified using Affi-Gel Blue and DEAE Sepharose chromatography. As shown in Figure 3.9 A, a few misfolded T7 RNA polymerases remain in the inclusion body pellet after centrifugation at 38K xg, and 0.2% streptomycin sulfate precipitates some T7 RNA polymerase after centrifugation at 27K xg by removing nucleic acids and ribonucleoproteins. Subsequently, as shown in Figure 3.9 B, the supernatant does not fully bind to the Affi-Gel Blue column, as indicated by the flow-through lane. The target protein elutes in peak 42, although some is washed out during washes 10 and 11; the resin-bound target protein is eluted with regeneration buffer in lane 4. After dialysis, the sample was further purified using DEAE Sepharose. As shown in lanes labeled with peak 25/26/27, the yield of T7 RNA polymerase is too low to be visible in the SDS-PAGE, and activity tests indicate that none of the fractions show transcription activity, except the supernatant after the centrifugation at 38K xg, which appears smeared, possibly due to polynucleotide phosphorylase from *E. coli*, as seen in Figure 3.9 C. Interestingly, commercial T7 RNA polymerase failed to transcribe the EGFP gene in buffer B, possibly because it is damaged or requires spermidine to activate. These results prompted the development of a second purification method for T7 RNA polymerase.

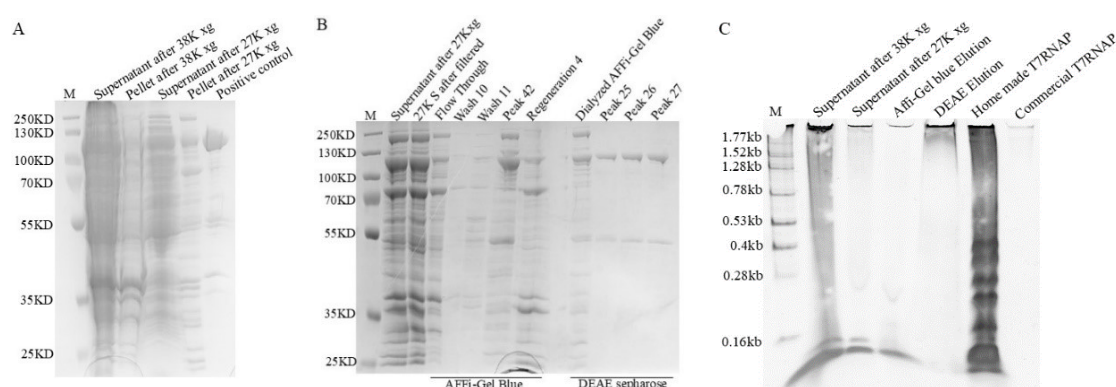


Figure 3.9 T7 RNAP purification and activity test

T7 RNAP was purified by AFFi-GEL Blue and DEAE Sepharose chromatography. Panels A and B show the different fractions produced from purification steps in SDS-PAGE stained with

Results

Coomassie brilliant blue solution. Panel C shows the denatured PAGE of the activity test assay from different fractions.

The resulting supernatant was gradient precipitated with ammonium sulfate following streptomycin sulfate precipitation. Each fraction was then purified using Affi-Gel Blue chromatography. As shown in Figure 3.10 A, the target protein could not be fully separated from other proteins. Surprisingly, after 20% ammonium sulfate precipitation, the impure fraction could transcribe the EGFP gene, as seen in the elution buffer-1 lane of Figure 3.10 B. The activity tests for T7 RNA polymerase optimized its activity at various concentrations. The trend matches that in Figure 3.8 C, because the lane labeled x0 in Figure 3.10 C shows that purified T7 RNA polymerase cannot transcribe the EGFP gene at high or very low concentrations.

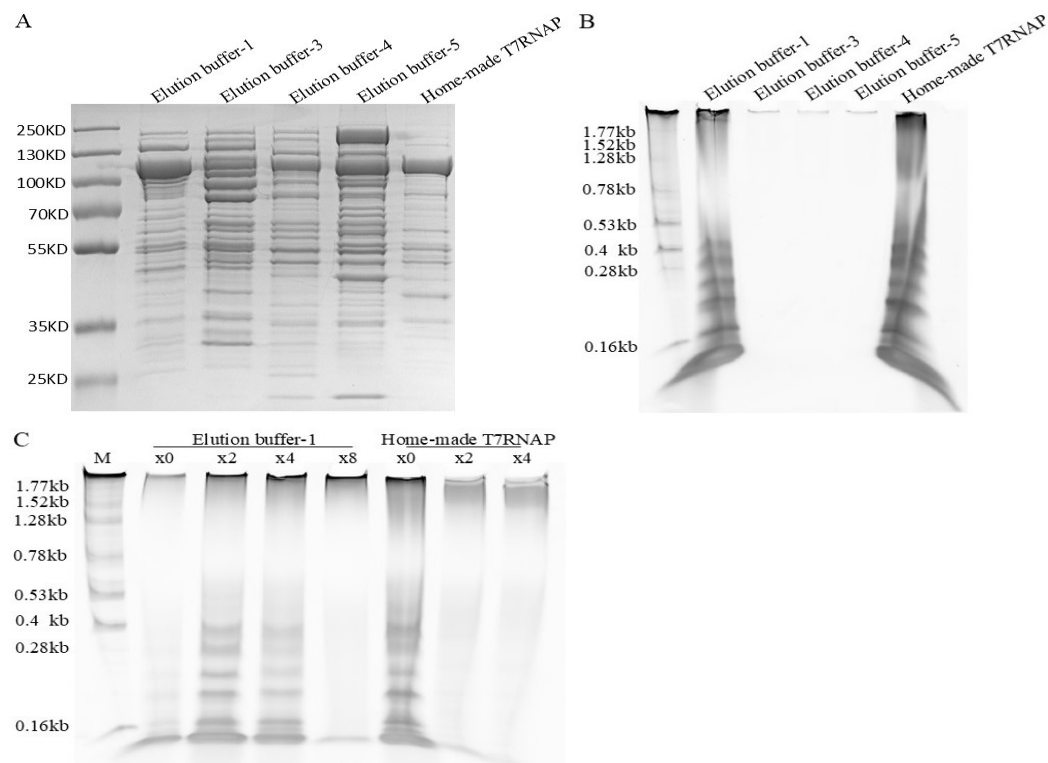


Figure 3.10 T7 RNAP purification and activity test

T7RNAP was purified by AFFi-Gel blue chromatography and ammonium sulfate precipitation. Panel A shows the various fractions obtained through purification procedures, as visualized by SDS-PAGE stained with Coomassie Brilliant Blue. Panel B shows the denatured PAGE results of the activity test assay across different fractions. Panel C presents the optimization process for the concentration of the purified T7 RNA polymerase.

3.1.5 Total-tRNA extraction

As the hypothetical strategy to improve the incorporation of unnatural amino acids into proteins, the total-tRNA containing the one carrying and transporting unnatural amino acids was extracted from BL21 (DE3) star. As shown in Figure 3.11 A and B, rRNA was separated from DNA and tRNA after being precipitated by rRNA precipitation buffer, because a high salt solution provides a high concentration of positively charged ions, which neutralize the negatively charged phosphate group of RNA precipitating rRNA containing >100 nucleotides, instead of tRNA, DNA and other small RNA (Guy Cathala 1983; Sarah E. Walker 2013). DNA was pelleted down from tRNA after being precipitated with DNA precipitation buffer, accompanied by a trace of tRNA. In addition, Figure 3.11 A also shows no changes after tRNA deacylation, and the purity of tRNA is lower than that shown in Figure 3.11 B. This may be because the acidic phenol solution was not fresh at the first extraction. The concentrations of total tRNA were 4850 ng/ μ L (A) and 12920 ng/ μ L (B), as measured by the nanophotometer from IMPLAN.

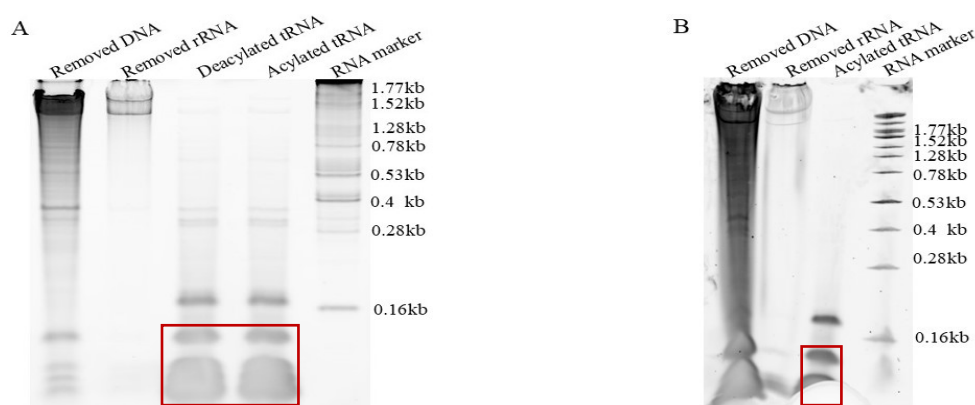


Figure 3.11 Denatured SDS-PAGE for extracted total tRNA

pEVOL-BpA-tRNA was transformed into BL21(DE3), and the total tRNA, containing the one carrying unnatural amino acids, was extracted, as visualized in the Denatured SDS-PAGE. Panel A is the result of the first extraction with an acidic phenol solution stored for over a month, and Panel B is the result of the extraction with a fresh acidic phenol solution. The total tRNA is enclosed within a red frame.

3.2 Protein Constructs

3.2.1 Analysis of γ -secretase

To identify an appropriate amino acid within PS1 for photocrosslinking with PEN2 at a distance of 6.0 Å, the structure of γ -secretase (PDB entry 5FN5) was utilized to conduct bioinformatic analysis. Considering the influence of the C-H bonds from lipids in the phospholipid bilayer on photocrosslinking with PS1, orientations of γ -secretase in membranes were examined using the PPM 2.0 Web Server from the OPM database, followed by PyMOL analysis to determine the distance between PS1 and PEN2. The amino acid asparagine at position 190 in PS1 was replaced with BpA via WinCoot. Subsequently, the distance between the BpA residue of PS1 and the interactive amino acid of PEN2 was assessed using PyMOL. As shown in Figure 3.12 A, the asparagine at position 190 of presenilin 1 is embedded on the extracellular side of the phospholipid bilayer, with its side chain being 3.5 Å away from the C-H bonds of leucine at position 98 of PEN2 (PEN2-L98). This 3.5 Å distance meets the criteria necessary for photocrosslinking (6 Å). Accordingly, this asparagine is identified as a potential candidate for mutation. When the 190th asparagine is substituted with the unnatural amino acid p-benzoyl-L-phenylalanine (PS1-N190X), the distance increases to 6.0 Å (Figure 3.12 B) or 4.6 Å (Figure 3.12 C), due to the steric hindrance introduced by the benzophenone moiety in BpA; however, this does not impede the formation of the covalent adduct. Furthermore, since this covalent adduct between PS1-N190X and PEN2-L98 is located on the extracellular side, it is unlikely to be influenced by the C-H bonds from the phospholipid bilayer. Therefore, this substitution mutation (PS1-N190X) will be prioritized in future studies involving γ -secretase photo-crosslinking.

Results

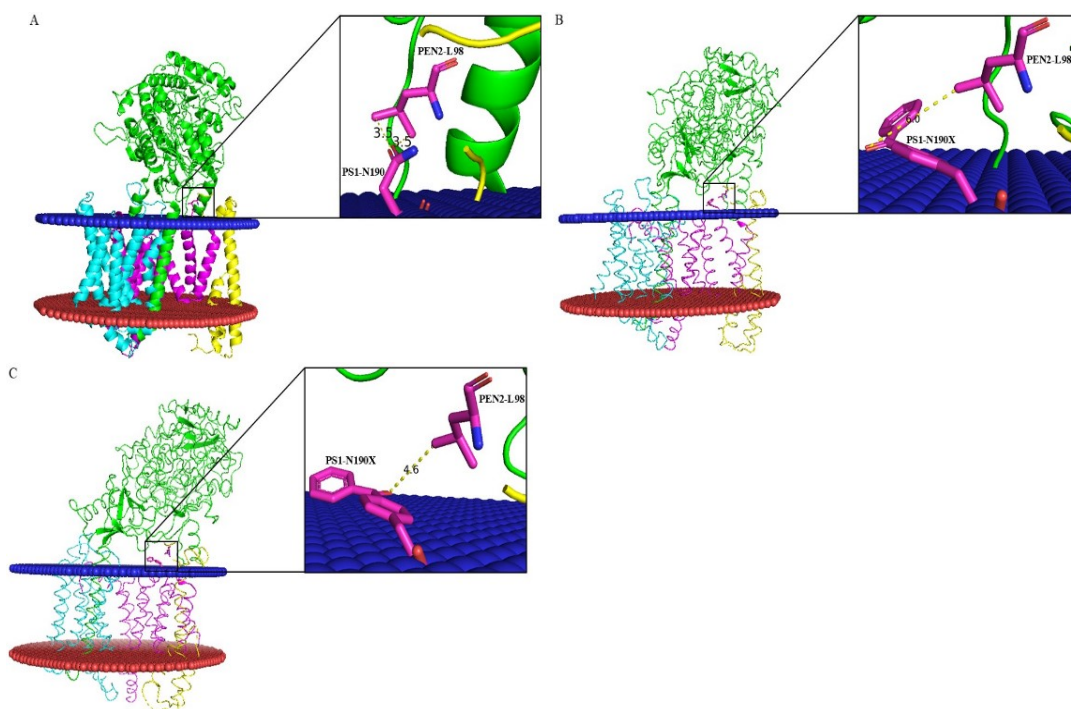


Figure 3.12 The distance between PS1-N190/PS1-N190X and PEN2-L98 in γ -secretase complex

All structures are based on the γ -secretase with the entry ID 5FN5 in the Protein Data Bank (PDB), whose orientation within membranes was examined using the PPM 2.0 Web Server from the Orientation of Proteins in Membranes (OPM) database. The distance between the wild-type PS1-N190 or mutant PS1-N190X and PEN2-L98 was measured utilizing PyMOL. The asparagine at position 190 of PS1 was substituted with BpA via WinCoot. Panel A shows the distance between PS1-N190 and PEN2-L98, with a magnified section shown within the highlighted box. Panels B and C show the magnified distance between the mutant PS1-N190X and PEN2-L98.

3.2.2 γ -secretase constructs

The PEN2/NCT/ApH1 gene was cloned into the vector pET-27b. As depicted in Figure 3.13 A, several recombinant plasmids are successful; for instance, the lane labeled with pET27b-Flag-PEN2 after digestion-1/2 shows that the sizes of the two bands align with those of the linearized pET27b and Flag-PEN2 individually. A comparable observation is evident in the lane labeled with pET27b-ApH1-HA after digestion-2/3/4/5 (Figure 3.13 B), as well as in the lane labeled with pET27b-NCT-rho after digestion-1/2/3 (Figure 3.13 C). These constructed plasmids will serve as templates in subsequent site-directed

Results

mutagenesis and in vivo/in vitro protein expression systems following verification through sequencing.

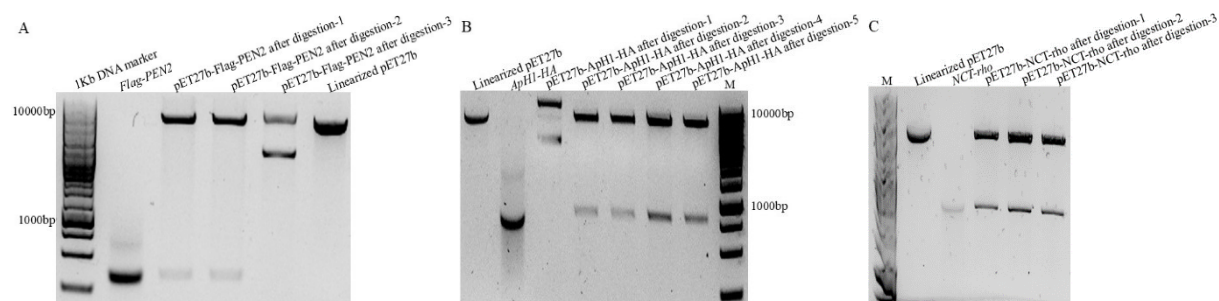


Figure 3.13 pET-27b-Flag-PEN2/ApH1-HA/NCT-rho Construct

XhoI and NdeI were used to perform double digestion assays to verify the constructed plasmids. As a control, pET27b was linearized using XhoI, and the inserted genes (Flag-PEN2, ApH1-HA, and NCT-rho) were amplified via PCR. As demonstrated in the gel electrophoresis, panels A, B, and C display the results for pET27b-Flag-PEN2, pET27b-ApH1-HA, and pET27b-NCT-rho, respectively.

3.2.3 sjGST constructs

To exploit the photocrosslinking assay, the sjGST gene was reconstructed into the vector pET-27b by the classical biomolecular techniques (PCR and double enzyme digestion). The reconstructed plasmids were extracted and verified by double enzyme digestion. As shown in Figure 3.14, the lane labeled “Digested pET27b-GST-1/2” includes two bands, which are aligned with the control bands of the lanes labeled “Digested pET27b” and “GST” individually. After sequencing, the confirmed plasmid will be used as a template in subsequent site-directed mutagenesis and in vitro protein expression systems to establish the photocrosslinking technique.

Results

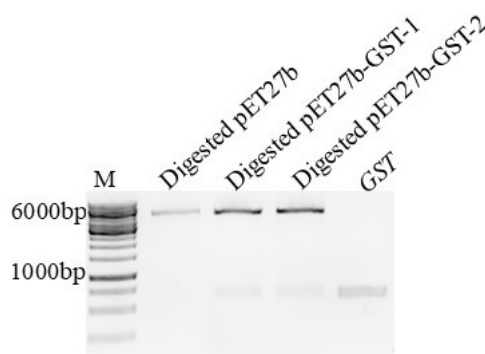


Figure 3.14 pET-27b-GST Construct

XhoI and NdeI were used to perform double digestion assays to verify the constructed plasmids. As a control, pET27b was linearized using XhoI, and the inserted gene (*GST*) was amplified via PCR.

3.3 In-vivo incorporation of UAAs

The pET/pEVOL co-transformed system in BL21(DE3) represents a strategic approach for introducing unnatural amino acids into the GFP protein. Given that the efficiency of UAA incorporation is affected by RF1, the RF1-deficient B95.ΔA strain was procured to utilize this system for expressing the membrane protein PS1-N190X. In vivo expression experiments were conducted, and the results are illustrated in Figure 3.15. As demonstrated in lanes PS1 and PS1-N190X, membrane proteins, namely PS1 (54.8 kDa) and the truncated PS1-N190X (23.9 kDa), were successfully expressed in B95.ΔA following induction with 1 mM IPTG. Conversely, there was no detection of full-length PS1 or truncated PS1-N190X after co-transformation with pET15b-PS1 or pET15b-PS1-N190X alongside pEVOL-pBpF into B95.ΔA, as indicated by lanes PS1+pBpF and PS1-N190X+pBpF. Furthermore, the full-length PS1-N190X was not expressed even upon supplementation with BpA, as shown in lane PS1-N190X+pBpF+BpA. These findings suggest that the pET/pEVOL co-transformed system is unsuitable for the incorporation of BpA into PS1-N190X within the B95.ΔA strain. Consequently, subsequent investigations focused on strategies for incorporating unnatural amino acids into proteins through an in vitro protein expression system.

Results

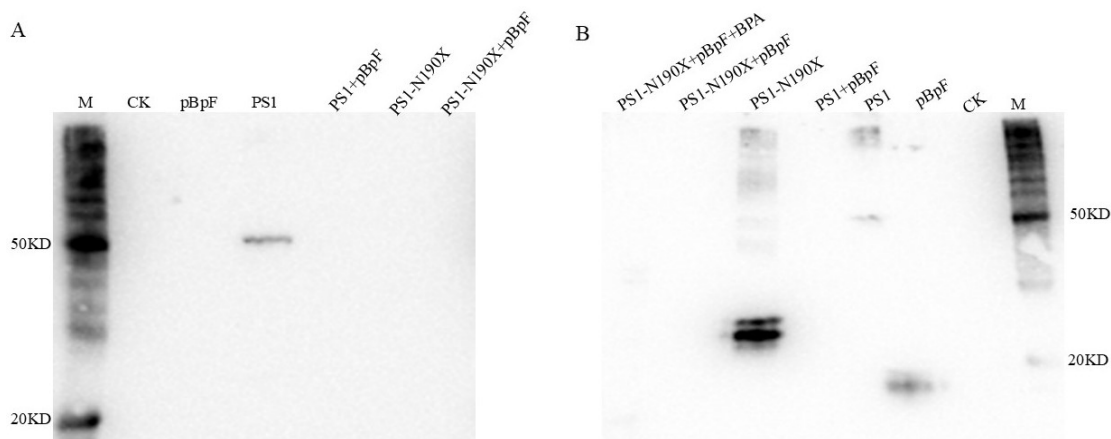


Figure 3.15 The incorporation of BpA to PS1 by the in vivo assay

Plasmids pET15b-PS1-N190X and pEVOL-pBpF were cotransformed into the B95.ΔA strain to incorporate BpA into PS1-N190X. As controls, pET15b-PS1, pET15b-PS1-N190X, and pEVOL-pBpF were transformed into the B95.ΔA strain individually, and pET15b-PS1 was cotransformed with pEVOL-pBpF. Panels A and B are the results of the Western blot. Panel A shows samples before induction. Panel B shows samples after induction. PS1-N190X means the truncated presenilin 1. PS1 means the full-length presenilin 1. pBpF means pEVOL-pBpF was transformed. BPA means that the unnatural amino acid BpA was added to the 2xYT media. CK means there were no plasmids transformed.

3.4 In-vitro incorporation of UAAs

Two methods were employed to incorporate BpA into PS1-N190X. The first is the co-expression of BpA-RS and PS1-N190X by combining the pET/pEVOL system with an in vitro protein expression system. The second is expressing PS1-N190X in the in vitro system, accompanied by the addition of purified BpA-RS and extracted BpA-tRNA.

3.4.1 The pET/pEVOL system

To incorporate UAAs into the target protein by the in vitro protein expression system, pEVOL-pBpF was transformed into B95.ΔA to overexpress BpA-RS and tRNA-BpA. Based on this, the lysate was prepared by incubation with preincubation buffer. As the control, commercial lysate was employed in the in vitro protein expression system. Considering the potential influence of pEVOL-pBpF on B95.ΔA strain, the growth curves of B95.ΔA and B95.ΔA containing pEVOL-pBpF were evaluated in 2xYT media. As illustrated in Figure 3.16 A,

Results

following the transformation of pEVOL-pBpF into B95.ΔA, its growth rate was notably sluggish, and the lag phase was prolonged when compared to the growth trend of B95.ΔA. Upon harvesting cells at OD600 = 3.0, the two cell types, B95.ΔA and B95.ΔA+pBpF, underwent distinct treatments to produce lysates, the expression efficiencies of which were analyzed via Western blot, as depicted in Figure 3.16 B. The commercial lysate successfully expressed PS1 but failed to express PS1-N190X. This indicates that the pET/pEVOL co-transformed system did not facilitate the incorporation of BpA into PS1-N190X through the in vitro system based on commercial lysate, which was observed in the in vivo expression assay. To address this issue, purified BpA-RS and extracted BpA-tRNA may present viable strategies for enhancing BpA incorporation. Furthermore, all homemade lysates, irrespective of the preincubation buffer treatment, were unsuccessful in expressing PS1-N190X, potentially attributable to the influence of cells harvested during the late exponential phase on the lysate preparation. Consequently, utilizing lysates from various growth stages may constitute a promising strategy to enhance the expression efficiency of the B95.ΔA-based in vitro protein expression system. These proposed strategies will be validated in subsequent experiments.

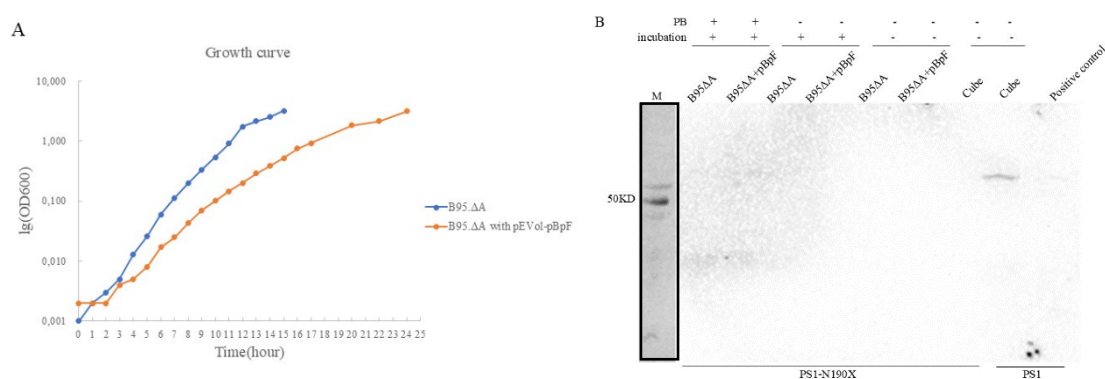


Figure 3.16 In-vitro incorporation of BpA to PS1-N190X using the pET/pEVOL system

In vitro incorporation of BpA into PS1-N190X was conducted by adding pET15b-PS1-N190X and pEVOL-pBpF into the in vitro reaction system, where the lysate derived from the B95.ΔA strain was treated with the preincubation buffer. Panel A shows the growth curve of B95.ΔA

Results

with (orange curve) or without (blue curve) plasmid pEVOL-pBpF in 2xYT media at 37°C and 120 rpm. Panel B shows the Western blot results from the in vitro assay based on variable treatments. B95.ΔA means the lysate from the B95.ΔA strain. B95.ΔA+pBpF means the lysate after transforming pEVOL-pBpF into B95.ΔA. Cube means commercial lysate from Cube Biotech. PB means preincubation buffer. “+” means this step was performed, while “-” means the opposite. PS1 means presenilin 1 protein was supposed to be expressed, and PS1-N190X means presenilin 1 mutant (PS1-N190X) was to be expressed. Positive control means the PS1 sample expressed from *E. coli*.

3.4.2 The addition of BpA-RS and BpA-tRNA

To validate the feasibility of the second proposed strategies, the constructed pET27b-His-TEV-BpA-RS/pET27b-His-TEV-AzF-RS and pEVOL-BpA-tRNA were employed to overexpress the synthetase and tRNA for unnatural amino acids, respectively. Considering potential factors, the B95.ΔA lysate was prepared using various methods, such as different media, cell harvesting times, cell disruption pressures, and the overexpression of T7 RNAP in B95.ΔA. Additionally, the concentrations of Mg^{2+}/K^{+} and T7 RNAP were optimized. To assess the expression efficiency of the in vitro system, EGFP was expressed under diverse conditions within the in vitro expression system, with its fluorescence measured using a plate reader.

3.4.2.1 Preparation of B95.ΔA Lysate

E. coli lysate includes all components of prokaryotic transcription and translation in the in-vitro protein expression system, especially 70S ribosomes. To maintain the activity of ribosomes, *E. coli* lysate is preferred to be opened by the French Press. As an alternative, a Microfluidics M-110P device was employed to disrupt cells in this work. However, the target proteins were not expressed using the lysate prepared by the previous method. Therefore, a new method was exploited, such as the application of 2xYTPG media, harvesting cells at different OD600, disrupting cells at low pressure, 700 bar, and pAR1219 transformation to B95. ΔA, and lysate lyophilization. As shown in Figure 3.17 A, B95. ΔA, the RF1 deficiency strain, grows at the same growth rate as BL21 in

Results

2xYTPG media, which could improve the protein expression efficiency of the in vitro system due to the production of ribosomes of high quality in the lysate.

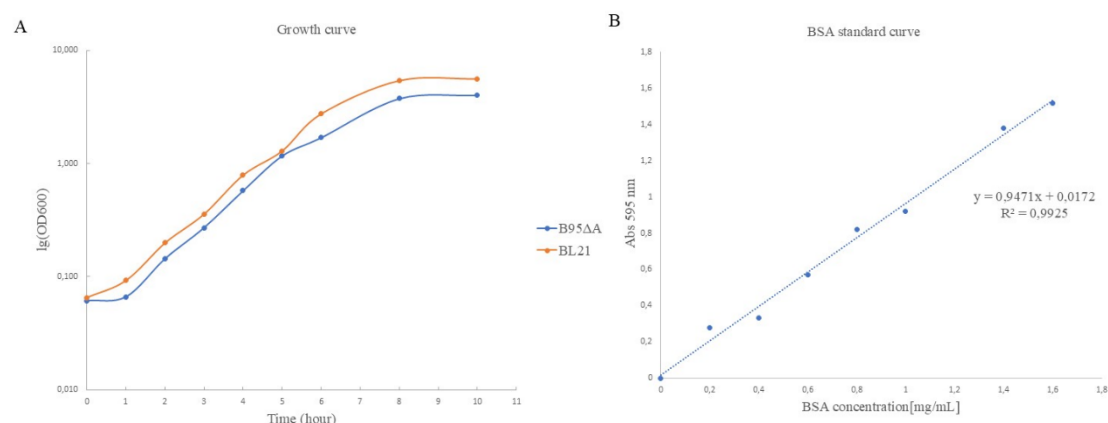


Figure 3.17 The growth curve and BSA standard curve

Panel A shows the growth curve of B95.ΔA and BL21 in 2xYTPG media at 37°C and 120 rpm. Panel B shows the BSA standard curve to measure the total protein concentration in the lysate.

The EGFP yield increases by 35.3% after utilizing 2xYTPG medium (Table 3.1 lysate 0 and 1), which demonstrates that the application of 2xYTPG medium enhances the synthesis efficiency of the in vitro protein expression system derived from the B95.ΔA strain. Compared with lysate 1 in Table 3.1, lyophilized lysate 2 increases the EGFP yield by 13-fold, despite its total protein concentration being around 2-fold higher. This suggests lyophilization of lysate could be one method to enhance the yield of the in-vitro protein expression system. The EGFP yield will not improve synchronously with the total protein concentration of the lysate. An alternative to lyophilization could be processing the lysate with the run-off step at 37°C, as the EGFP yield increases 10-fold, as shown in lysate 4, compared to 42°C, as demonstrated in lysate 3. This suggests that 42°C may inhibit the activity of endogenous T7 RNA polymerase, which is supported by experiments with lysates 5, 6, and 7, where all samples overexpressed T7 RNA polymerase and were processed with the run-off step at 37°C after harvesting cells at different OD600. This strategy significantly increases the EGFP yield, approximately 58-fold compared to that produced

Results

from lysate 4, demonstrating that the amount of T7 RNA polymerase plays a crucial role in the in vitro protein expression system. Collectively, cells should be harvested before the OD600 reaches 2.5, and the addition of T7 RNA into CFPS could increase the yield of expressed proteins. Therefore, the following experiment tries to determine how to enhance the efficiency of the in vitro protein expression system by the addition of T7 RNA polymerase.

Table 3.1 Factors on lysate preparation

Samples name	Lyophilization	Run-off (°C)	OD600	pAR 1219	Total protein Concentration(mg/ml)	EGFP yield(ng/μl)
Lysate 0	-	-	3.0	-	39.47 ± 1.25	0.17 ± 0.42
Lysate 1	-	-	3.0	-	36.33 ± 2.07	0.23 ± 0.08
Lysate 2	+	-	3.0	-	70.74 ± 3.45	3.14 ± 0.16
Lysate 3	-	42	3.0	-	69.80 ± 1.78	0.00 ± 0.00
Lysate 4	-	37	3.0	-	49.78 ± 0.98	2.30 ± 0.42
Lysate 5	-	37	3.0	+	78.02 ± 3.52	119.01 ± 2.27
Lysate 6	-	37	2.5	+	38.12 ± 1.76	126.04 ± 3.33
Lysate 7	-	37	2.0	+	58.35 ± 1.05	124.05 ± 2.37

Notes: “-” means the treatment isn’t employed. “+” means the treatment is employed. Lysate 0 is the lysate prepared from the B95.ΔA growing in 2xYT medium. Lysate 1-7 is the one from the B95.ΔA growing in 2xYTPG medium.

3.4.2.2 Optimizing T7 RNA polymerase

In previous T7 RNAP purification experiments, we found that T7 RNAP inhibited the transcription process at high concentrations. This high-concentration inhibitory effect is observed in the in vitro protein reaction system, as shown in Table 3.2, where the yield of EGFP decreases as the concentration of purified T7 RNA polymerase increases in the in vitro protein expression system. The optimal amount of T7 RNA polymerase is 4.25 μg in 50 μl batch mode CFPS,

Results

namely 85 ng/μl, when the EGFP yield is similar to the yield from lysate 6/7 in Table 3.1. Therefore, the in vitro protein expression system will replace the commercial T7 RNAP with the purified T7 RNA polymerase in further experiments.

Table 3.2 The yield of EGFP in various in vitro systems

Samples	Amount μg / 50μl	EGFP Yield (ng/μl)
Homemade T7 RNAP	12.5	90.05 ± 0.35
Purified T7 RNAP	2.5	87.23 ± 0.23
Purified T7 RNAP	4.25	123.11 ± 0.19
Purified T7 RNAP	8.45	78.37 ± 0.22
Purified T7 RNAP	12.5	65.92 ± 0.41
Purified T7 RNAP	33.75	41.47 ± 0.33

3.4.2.3 The optimization of Magnesium and Potassium ions

Magnesium ions and potassium ions concentrations are vital factors in the yield of the in-vitro protein expression system, so the optimum experiment was performed in this part. As shown in Figure 3.18 A and B, the EGFP concentration rises first and then decreases with the increase of magnesium acetate. The optimal concentration of magnesium ions is 12 mM, less than the original concentration of 16 mM in the in vitro protein expression system, which corroborates the previous research that at the magnesium concentration of 12 mM, CTA yield was the highest when 33 mM PEP was used as the energy source in CFPS (Tae-Wan Kim 2006). Based on this result, the concentration of potassium ions was optimized. Figure 3.18 C shows that it is beneficial to increase the EGFP concentration when the concentration of potassium ions is less than 270 mM in the in vitro protein expression system, and the optimum is 170 mM. These results demonstrate that a high concentration of magnesium or

Results

potassium decreases the yield of the in-vitro protein expression system. Therefore, 12 mM magnesium acetate and 170 mM potassium acetate will be employed in further in vitro protein expression systems.

Until now, the in vitro protein expression system has been established to perform protein expression. The following task is to incorporate BpA into PS1-N190X with the addition of BpA-RS and BpA-tRNA.

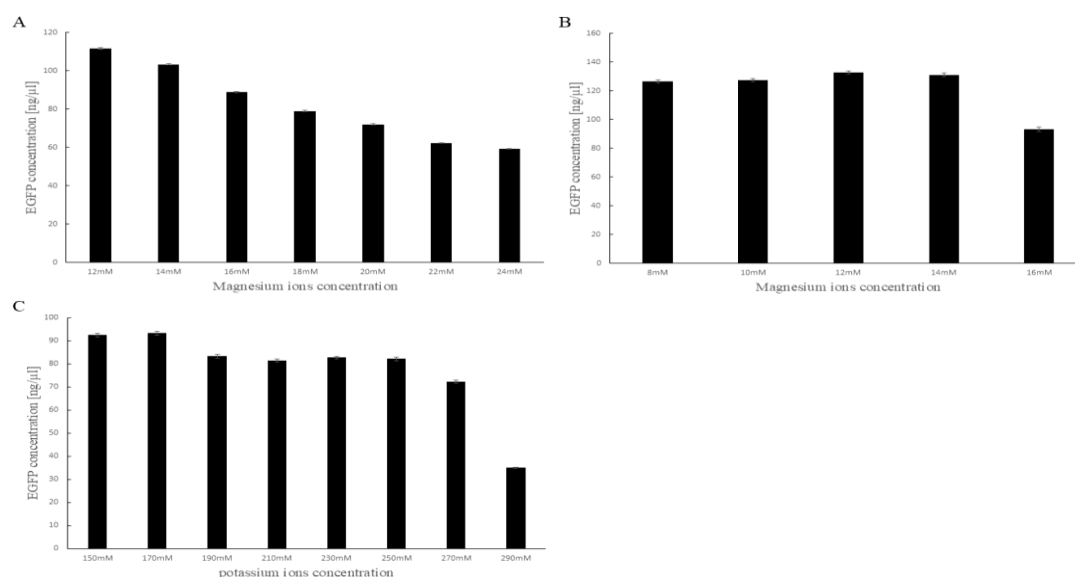


Figure 3.18 The optimization of magnesium and potassium ions

Bar charts A and B show the screening of magnesium ion concentration in the in-vitro protein expression system. Bar chart C shows the screening of potassium ion concentration in the in-vitro protein expression system.

3.4.2.4 The addition of BpA-RS and BpA-tRNA

After establishing the in vitro protein expression system, unnatural amino acid BpA was incorporated into membrane protein PS1-N190X by adding BpA, BpA-tRNA, and BpA-RS. To provide the native environment, liposomes were added to the in vitro reaction system to stabilize the expressed PS1-N190X. Additionally, due to TEV residing in the purified BpA-RS (Figure 3.5 C), excess TEV was added to the in vitro protein expression system to analyze the effect on the PS1-N190X expression.

Results

In terms of the influence of TEV, Lanes 4 and 6 indicate no negative effect of TEV on PS1-N190X expression in Figure 3.19. Subsequently, the influence of the pET-plasmid was detected. Lanes 4 and 8 show that the established in vitro protein expression system could express full-length PS1 when only pET15b-PS1 was added into the reaction system, as the *E. coli*-derived in vivo expression system does. This is confirmed by the addition of pET15b-PS1-N190X (lanes 1 and 9). These findings indicate that the pET-plasmid is a suitable template to express the target protein in CFPS.

In terms of the influence of the addition of BpA-RS, BpA-tRNA, and BpA, lanes 1, 2, and 3 indicate that the in vitro protein expression system expresses truncated PS1-N190X without the addition of BpA, and the amount of truncated PS1-N190X decreases only when the components for BpA incorporation, BpA, BpA-tRNA, and BpA-RS, are supplemented. However, in lanes 1 and 3, an uncertain band of the same size as the full-length PS1 is likely due to the misrecognition of BpA-tRNA as a natural amino acid.

These results indicate that the method of adding BpA-RS and BpA-tRNA into the established CFPS is an available strategy to incorporate unnatural amino acids into the target protein. Therefore, the strategy of photocrosslinking will be tested in the next chapter.

Results

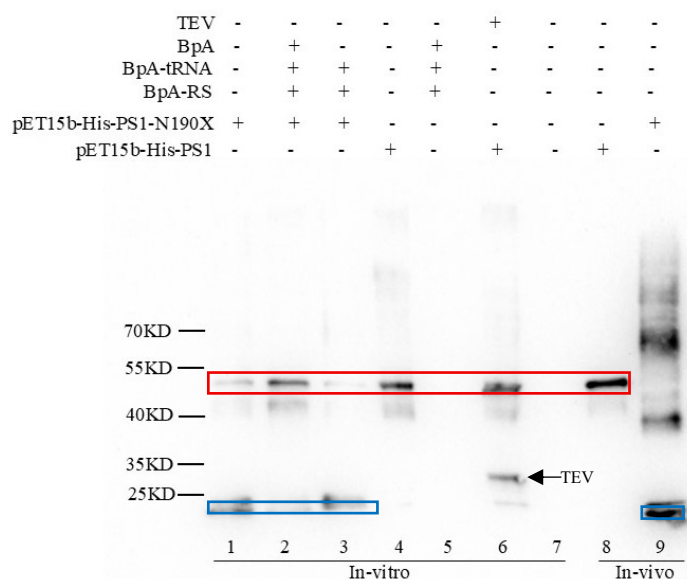


Figure 3.19 Incorporating BpA to PS1-N190X

The expression of PS1-N190X or PS1 under different conditions by in vitro or in vivo protein expression systems. “+” means the component is added into the in-vitro reaction system while “-” means the opposite of “+”. The full-length PS1/PS1-N190X is enclosed in the blue frame, and the truncated PS1-N190X is enclosed in the red frame. The TEV band is indicated by a black arrow.

3.5 Photo-crosslinking test with sjGST-F52X

According to the crystal structure of Glutathione S-transferase from *Schistosoma japonica* (Michele A. McTigue 1995), Jason W. Chin and his colleague used the dimeric sjGST, as shown in Figure 3.20, to perform a model crosslinking assay (Jason W. Chin 2002a). They proved that >50% of mutant sjGST-F52X were crosslinked after irradiation for 5 minutes with a UV lamp at 360 nm. Therefore, this mutant was employed to set up the photo-crosslinking assay in this work.

Figure 3.21 A shows gene *sjGST* is successfully reconstructed to plasmid pET27b, which is expressed in *E. coli* as shown in the lane GST of Figure 3.20 B, where 20% of sjGST-F52X is crosslinked after irradiation for 5 minutes with a 352-nm UV lamp, and the efficient photo-crosslink increases with the extension of illumination time. As shown in Figure 3.21 C, the photo-crosslink efficiency is up to 100% after irradiation for >45 minutes; therefore, this

Results

illumination time will be used for further experiments. The concentration of crosslinked sjGST-F52X is 0.286 mg/ml after purification with Ni-Indigo MagBeads. Surprisingly, two photocrosslinked bands appear after the sample is illuminated for longer than 5 minutes, likely due to conformation changes in the sjGST-F52X dimer from the extended to the compact form, as one covalent bond is generated, to two covalent bonds are generated.

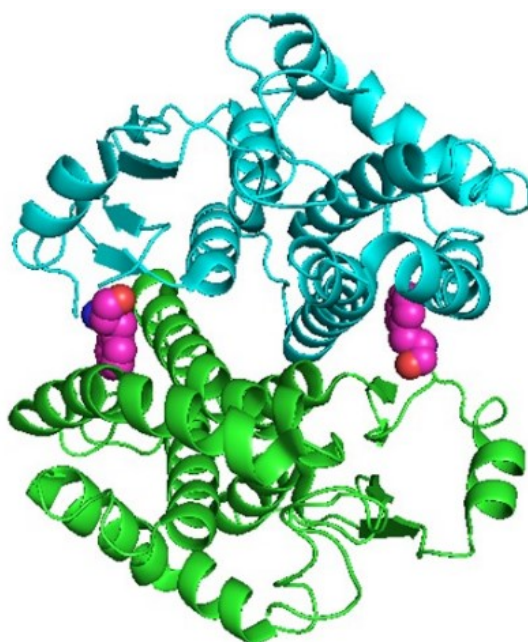


Figure 3.20 The crystal structure of sjGST, adapted from (Jason W. Chin 2002a)

Monomers of the sjGST dimer are shown in green and cyan. The residue F52 is shown in a sphere mode and colored by elements for each monomer.

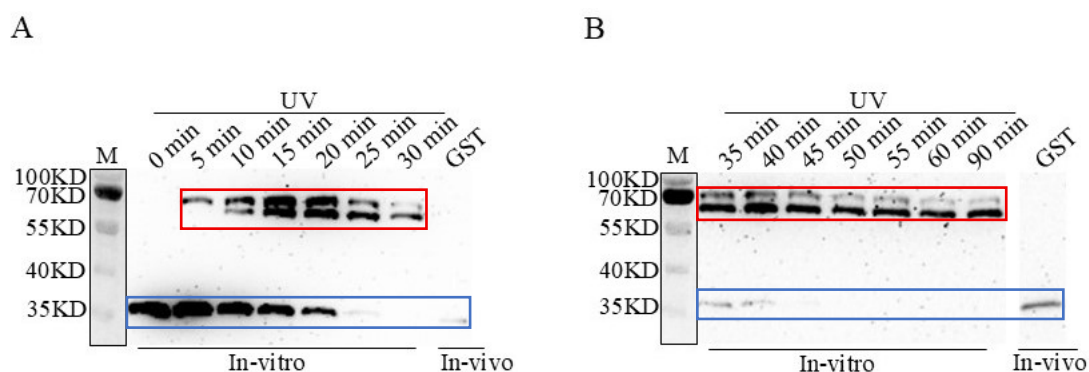


Figure 3.21 Photo-crosslinking test with sjGST-F52X

BpA was incorporated into sjGST by the established CFPS, followed by irradiation with a 352-nm lamp at different times. Subsequently, all irradiated samples were purified using Ni-Indigo MagBeads from Cube Biotech and visualized by Western blot. Panels A and B show the

Results

Western blot results of purified sjGST-F52X after irradiation. The monomer sjGST-F52X is enclosed in the blue frame, and the covalent dimer is enclosed in the red frame.

3.6 Photo-crosslinking of PS1-N190X and PEN2

To crosslink PS1-N190X and PEN2, the first challenge is folding membrane proteins in the right conformation. Therefore, different hydrophobic environments were exploited to be the receiver of PS1-N190X and PEN2 subcomplex.

3.6.1 Preparation of liposomes and CHAPSO-destabilized liposomes

Liposomes or CHAPSO-destabilized liposomes were prepared by sequential extrusion combined with thin-film hydration and Freeze-thaw cycles. To prepare CHAPSO-destabilized liposomes, 100 nm liposomes were titrated with CHAPSO buffer in different ratios. In this process, as shown in Figure 3.22 A, the optical density trends are consistent across wavelengths from 250 nm to 600 nm upon titrating liposomes with CHAPSO in differing ratios. Additionally, since CHAPSO exhibits no absorption at 280 nm, we focused our data analysis on the optical density measurements at this wavelength.

Figure 3.22 B illustrates three distinct phases: Phase I, incorporation of CHAPSO monomers into the bilayer until the CHAPSO concentration reaches the critical micelle concentration (CMC), where the liposomes are saturated with CHAPSO (R_{sat}), causing their diameter to increase to a maximum. This results in the observed peak of optical density at 280 nm; Phase II, gradual solubilization of liposomes occurs as CHAPSO concentration continues to rise. At this stage, CHAPSO-lipid micelles start to form, leading to a significant reduction in optical density; Phase III, the complete solubilization of liposomes (R_{sol}) causes the micelles to become so small and the diameter was around 12 nm (Josep Cladera 1997) that the solution cannot scatter light and the optical

Results

density decreases to the minimum. These results show that 3.8 μM of CHAPSO is the optimum concentration to destabilize the 100 nm liposomes.

Dynamic Light Scattering (DLS) was performed to assess the size distribution and homogeneity of liposomes, CHAPSO-titrated liposomes, or CHAPSO-destabilized liposomes. In Figure 3.23, the intensity-based size distributions reveal a dominant single peak for all samples at 20°C, confirming their stability and uniformity. In addition, a minor peak near 1 nm is observed in untreated liposomes (Figure 3.23 A), likely due to residual contaminations, such as unincorporated lipids or debris. Notably, this small peak is absent in CHAPSO-titrated liposomes or CHAPSO-destabilized liposomes before or after dialysis, suggesting effective removal of contaminants with CHAPSO.

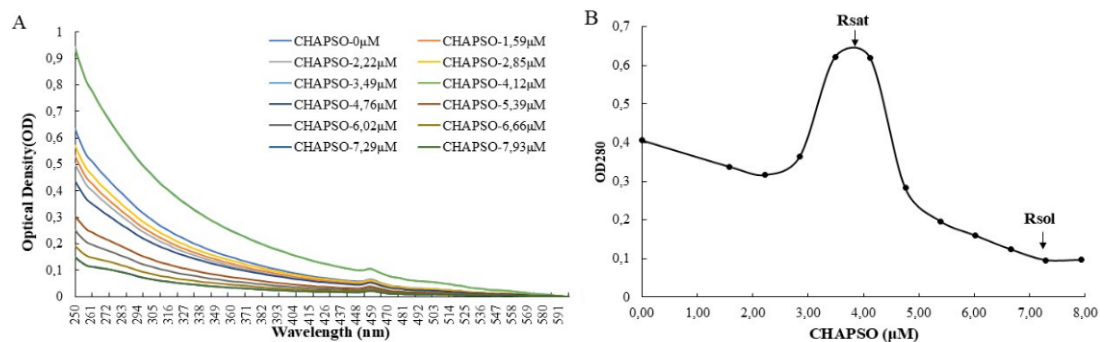
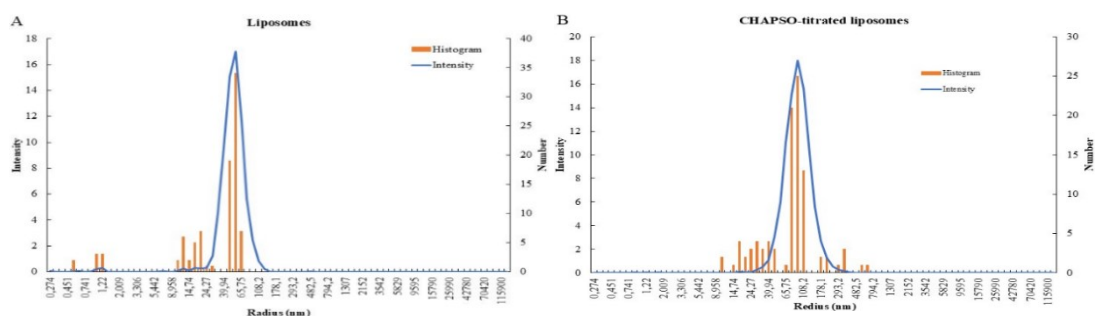


Figure 3.22 Liposome titration with CHAPSO

Panel A shows the absorbances of the liposomes-CHAPSO mixture with variable ratios at different wavelengths ranging from 250nm to 600nm. Panel B shows the absorbances of the liposomes-CHAPSO mixture in different ratios at 280nm, whose plot is derived from data in panel A. Rsat: the detergent-to-lipid ratio concentration where the liposomes are saturated. Rsol: the detergent-to-lipid ratio where the liposomes are solubilized and fully converted into mixed micelles.



Results

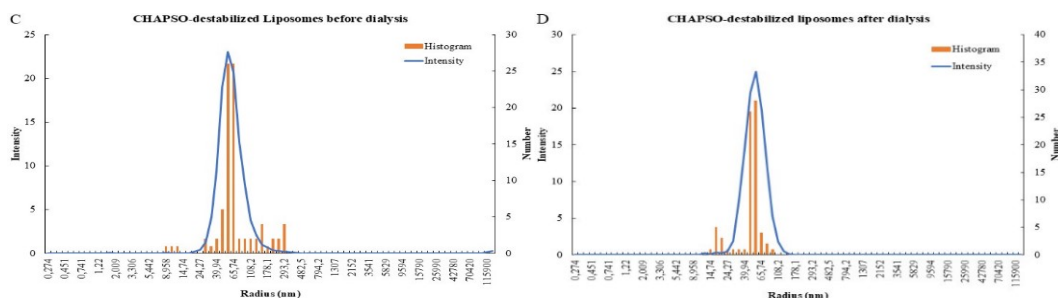


Figure 3.23 Dynamic Light Scattering for prepared liposomes

The intensity chart (blue line) and the histogram chart (orange bar) of different samples after 10-fold dilution. Panel A shows the result of liposomes; Panel B is CHAPSO-titrated liposomes; Panels C and D are CHAPSO-destabilized liposomes, with the difference being that D is the sample after overnight dialysis at 4°C.

Table 3.3 shows that all samples exhibit a polydispersity index (PDI) less than 0,3, indicative of moderate heterogeneity. While moderate PDI values can signal aggregation or small particles, the presence of a predominated single peak in size distributions rules out aggregation, attributing the polydispersity to trace contaminants or inherent sample diversity.

3.8 μM of CHAPSO induces significant liposome swelling, increasing the average radius from 52,15 nm (untreated liposomes) to 89.38 nm, consistent with CHAPSO-mediated liposome titration. In contrast, CHAPSO-destabilized liposomes show a post-dialysis radius reduction from 56.57 nm to 53.14 nm, closely matching the size of untreated liposomes (~ 52 nm). Additionally, their molecular weight (30.40 MDa) aligned with that of untreated liposomes, strongly suggesting CHAPSO removal during dialysis. These findings imply that overnight dialysis eliminates CHAPSO from destabilized liposomes, effectively reverting them to their native state. To preserve CHAPSO-induced destabilization in future preparations, overnight dialysis should be omitted from the protocol. However, this step can be employed to strip CHAPSO from protein-inserted liposomes, thereby reversing the protein to the native membrane environment.

Table 3.3 DLS results after individual measurements

Sample name	Radius (nm)	PDI (%)	Molecular weight (MDa)
Liposomes	52.15 ± 3.35	22.3	29.13
CHAPSO-destabilized Liposomes	89.38 ± 8,43	26,6	100,6
CHAPSO-destabilized Liposomes before dialysis	56,57 ± 5,75	21,7	35,12
CHAPSO-destabilized Liposomes after dialysis	53,14 ± 4,94	22,9	30,41

3.6.2 PS1-N190X/PEN2 in liposomes or CHAPSO-destabilized liposomes

Liposomes and CHAPSO-destabilized liposomes were utilized to co-express PS1-N190X and PEN2 via the established in vitro protein expression system. After irradiation, samples were analyzed by Western blot. Figure 3.24 A shows that PS1-N190X tetramers are present when it is expressed without PEN2, rather than when it is co-expressed with PEN2. This phenomenon may be attributed to the presence of PEN2, which diminishes the formation of PS1-N190X tetramers, potentially indicating that PEN2 replaces the PS1-N190X component of this tetramer to form a PS1-N190X and PEN2 subcomplex within liposomes or CHAPSO-destabilized liposomes.

Concerning the modification of the PEN2 band during irradiation, PEN2 fragments exhibit shifts when coexpressed with PS1-N190X (Figure 3.24 B). This may imply that PEN2 interacts with either the N-terminal or C-terminal fragments of PS1-N190X. Since Triton X-100 can dissociate the γ -secretase complex to verify the specificity of BpA-based photo-crosslinking (Steiner, 2016), samples were treated with 1% Triton X-100. The shifted PEN2 fragment was only observed in the liposome sample post-irradiation, instead of the CHAPSO-destabilized sample (Figure 3.24 C). However, this shifted band was not apparent in the western blot of the His-tag in PS1-N190X when liposomes served as the hydrophobic environment. This may be attributable to the low

Results

incorporation efficiency of the PS1-N190X and PEN2 subcomplex when liposomes are used as the hydrophobic environment in CFPS, or the contaminating band. Accordingly, additional purification experiments were conducted using size exclusion chromatography (SEC) in the following section to identify the incorporation efficiency of protein into liposomes.

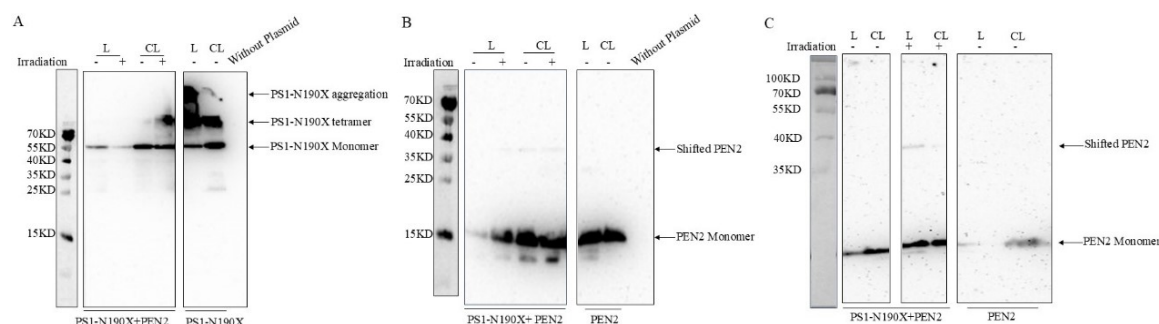


Figure 3.24 Photo-crosslinking PS1-N190X/PEN2 in liposomes and CHAPSO-destabilized liposomes

PS1-N190X and PEN2 were co-expressed with liposomes or CHAPSO-destabilized liposomes, followed by a photo-crosslinking assay. As controls, Panel A shows the Western blot result of PS1-N190X, labeled with a His-tag at the N-terminus, where the samples were not treated with 1% Triton X-100. Panels B and C show the western blot results of Flag-PEN2, labeled with Flag-tag at the N-terminus and treated without or with 1% Triton X-100, respectively. “-” indicates that samples were not irradiated with Crosslinker CL-1 at 352 nm. “+” signifies that samples were treated with irradiation. “L” stands for liposomes, where samples were expressed with liposomes. “CL” is the abbreviation of CHAPSO-destabilized liposomes, where samples were expressed with CHAPSO-destabilized liposomes.

3.6.3 SEC purification of crosslinked PS1-N190X/PEN2

Size-exclusion chromatography (SEC) was performed to determine whether proteins were inserted into liposomes. The profile (Figure 3.25 A) indicates that the liposomes were eluted first from the column (the red line); therefore, the incorporated protein-liposomes are likely to be eluted simultaneously with the same amount of elution buffer. However, the western blot (Figure 3.25 B) indicates that there are no PS1-N190X bands on the polyvinylidene difluoride (PVDF) membrane after utilizing immune-detection techniques. Additionally, PS1-N190X exists in the pellet fractions. These results imply that the

Results

incorporation of PS1-N190X and PEN2 into liposomes is unsuccessful. Therefore, the aggregated protein will be solubilized with detergents in the following experiment.

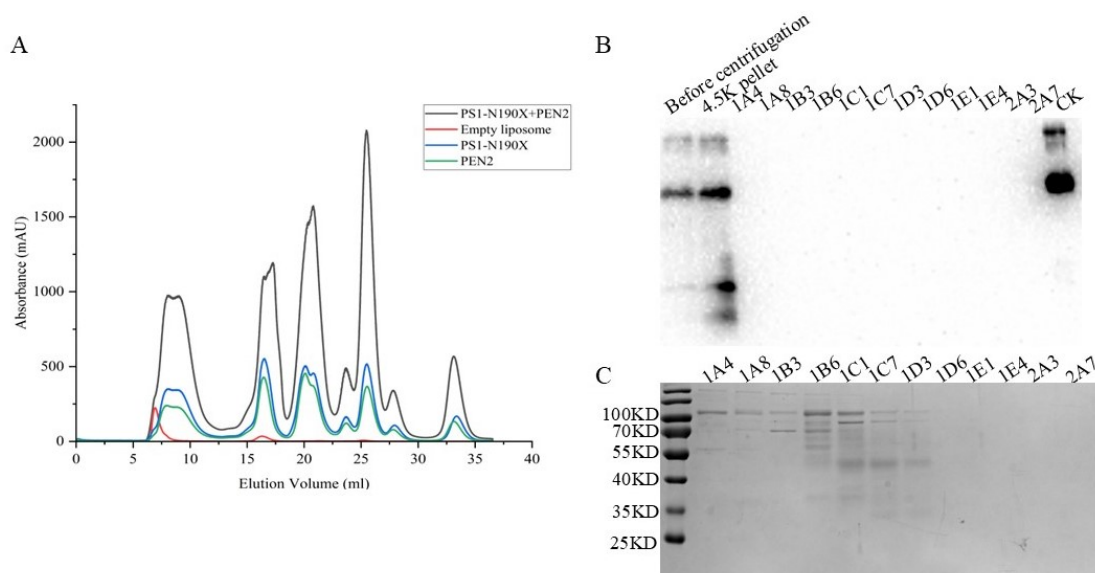


Figure 3.25 SEC Purification of PS1-N190X/PEN2 complex in liposomes

PS1-N190X and PEN2 were co-expressed with liposomes overnight at 30°C, followed by size-exclusion chromatography with a Superose 6 column. As controls, the expressed PS1-N190X, PEN2, and liposomes were purified with the Superose 6 column. Panel A shows the profile of the SEC. Panels B and C show the Western blot and blue stain results of several samples from panel A, respectively. CK means the expressed PS1 in *E. coli* cells.

3.6.4 Photocrosslink of solubilized PS1-N190X/PEN2

To purify the PS1-N190X and PEN2 subcomplex, the expressed PS1-N190X was solubilized using 11 different detergents to determine an optimal solution. As shown in Figure 3.26, phosphocholine group zwitterionic detergents, 1% FOS-choline-14 or FOS-choline-16, except for FOS-choline-12, can completely solubilize PS1-N190X to generate a water-soluble protein-detergent-lipids complex. However, glucoside group non-ionic detergents (OG and NG), maltoside group non-ionic detergents (DM, DDM, and Cy6), and zwitterionic detergents (LDAO, CHPAS, and CHAPSO) exhibit low solubilization efficiency, which is similar to the study from the previous PhD students, Ge and Kun (Ge

Results

Yang 2018). Since Kun found FOS-choline-16 purified presenilin 1 was more compact than FOS-choline-14 purified presenilin 1 (Yu 2016), FOS-choline-14 will be used to solubilize the PS1-N190X and PEN2 subcomplex.

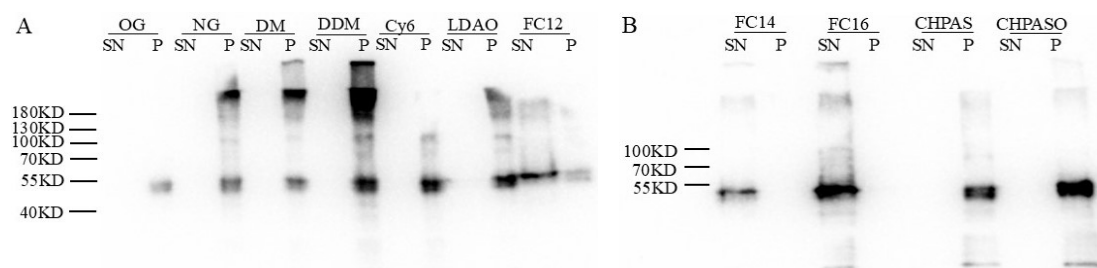


Figure 3.26 PS1-N190X solubilization

After solubilization with different detergents, samples were centrifuged. Panels A and B show the Western blot results on the His-tag of PS1-N190X after solubilization with different detergents. OG is the abbreviation of n-Octyl- β -D-glucopyranoside; NG stands for n-Nonyl- β -D-glucopyranoside; DM for n-Decyl- β -D-maltopyranoside; DDM for n-Dodecyl- β -D-maltoside; Cy6 for 6-Cyclohexyl-hexyl- β -D-maltoside; LDAO for N,N-Dimethyldodecylamine-n-oxide; FC12 for FOS-choline-12; FC14 for FOS-choline-14; FC16 for FOS-choline-16; CHAPS for 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate; CHAPSO for 3-[(3-Cholamidopropyl) dimethylammonio]-2-hydroxy-1-propanesulfonate. SN means the supernatant, and P stands for the pellet.

After solubilization with Fos14, co-expressed PS1-N190X/PEN2 was photo-crosslinked and purified using Ni-Indigo Magbeads. Figure 3.27 A shows the SDS-PAGE of purified fractions, illustrating the potential photo-crosslinked subcomplex bound to Ni-Indigo resin, with most being eluted from the resin by the elution buffer, although the wash buffer washed away a small quantity. Figure 3.27 B shows that the band from the eluted fraction is higher than the uncrosslinked PEN2 (13.6 kDa). Furthermore, the Western blot results for the anti-His tag (Figure 3.27 C) indicate that PS1-N190X (54.8 kDa) shifts to a band approximately 70 kDa in size, which is the same as the shifted PEN2. Therefore, these results suggest that PS1-N190X was photo-crosslinked with PEN2 following the solubilization of the PS1-N190X and PEN2 subcomplex. However, this crosslinked band differs from the shifted PEN2 seen in Figure 3.24 C, possibly due to either a shorter co-expression time or the detergent Fos-14,

Results

which inhibits the interaction between PS1-N190X and PEN2. Unfortunately, the mass spectrometry was unsuccessful. Therefore, the photo-crosslinking between PS1-N190X and PEN2 is uncertain.

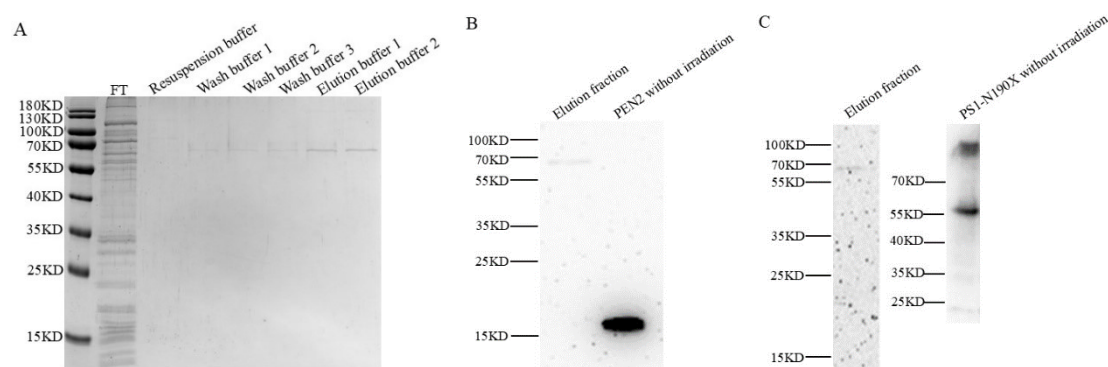


Figure 3.27 Purification of crosslinked PS1-N190X/PEN2 liposome complex

PS1-N190X and PEN2 were co-expressed with liposomes for 4 hours at 30°C. After solubilization, samples were irradiated and purified with Ni-Indigo Magbeads. To observe the shifted bands, the resultant elution fractions were visualized with the Western blot. Panel A shows the blue stain result of Ni-Indigo purification. Panels B and C show the Western blot results of Flag-tag and His-tag, respectively.

3.6.5 PS1-N190X/PEN2 in Sulfo-DIBMA-DMPC or MSP-2N2-DMPC nanodiscs

MSP-2N2-DMPC (diameter approximately 17 nm), and Sulfo-DIBMA-DMPC (diameter approximately 10 nm) nanodiscs were utilized to co-express PS1-N190X and PEN2 via the established in vitro protein expression system. To evaluate whether all expressed proteins were successfully incorporated into the Sulfo-DIBMA-DMPC nanodiscs, all samples underwent ultracentrifugation.

As shown in Figure 3.28, the resulting supernatants and pellets were analyzed via Western blot. These findings suggest that the majority of proteins are integrated into the sulfo-DIBMA-DMPC nanodisc, although the presence of some unfolded proteins across all samples was noted. Notably, full-length PS1-N190X was not detected in the pellet. Additionally, it is significant that PEN2 was only expressed in the presence of PS1 or PS1-N190X, rather than

Results

independently (Figure 3.28 B and D), which diverges from the observed results when PEN2 is expressed within liposomes or CHAPSO-destabilized liposomes. This discrepancy is likely attributable to the less amount of pET27b-Flag-PEN2 added during CFPS when utilizing Sulfo-DIBMA-DMPC nanodiscs as the hydrophobic environment than when using liposomes or CHAPSO-destabilized liposomes. This is also consistent with the theory that PEN2 accumulates depending on PS1 (Naoto Watanabe 2005).

What's more, the shifted PEN2 bands are independent of the irradiation, which is different from the one when liposomes are the receiver. This may imply that PEN2 interacts with the C-terminal of PS1-N190X.

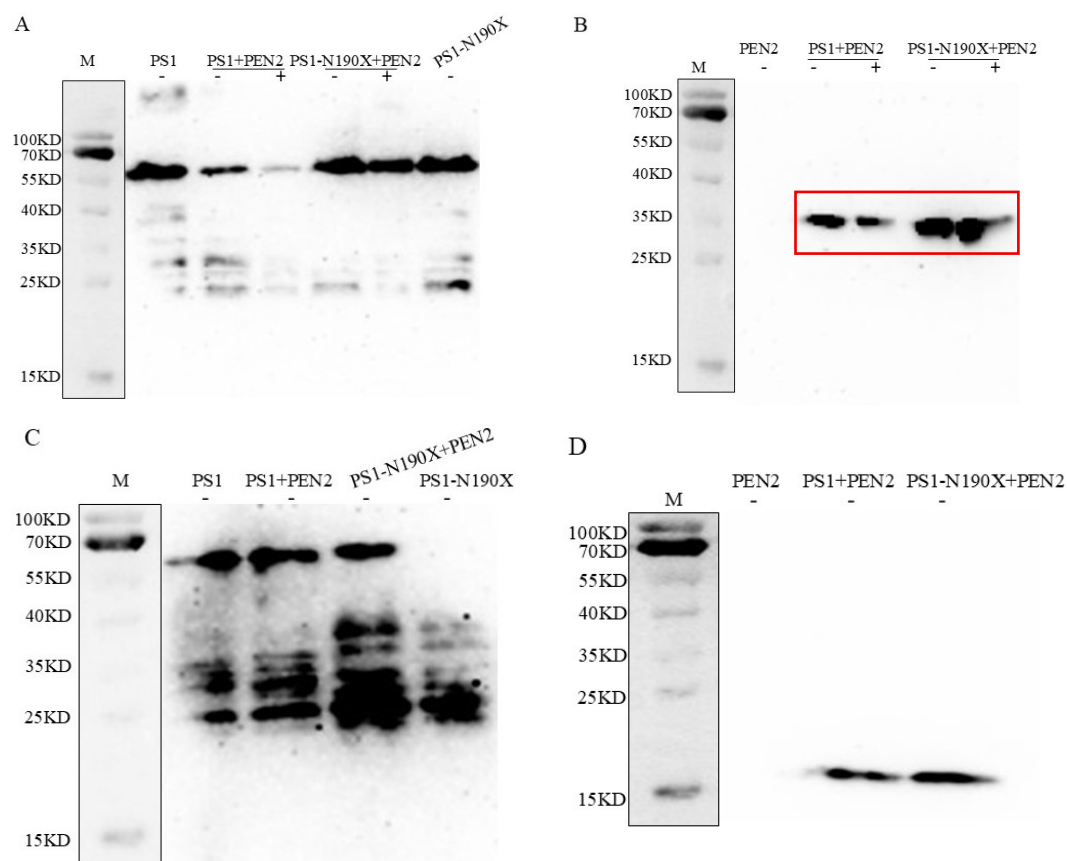


Figure 3.28 The expression of PS1/PS1-N190X and PEN2 subcomplex in Sulfo-DIBMA-DMPC nanodiscs

All proteins were expressed with Sulfo-DIBMA-DMPC nanodiscs by CFPS, followed by ultracentrifugation. Western blot was performed without the treatment with 1% Triton X-100. Panels A and B show the Western blot result of the supernatant for the His-tag and Flag-tag, respectively. Panels C and D show the pellets for His-tag and Flag-tag separately. “-” indicates

Results

that samples were not irradiated with Crosslinker CL-1 at 352 nm. “+” signifies that samples were treated with irradiation. The shifted PEN2 bands are enclosed in the red frame.

After a 1% Triton X treatment, one band of the PS1-N190X sample (red frame in Figure 3.29 A) disappears, which implies that the shifted PEN2 in Figure 3.28 B is not shown in Figure 3.29 B, which is consistent with the former observation when the CHAPSO-destabilized liposomes were the receiver (Figure 3.24). These observations imply that the shifted PEN2 band is not the contaminant.

To quantify the yield of PS1-N190X in CFPS, the home-made purified PS1 was diluted at different times (the right part of Figure 3.29 A). Since the concentration of home-made PS1 is 0.574 mg/ml, the concentration of purified PS1-N190X is around 57.4 ng/μl in the presence of Sulfo-DIBMA-DMPC nanodiscs.

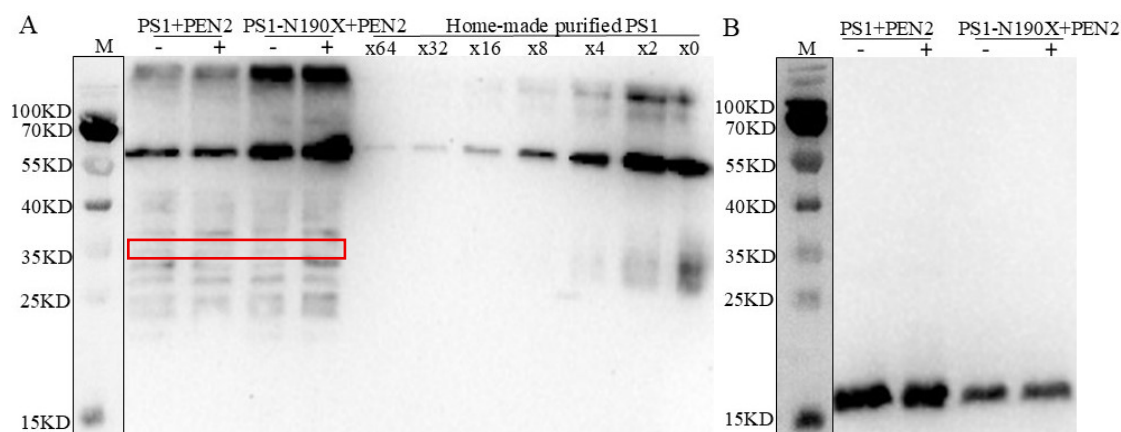


Figure 3.29 The supernatant of PS1/PS1-N190X and PEN2 subcomplex in sulfo-DIBMA-DMPC nanodiscs

All proteins were treated with 1% Triton X-100. Panels A and B show the Western blot result for His-tag and Flag-tag, respectively. “-” indicates that samples were not irradiated with Crosslinker CL-1 at 352 nm. “+” signifies that samples were treated with irradiation. “x numbers” means the dilution multiple.

There are two possible reasons why photocrosslinking between PS1-N190X and PEN2 in the presence of Sulfo-DIBMA-DMPC has failed: either there is no interaction between PS1-N190X and PEN2, or a different configuration has

Results

occurred. To address this question, the irradiated samples were purified by Ni-indigo affinity chromatography. The purification results show that PS1-N190X and PEN2 are present in the elution fraction (Figure 3.30), which is also confirmed by mass spectrometry. The mass spectrometry shows that 46.47% coverage was matched to the PS1 sequence (Figure 3.31 A) and 36.84% coverage was matched to the PEN2 sequence (Figure 3.31 B). The matched sequences are colored in red (Figure 3.31). Therefore, these results exclude the possibility that there is no interaction in the subcomplex, and the high likelihood of unsuccessful photocrosslinking is due to conformational changes in the PS1-N190X and PEN2 subcomplexes.

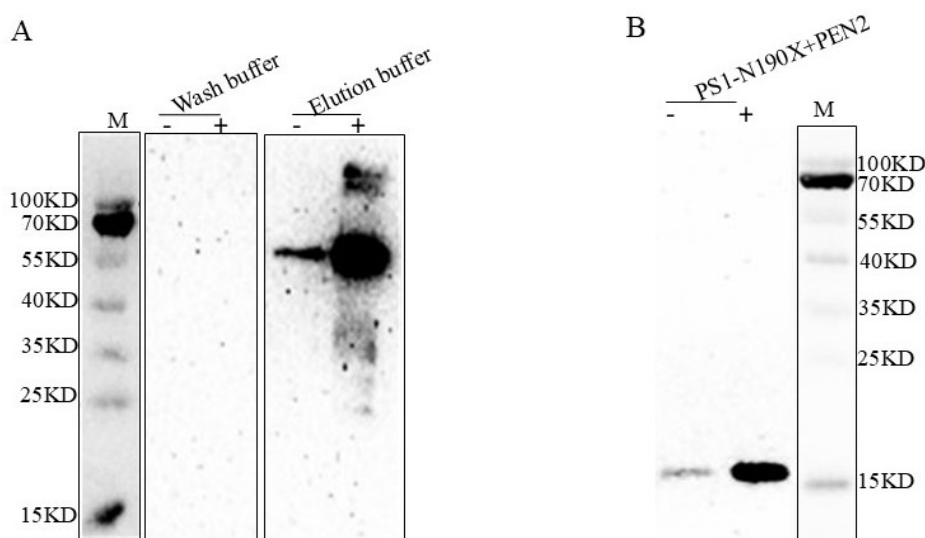


Figure 3.30 Purification of crosslinked PS1-N190X/PEN2/sulfo-DIBMA-DMPC

Ni-indigo magbeads purified Subcomplex PS1-N190X and PEN2 after photo-crosslinking. Panels A and B show the Western blot result for His-tag and Flag-tag, respectively. "-" indicates that samples were not irradiated with Crosslinker CL-1 at 352 nm. "+" signifies that samples were treated with irradiation.

Results

A

MGSSHHHHHSSGLVPRGSHMTELPAPLSYFQNAQMSEDNHLSTVRSQNDNRERQEHNDRRSLGHPEP
LSNGRPQGNSRQVVEQDEEDEELTKYGAKHVIMLFVPVTL~~CMVVVVATIKSVSFYTRKDGQLIYTPFTE~~
~~DTETVGQR~~ALHSILNAAIMISVIVVMTILLVVLYKYRCYKVIHAWLISSLLLLFFFSFIYLGEVFKTYN~~V~~AVD
YITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKG
PLRMLVETAQERNETLFPAL~~IYSS~~TMVWLVNMAEGDPEAQRRVSKNSKYNAESTERESQDTVAENDDGGF
~~SEEWEAQRD~~SHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSLVGKASATASGDWNTTI
ACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYATDYLVPFMDQLAFHQFYI

B

MDYKDDDDKAIEGRNLERSNEEK~~LNL~~CRKY~~Y~~LGGAFLPFLWLNVNIFWFFREAF~~LP~~PAYTEQSQIKGYVWRSVAVGFLFWVIVLTSWI
TIFIYRPRWGALGDYLSFTIPLGTP

Figure 3.31 Peptide identification of purified PS1-N190X/PEN2/Sulfo-DIBMA-DMPC

PS1(A) and PEN2(B) protein sequences are shown. The matched peptides are in red, and the mutant amino acid asparagine (N) is highlighted in yellow.

PS1 undergoes endoproteolysis into PS1-NTF (~28 kDa) and PS1-CTF (~22 kDa) (Gopal Thinakaran and Samuel E. Gandy, 1996), a process partially facilitated by PEN2 (Nobumasa Takasugi 2003; Oliver Holmes 2014). Interestingly, this uncertain endoproteolysis in Sulfo-DIBMA-DMPC is evident in the presence of MSP-2N2-DMPC nanodiscs, but not in liposomes or CHAPSO-destabilized liposomes. This discrepancy may be influenced by the unique structure of nanodiscs. Compared with the random orientation incorporation of PS1/PEN2 into liposomes, nanodiscs permit protein incorporation in two orientations, therefore enhancing the PS1 endoproteolysis with the increasing likelihood of the correct conformation of the PS1-N190X and PEN2 subcomplex.

The PS1-NTF band reappears in the MSP-2N2-DMPC/Sulfo-DIBMA-DMPC-based samples upon western blot analysis following Triton X-100 treatment. Nonetheless, this band disappears after irradiation, suggesting that PS1-NTF may bind to PS1-CTF to form a heterodimer post-endoproteolysis (Figure 3.23 D and 3.25 A). This observation aligns with prior research indicating that PS1-NTF and PS1-CTF, or their homologues, bind to each other (Gang Yu, 1998; Sebastien S. Hebert, 2003).

Results

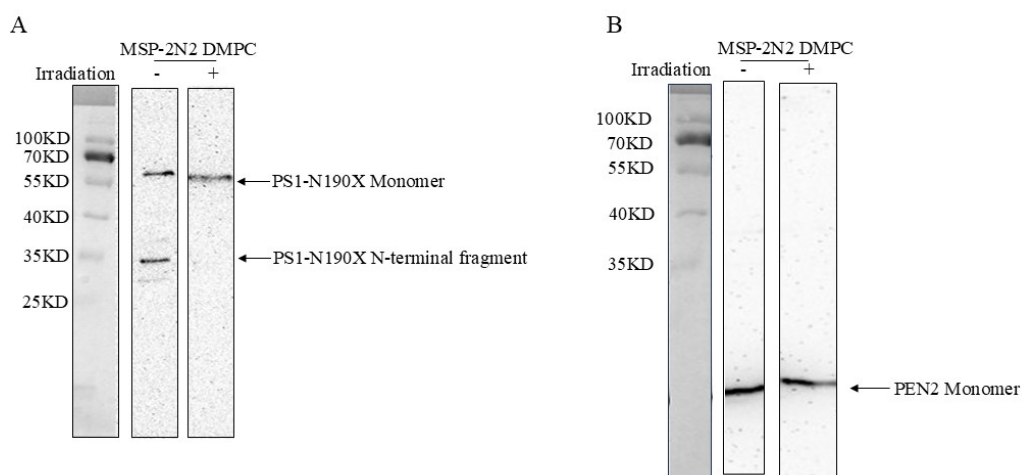


Figure 3.32 photo-crosslinking of PS1-N190X/PEN2/MSP-2N2-DMPC

PS1-N190X and PEN2 were co-expressed with MSP-2N2 DMPC, followed by a photo-crosslinking assay. Panel A shows the western blot result of PS1-N190X labeled with His-tag at the N-terminal, which was treated with 1% Triton X-100. Panel B shows the western blot result of PEN2, labeled with Flag-tag at the N-terminal and treated with 1% Triton X-100. “-” indicates that samples were not irradiated with Crosslinker CL-1 at 352 nm. “+” signifies that samples were treated with irradiation.

4.1 Incorporation of unnatural amino acid into PS1 by in vivo assay

The pEVOL series plasmids, which possess a mid-copy-number p15A origin, are capable of compatible replication with high-copy-number origins such as ColE/pMB1/pBR322/pUC (Mary Munson 1994). Furthermore, a single copy of optimized tRNA and both inducible and constitutive RNA synthetases for unnatural amino acids derived from *M. jannaschii* presents a significant potential for the incorporation of unnatural amino acids into the target proteins expressed in *E. coli*, particularly in the BL21(DE3) strain (Travis S. Young 2010).

However, this strategy proved unsuccessful in this study when B95.ΔA, which is derived from BL21(DE3), was employed to incorporate BpA into PS1-N190X. The most plausible explanation for this outcome is the overexpression of BpA-RS and BpA-tRNA in B95.ΔA cells, which may be toxic to endogenous translation due to potential interactions with the elongation factor Tu (EF-Tu) or ribosomal components (Travis S. Young 2010). This assertion is backed up by the results from the Western blot (Figure 3.3), showing that PS1-N190X wasn't expressed after cotransforming the pEVOL-pBpF plasmid into B95.ΔA. Additionally, the growth curve (Figure 3.4 A) reveals that the B95.ΔA strain grew more slowly in the 2xYT media following the pEVOL-pBpF transformation. These findings suggest that B95.ΔA may not represent an optimal strain for incorporating unnatural amino acids into target proteins using the pET/pEVOL co-transformed system, which requires further validation through additional experiments. One potential approach is to modify the induction conditions, including induction time, temperature, and the order of addition of inducers (Travis S. Young 2010). An alternative strategy would involve the construction of a new pET plasmid that encodes the genes for both BpA-RS and BpA-tRNA, as well as the gene for the target proteins (Takahito Mukai 2015).

4.2 Incorporation of unnatural amino acid BpA into PS1-N190X by in vitro assay

Two strategies were employed to incorporate BpA into PS1-N190X via the coupled transcription-translation cell-free protein synthesis (CFPS) method. The first approach is founded upon the pET/pEVOL co-transformed system, which was intended to co-express BpA-tRNA/BpA-RS alongside PS1-N190X; however, this approach was unsuccessful. It is suggested that B95.ΔA is not the appropriate strain for utilizing the pET/pEVOL co-transformed system for protein co-expression purposes. The alternative strategy involves the combination of purified BpA-RS and extracted BpA-tRNA within the in vitro expression system. During the implementation of this strategy, five factors exert influence on the expression efficiency of CFPS utilizing the B95.ΔA strain.

4.2.1 The 2 x YTPG medium

Phosphatases present in the *E. coli* S30 lysate are capable of degrading PEP to pyruvate and inorganic phosphate (Dong-Myung Kim 1999), as well as dephosphorylating ATP, GTP, CTP, and UTP, which inhibits the regeneration of ATP and NTPs (R.G. Kim 2001). Consequently, the deactivation of phosphatases is utilized to enhance the synthesis efficiency of Cell-Free Protein Synthesis (CFPS). In comparison to employing Cu(OAc)₂ to diminish the activity of acid phosphatases (Yasuaki Kawarasaki 1996) and the immunodepletion of phosphatases (Yasuaki Kawarasaki 1998) in wheat-germ CFPS, the use of 2 x YTPG medium, which contains a high concentration of glucose and inorganic phosphate, represents the simplest approach to inhibit the expression of phosphatases during the preparation of *E. coli* extract. Except for diminishing the activity of phosphatases, 2 x YTPG medium rich in nutrients provides the possibility of rapid culture growth rates, which is important for the preparation of highly productive extract due to its richness in 70S ribosomes

(James Zawada 2006). In this work, 2 x YTPG medium witnesses the enhancement of CFPS as the EGFP yield increased by around 35.3%.

4.2.2 Cell harvesting time

E. coli extract, representing the cytoplasmic content, provides essential components for protein synthesis, including ribosomes, proteins, membrane lipids, and other molecules (Reuel, 2018). Monomers of 70S ribosomes, comprising 50S and 30S subunits, can convert into 100S ribosome dimers, which lack translational activity (Akira Wada, 1995). This transformation occurs when *E. coli* cells transition from the logarithmic phase to the stationary phase, due to the expression of ribosome modulation factor (RMF) encoded by the *rmf* gene under nutritional deficiency conditions, or during slow growth of *E. coli* cells (Masahiro Yamagishi, 1993). Consequently, *E. coli* cells are typically harvested at an OD600 of 3.0, indicative of the mid-logarithmic phase. Interestingly, the extract obtained from *E. coli* A19 cells harvested during the non-growing phase exhibits comparable synthesis efficiency to that harvested during the mid-logarithmic phase (Jurek Failmezger, 2017). In this study, the peak synthesis efficiency of CFPS was observed when the extract was derived from B95.ΔA cells harvested at approximately OD600 2.5, which may be associated with the overexpression of T7 RNA polymerase, as Jared L. Dopp identified that optimizing IPTG induction timing and cell growth duration was a key factor in preparing highly productive extracts from BL21-Gold (DE3) strain (Reuel, 2018).

4.2.3 Cell disruption

The methodologies employed to disrupt the plasma membrane of cells include the French press, sonication, high-pressure homogenizer, and multi-bead shocker. Compared to other methods, the French press is considered a traditional approach for preparing extracts for CFPS. However, it is difficult to disrupt cells uniformly and is not suitable for scaling up the extract preparation

(Takanori Kigawa 2004). In our protocol, a high-pressure homogenizer was utilized to lyse the cells, owing to its advantages such as scalability of cell disruption, reproducibility of the process, and the incorporation of a cooling system that protects the extract from heat-induced damage. During the process, we find that the lower pressure (700 bar/10,152 psi) is necessary for the highly productive extract from B95.ΔA, which differs from the pressure used habitually (15,000 psi). Besides, 500 bar pressure is not suitable for the extract preparation (The experiment was done by my colleague, which is not shown in this work). The reason for the application of lower pressure could be the changes in the cell wall rigidity of B95.ΔA caused by multiple genomic mutations (Eiko Seki 2018).

4.2.4 T7 RNA polymerase

T7 RNA polymerase (T7 RNAP), a small chain of DNA-dependent RNA polymerase from a virulent bacteriophage T7 (Fano 1945), synthesizes RNA from DNA. T7 RNAP is highly specific for the T7 promoter, capable of initiating and extending the RNA chain, and accurately terminating transcription by releasing from the DNA template, without requiring auxiliary factors (Chun Jung Chen 1999). Additionally, the T7 promoter/T7 RNAP system is capable of high-level overexpression of a heterogeneous protein (Richardson 1985). Therefore, the overexpression of T7 RNAP in the extract is a method to improve the productivity of CFPS (Reuel 2018). This is supported by our study, which shows that T7 RNAP overexpression increases the EGFP yield by 58-fold (lysate 5, Table 3.1), compared with the one without T7 RNAP overexpression (lysate 4, Table 3.1).

4.2.5 Mg²⁺ and K⁺ concentration

Mg²⁺ and K⁺ are essential effectors on the protein synthesis efficiency of CFPS, playing indispensable functions in energy regeneration, transcription, and translation in CFPS. In the current CFPS, phosphoenolpyruvate (PEP) and

Discussion

acetyl phosphate (AcP) complete energy regeneration. The phosphate group from PEP/AcP is transferred to ADP, catalyzed by exogenous pyruvate kinase/endogenous acetate kinase, to generate ATP and pyruvate/acetate, respectively. Subsequently, the generated ATP, as the primary energy resource, is consumed to drive the transcriptional and translational process and release inorganic phosphate (Dong-Myung Kim 1999). In this process, Mg^{2+} binds to pyruvate kinase to accelerate the phosphate group transfer, and K^+ arranges its residues to bind the nucleotide ADP (Jesús Oria-Hernández 2005). Similarly, Mg^{2+} and K^+ catalyze the reaction of acetate kinase (Roseman 1986). Transcription is done by T7 RNAP, whose activity is influenced by the concentration of Mg^{2+} and K^+ . T7 RNAP is most active in the presence of 20 mM Mg^{2+} , while its activity decreases in the presence of K^+ and is even lost when the concentration is above 0.2 M (Ring 1973). The translational process is involved in the ribosomes from the B95.ΔA extract, whose association of 50S and 30S subunits to the activatable 70S ribosome is facilitated by 15mM Mg^{2+} (A. Tissières 1959). However, when the concentration of Mg^{2+} is around 10 mM, the 70S ribosome conforms to the inactive 100S ribosome in vitro, and ribosomes break down when the concentration is above 20 mM or less 1 mM (A. Tissières 1959). In like manner, the subunits 50S and 30S of ribosomes degrade and convert to particles with the sedimentation coefficients of 14S and 18S when the absence of K^+ in the *E. coli* B207 strain (Lubin 1965), because K^+ plays an important role in the association of ribosome subunits and stabilization of rRNAs, tRNAs, and r-proteins (Alexey Rozov 2019). However, the ribosome dissociates when the K^+ concentration is above 0.5 mM (Richard S. Zitomer 1972) due to competition with Mg^{2+} at the metal binding sites of ribosomal subunits (Daniel J. Klein 2004). Since the activity of many proteins is affected by Mg^{2+} and K^+ , the current CFPS method indicates that the optimal concentrations of Mg^{2+} and K^+ are 12 mM and 170 mM, respectively. After the addition of 1 μ M BpA-tRNA, 1 μ M BpA-RS, and 1mM BpA, the optimized CFPS could express the full-length PS1-N190X, although the truncated PS1 still exists.

4.3 Photo-crosslinking assay

A variety of factors influence the efficiency of photo-crosslinking, including external elements such as irradiation duration - affected by the variability in UV bulb intensity and the distance between the sample and the UV source (Janhavi A. Kolhe 2023), as well as internal factors such as low yields of UAA incorporation, diminished fidelity of UAA incorporation, poor reactivity of UAA with amino acids of the interacting partner, suboptimal positioning of UAA, internal quenching from neighboring amino acids (Jody K. Lancia 2014), or lipids (Melissa N. Webby 2022), and the correct conformation of membrane proteins (Raghavendar Reddy Sanganna Gari 2021).

To mitigate the influence of external factors, the model utilized was sjGST-F52X. Experiments demonstrate that when the concentration of purified sjGST-F52X reaches 286 ng/μl, the optimal irradiation time exceeds 45 minutes at an illumination distance of approximately 2 cm from the UV source, using the crosslinker CL-1.

Given the low yield of purified PS1-N190X (57.4 ng/μl), the irradiation time for the photo-crosslinking assay between PS1-N190X and PEN2 was set at 90 minutes. However, this photo-crosslinking attempt was unsuccessful under four different hydrophobic environments, which is likely due to conformational changes in the PS1-N190X and PEN2 subcomplexes. The altered configuration is likely caused by PS1 endoproteolysis and the incorporation of subcomplexes into liposomes or nanodiscs.

4.3.1 PS1 endoproteolysis

PS1 undergoes endoproteolytic cleavage between residues M292 and V293 to generate the stable PS1-NTF and PS1-CTF fragments (Anne L. Brunkan 2005), especially when binding to PEN2, which is also enhanced by lipid-raft membrane compartments rather than non-lipid raft membrane compartments (Min Suk Kang 2013). Similarly, when we expressed the PS1-N190X and PEN2

subcomplex in the presence of typical lipid-raft imitation, MSP-2N2-DMPC and Sulfo-DIBMA-DMPC nanodiscs, the PS1-NTF fragment was observed. The presence of the PS1-NTF fragment indicates that the self-activation of PS1 occurred in the PS1-N190X/PEN2/MSP-2N2-DMPC complex, which is different from the previous finding in our lab when Fos14-purified PS1 and PEN2 were integrated into MSP-1D1 nanodiscs with a diameter of 9.5 nm (Tao, 2022). PS1 self-activation leads to the form of heterodimer PS1-NTF/PS1-CTF, homodimer PS1-NTF/PS1-NTF, or PS1-CTF/PS1-CTF, which has been demonstrated in previous studies using the DSP crosslinking assay, co-immunoprecipitation, or the GST pull-down assay (Gang Yu 1998; Gopal Thinakaran 1998; Sebastien S. Hebert 2003)

However, the consequence of PS1 endoproteolysis, the heterodimer PS1-NTF/PS1-CTF, is a high possibility of why there was no crosslinked band between PS1-N190X and PEN2 after illumination. As this work shows, the PS1-N190X-NTF fragment disappeared after irradiation in the presence of MSP-DMPC or Sulfo-DIBMA-DMPC nanodiscs (Figure 3.23 D; Figure 3.25 A), and no crosslinked PEN2 bands were observed (Figure 3.23C; Figure 3.25 B). Therefore, it is highly possible that the photocrosslinking occurs between PS1 fragments.

4.3.2 The incorporation of the PS1-PEN2 subcomplex

The PS1 and PEN2 subcomplexes were inserted into nanodiscs made with DMPC (Figure 3.25), instead of liposomes containing 25% (w/w) brain extract total lipids, 65% (w/w) egg phosphatidylcholine (eggPC), and 10% (w/w) Cholesterol (Figure 3.28). DMPC is a synthetic phospholipid with two 14-carbon saturated acyl chains at the sn-1 and sn-2 positions of glycerophosphocholine, forming a liquid-order lipid domain when tightly packed by MSP or Sulfo-DIBMA. EggPC, on the other hand, is a mixture that includes both saturated fatty acids (such as palmitic and stearic acids) and unsaturated

Discussion

fatty acids (like oleic and linoleic acids), resulting in a more fluid membrane and creating a liquid-disordered lipid domain in the liposomes.

Overall, this suggests that PS1 and PEN2 subcomplexes prefer to embed in a liquid-ordered membrane rather than a liquid-disordered one. As Marilia Barros noted, γ -secretase is highly localized in liquid-ordered lipid domains, which are more ordered and rigid phases within a cell membrane compared to liquid-disordered lipid domains (Marilia Barros 2020).

4.4 Conclusion

To introduce unnatural amino acids into the membrane protein PS1, I tried both in vivo and in vitro protein expression systems. Based on the release factor 1-deficient B95.ΔA strain, the in vitro protein expression system can express 57.4 ng/μl of PS1-N190X in the batch mode. The expression efficiency of CFPS depends on the following conditions. First, the B95.ΔA cells should be harvested before the early-logarithmic phase (OD₆₀₀ ~ 2.5) to achieve ribosome-rich cells. Second, B95.ΔA should be disrupted only one time under lower pressure (700 bar) with Microfluidics M-110P, followed by overnight dialysis at 37°C to produce an efficient S30 lysate. Third, the T7 RNAP should not only be added into the reaction chamber of CFPS, but also overexpressed in B95.ΔA cells to enhance its expression efficiency. Fourth, the concentration of magnesium and potassium should be optimized. Furthermore, the addition of purified BpA-RS and extracted BpA-tRNA is advantageous for the incorporation of unnatural amino acid BpA into PS1.

To photo-crosslink PS1 and PEN2, I first employed sjGST-F52X to explore an appropriate crosslinking method. Second, the potential residues of PS1 were analyzed by the PPM 2.0 Web Server from the OPM database (<https://opm.phar.umich.edu/>), the PyMOL software (<https://pymol.org/edu/>), and the structural analysis software WinCoot. Third, different hydrophobic environments were employed to support the correct conformation of PS1 and PEN2. However, no crosslinked bands were certified by Western blot and mass spectrometry. The possible reasons are PS1 endoproteolysis in the liquid-ordered domain of DMPC-nanodiscs and the poor incorporation efficiency in the liquid-disordered lipid domain of prepared liposomes.

Summary

The in vitro protein expression system (CFPS) serves as a robust methodology for synthesizing proteins in an open and accessible environment. The utilization of Genetic Code Expansion (GCE) through the orthogonal tRNA/tRNA-synthetase system incorporating unnatural amino acids (UAAs) enables the study of protein interactions by photocrosslinking. Peter G. Schultz employed this GCE approach to incorporate *p*-benzoyl-L-phenylalanine and *p*-azido-L-phenylalanine into sjGST, achieving a photocrosslinking efficiency of approximately 30%. This relatively low efficiency is attributed to the competition between release factor 1 and the orthogonal tRNA to the amber codon. The use of a release factor 1 (RF1)-deficient strain leverages the CFPS platform to be an effective tool for the incorporation of UAAs into target proteins. The CFPS derived from the RF1-deficient strain B95.ΔA increased the photocrosslinking efficiency.

To incorporate UAAs into the membrane protein PS1, first of all, the factors relevant for the preparation of S30 lysate were studied, whose optimum conditions are culturing cells in 2xYTPG medium; harvesting T7 RNAP-overexpressed B95.ΔA cells before the early-logarithmic phase (OD₆₀₀ ~ 2.5); disrupting cells only once under lower pressure (700 bar with Microfluidics M-110P); and overnight dialyzing the crude S30 lysate at 37°C. Furthermore, three components of the CFPS in batch mode had to be optimized: The optimal conditions were 85 ng/μl T7 RNAP, 12mM Mg²⁺, and 170 mM K⁺, yielding around 100 ng/μl EGFP. By addition of 1 μM BpA-RS, 1 μM BpA-tRNA, and 1 mM BpA into the CFPS was significant in incorporating UAAs into target membrane proteins. Overall, the established CFPS can yield 57.4 ng/μl PS1-N190X in batch mode.

Subsequently, photocrosslinking was used on the soluble protein sjGST, which showed that the photocrosslinking efficiency was nearly 100% once the irradiation time exceeded 45 minutes and the illumination distance was

Summary

approximately 2 cm between the 96-well plate and the UV source when using the photon source.

Various hydrophobic receivers for solubilization of the membrane protein PS1, such as liposomes, CHAPSO-destabilized liposomes, Sulfo-DIBMA-DMPC nanodiscs, and MSP-2N2-DMPC nanodiscs, were used to mimic the native environment. The right conformation of PS1-N190X and PEN2 was confirmed indirectly by the self-activation of the PS1-N190X/PEN2/MSP-2N2-DMPC nanodiscs complex. After the 90-minute irradiation of self-activated (cleaved) PS1/PEN2 in the presence of MSP-2N2-DMPC nanodiscs in the established CFPS. The full-length PS1-N190X bands were recovered, thereby showing that the NTF- and CTF- fragments of PS1 were crosslinked.

Zusammenfassung

Das In-vitro-Proteinexpressionssystem (CFPS) dient als robuste Methode zur Proteinsynthese in einer offenen und zugänglichen Umgebung. Die Nutzung der genetischen Code-Expansion (GCE) durch das orthogonale tRNA/tRNA-Synthetase-System unter Einbeziehung nichtnatürlicher Aminosäuren (UAAs) ermöglicht die Untersuchung von Proteininteraktionen durch Photovernetzung. Peter G. Schultz nutzte diesen GCE-Ansatz, um p-Benzoyl-L-phenylalanin und p-Azido-L-phenylalanin in sjGST einzubauen und erreichte dabei eine Photovernetzungseffizienz von etwa 30 %. Diese relativ geringe Effizienz wird auf die Konkurrenz zwischen Release Factor 1 und der orthogonalen tRNA um das Amber-Codon zurückgeführt. Die Verwendung eines Release Factor 1 (RF1)-defizienten Stammes macht die CFPS-Plattform zu einem effektiven Werkzeug für den Einbau von UAAs in Zielproteine. Das aus dem RF1-defizienten Stamm B95.ΔA gewonnene CFPS erhöhte die Photovernetzungseffizienz. Um UAAs in das Membranprotein PS1 einzubauen, wurden zunächst die für die Herstellung des S30-Lysats relevanten Faktoren untersucht. Optimale Bedingungen hierfür sind die Kultivierung der Zellen in 2xYTPG-Medium, die Ernte von T7-RNAP-überexprimierten B95.ΔA-Zellen vor der frühen logarithmischen Phase ($OD_{600} \sim 2,5$), das einmalige Aufbrechen der Zellen bei niedrigem Druck (700 bar mit Microfluidics M-110P) und die Dialyse des rohen S30-Lysats über Nacht bei 37 °C. Darüber hinaus mussten drei Komponenten des CFPS im Batch-Modus optimiert werden: Die optimalen Bedingungen waren 85 ng/μl T7-RNAP, 12 mM Mg^{2+} und 170 mM K^{+} , was etwa 100 ng/μl EGFP ergab. Die Zugabe von 1 μM BpA-RS, 1 μM BpA-tRNA und 1 mM BpA zum CFPS war für den Einbau von UAAs in Zielmembranproteine signifikant. Insgesamt kann das etablierte CFPS im Batch-Modus 57,4 ng/μl PS1-N190X liefern. Anschließend wurde das lösliche Protein sjGST photovernetzt. Dabei zeigte sich, dass die Photovernetzungseffizienz nahezu 100 % betrug, sobald die Bestrahlungszeit

Summary

45 Minuten überschritt und der Beleuchtungsabstand zwischen der 96-Well-Platte und der UV-Quelle bei Verwendung der Photonenquelle etwa 2 cm betrug. Verschiedene hydrophobe Empfänger zur Solubilisierung des Membranproteins PS1, wie Liposomen, CHAPSO-destabilisierte Liposomen, Sulfo-DIBMA-DMPC-Nanodiscs und MSP-2N2-DMPC-Nanodiscs, wurden verwendet, um die native Umgebung nachzubilden. Die korrekte Konformation von PS1-N190X und PEN2 wurde indirekt durch die Selbstaktivierung des PS1-N190X/PEN2/MSP-2N2-DMPC-Nanodisc-Komplexes bestätigt. Nach 90-minütiger Bestrahlung von selbstaktiviertem (gespaltenem) PS1/PEN2 in Gegenwart von MSP-2N2-DMPC-Nanodiscs im etablierten CFPS wurden die PS1-N190X-Bänder in voller Länge wiederhergestellt, was zeigt, dass die NTF- und CTF-Fragmente von PS1 vernetzt waren.

Abbreviations

AD	Alzheimer's disease
AMP	Adenosine monophosphate
ApH1	anterior pharynx 1
apo A-I	apolipoprotein A-I
APP	amyloid β -protein precursor
APPC99	The membrane-inserted C99 fragment of amyloid β -protein precursor
ATP	Adenosine 5'- triphosphate
AzF	p-azidophenylalanine
BACE1	β -site APP-cleaving enzyme-1
bp	Base pair
BpA	Benzoylphenylalanine/p-benzoyl-L-phenylalanine
BpA-RS	BpA-tRNA synthetase is responsible for attaching BpA to its corresponding BpA-tRNA
BpA-tRNA	transfer RNA (tRNA) that carries the unnatural amino acid p-benzoyl-L-phenylalanine (BpA)
CFPS	cell-free protein synthesis
CHAPS	3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate
CHAPSO	3-[(3-Cholamidopropyl) dimethylammonio]-2-hydroxy-1-propanesulfonate
CMC	Critical Micelle Concentration
Cryo-EM	cryogenic electron microscopy
CTP	Cytidine 5'- triphosphate
Cy6	6-Cyclohexyl-hexyl- β -D-maltoside
DDM	Dodecyl Maltoside
DDM	n-Dodecyl- β -D-maltoside
DIBMA	diisobutylene/maleic acid co-polymer

Abbreviations

DizPK	Nε-(3-(3-methyl-3H-diazirine-3-yl)-propamino-carbonyl-lysine
DLS	Dynamic light scattering
DM	n-Decyl-β-D-maltopyranoside
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DMPG	1,2-Dimyristoyl-sn-glycero-3-phosphoglycerol
DMSO	Dimethyl sulfoxide
DOPC	1,2-dioleoyl-sn-glycero-3-phosphocholine
DOPE	1,2-dioleoyl-sn-glycero-3-phosphoethanolamine
DOPG	1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
ER	endoplasmic reticulum
FC12	FOS-choline-12
FC14	FOS-choline-14
FC16	FOS-choline-16
fL	femtoliter
GCE	genetic code expansion
GPCR	G protein-coupled receptor
GTP	Guanosine 5'- triphosphate
GUVs	giant unilamellar vesicles
Kb	Kilo base
kDa	Kilodalton
LDAO	N, N-Dimethyldodecylamine-n-oxide
LUVs	uni-lamellar vesicles
mRNA	Messenger ribonucleic acid
MSP	membrane scaffold proteins

Abbreviations

MWCO	Molecular Weight Cut-Off
ncAA	non-canonical amino acids
NCT	nicastatin
NG	n-Nonyl- β -D-glucopyranoside
OD	Optical density
OG	n-Octyl- β -D-glucopyranoside
ORFs	Open reading frames
PAGE	Polyacrylamide Gel Electrophoresis
PC	Phosphatidylcholine
PCR	Polymerase Chain Reaction
PEN2	presenilin enhancer 2
PEP	Phosphoenol pyruvic acid monopotassium salt
photo-Leu	photo-leucine
photo-Lys	photo-lysine
photo-Met	photo-methionine
PK	Pyruvate kinase
PS1	presenilin 1
PS1-CTF	C-terminal fragment of presenilin 1
PS1-NTF	N-terminal fragment of presenilin 1
RF1	Release factor 1
rRNA	Ribosomal ribonucleic acid
sAPP β	APP β fragment
SEC	Size exclusion chromatography
sfGFP	Superfolder green fluorescent protein
sjGST	glutathione S-transferase
SMA	Styrene maleic acid
T7 RNAP	T7 RNA Polymerase

Abbreviations

TfmdF	p-trifluoromethyl-diazirinyL-L-phenylalanine
TMDs	transmembrane domains
tRNA	Transfer ribonucleic acid
UAAs	Unnatural amino acids
UTP	Uridine 5' - triphosphate
α -Syn	α -synuclein

Reference

1. A. D. Bangham, M. M. Standish and J. C. Watkins. 1965. 'Diffusion of univalent ions across the lamellae of swollen phospholipids', *J. Mol. Biol.*, 13: 17.
2. A. Tissières, D. Schlessinger, and Françoise Gros. 1960. 'Amino acid incorporation into proteins by Escherichia Coli ribosomes ', *Proc Natl Acad Sci USA*, 46: 14.
3. A. Tissières, J.D. Watson, D. Schlessinger, B.R. Hollingworth. 1959. 'Ribonucleoprotein Particles from Escherichia coli', *Journal of molecular biology*, 1: 13.
4. Abraham Olusegun Oluwole, Bartholomäus Danielczak, Annette Meister, Jonathan Oyebamiji Babalola, Carolyn Vargas, and Sandro Keller. 2017. 'Solubilization of membrane proteins into functional lipid-bilayer Nanodiscs Using a diisobutylene/maleic acid copolymer', *Angew Chem Int Ed Engl*, 56: 6.
5. Adarshi Welegedara, Luke A. Adams, Bim Graham, Thomas Huber, and Gottfried Otting. 2018. 'Site-specific incorporation of selenocysteine by genetic encoding as a photocaged unnatural amino acid', *Bioconjugate Chemistry* 29: 8.
6. Alexander Deiters, Bernhard H. Geierstanger, and Peter G. Schultz. 2005. 'Site-specific in vivo labeling of proteins for NMR studies', *ChemBioChem*, 6: 4.
7. Alexander S. Spirin, Vladimir I. Baranov, Lbov A. Ryabova, Sergey Yu. Ovodov, Yuly B. Alakhov. 1988. 'A continuous cell-free translation system capable of producing polypeptides in high yield', *Science*, 242: 3.
8. Alexey Rozov, Iskander Khusainov, Kamel El Omari, Ramona Duman, Vitaliy Mykhaylyk, Marat Yusupov, Eric Westhof, Armin Wagner & Gulnara Yusupova. 2019. 'Importance of potassium ions for ribosome structure and function revealed by long-wavelength X-ray diffraction', *nature communications*, 10: 12.
9. Ana C.A. Roque, Claudia S.O. Silva, M. Angela Taipa. 2007. 'Affinity-based methodologies and ligands for antibody purification: Advances and perspectives', *Journal of chromatography A*: 12.
10. Anne L. Brunkan, Maribel Martinez, Emily S. Walker, and Alison M. Goate. 2005. 'Presenilin endoproteolysis is an intramolecular cleavage', *Molecular and Cellular Neuroscience*, 29: 9.
11. Anton A.A. Smith, Henriette E. Autzen, Bryan Faust, Joseph L. Mann, Benjamin W. Muir, Shaun Howard, Almar Postma, Andrew J. Spakowitz, Yifan Cheng, and Eric A. Apper. 2020. 'Lipid Nanodiscs via Ordered Copolymers', *Chem*, 6: 15.
12. Bachman, Julia. 2013. 'Site-directed mutagenesis', *Methods in enzymology*, 529: 8.
13. Barbara Mui, Laurie Chow and Michael J. Hope. 2003. 'extrusion technique to generate liposomes of defined size', *Methods in enzymology*, 367: 12.

References

14. Barrett, G. C. 1985. Chemistry and biochemistry of the amino acids (Chapman and Hall: London New York).
15. Bilgimol C Joseph, Suthakaran Pichaimuthu, Sankaranarayanan Srimeenakshi, Musti Murthy, Kalimuthu Selvakumar, Ganesan M and Sadananda Rao Manjunath. 2015. 'An Overview of the Parameters for Recombinant Protein Expression in Escherichia coli', Journal of cell science & therapy, 6: 7.
16. Bin Shen, Zhang Xiang, Barbara Miller, Gordon Louie, Wen Yuan Wang, Joseph P. Noel, Fred H. Gage, and Lei Wang. 2011. 'Genetically encoding unnatural amino acids in neural stem cells and optically reporting voltage-sensitive domain changes in differentiated neurons', Stem cells, 29: 20.
17. Blaskovich, Mark Arnold Thomas. 2016. 'unusual amino acids in medicinal chemistry', Journal of medicinal chemistry, 59: 30.
18. Boerje Lindqvist, Torsten Storgards. 1995. 'Molecular-sieving properties of starch', nature, 175: 2.
19. Booth, Heather E. Findlay & Paula J. 2017. 'The folding, stability and function of lactose permease differ in their dependence on bilayer lipid composition', Scientific reports, 7: 12.
20. Bruce E. Cohen, Tim B. McAnaney, Eun Sun Park, Yuh Nung Jan, Steven G. Boxer, Lily Yeh Jan. 2002. 'Probing protein electrostatics with a synthetic fluorescent amino acid', Science, 296: 5.
21. Cédric Eichmann, Silvia Campioni, Julia Kowal, Innokentiy Maslennikov, Juan Gerez, Xiaoxia Liu, Joeri Verasdonck, Nadezhda Nespovitaya, Senyon Choe, Beat H. Meier, Paola Picotti, Josep Rizo, Henning Stahlberg, and Roland Riek. 2016. 'Preparation and characterization of stable α -synuclein lipoprotein particles', The Journal of Biological Chemistry, 291: 12.
22. Chiara De Faveri, Jordan M. Mattheisen, Thomas P. Sakmar, and Irene Coin. 2024. 'Noncanonical amino acid tools and their application to membrane protein studies', Chemical reviews, 124: 53.
23. Chloe M. Jones, D. Miklos Robkis, Robert J. Blizzard, Mika Munari, Yarra Venkatesh, Tiberiu S. Mihaila, Alex J. Eddins, Ryan A. Mehl, William N. Zagotta, Sharona E. Gordon and E. James Petersson. 2021. 'Genetic encoding of a highly photostable, long lifetime fluorescent amino acid for imaging in mammalian cells', Chemical science, 12: 11.
24. Chong, Shaorong. 2014. 'Overview of cell-free protein synthesis: historic landmarks, commercial systems, and expanding applications', Current protocols in molecular biology, 108: 14.
25. Christian Klammt, Daniel Schwarz, Klaus Fendler, Winfried Haase, Volker Dötsch and Frank Bernhard. 2005. 'Evaluation of detergents for the soluble expression of α -helical and β -barrel-type integral membrane proteins by a preparative scale individual cell-free expression system', The FEBS journal, 272: 6024.

References

26. Christopher J. Noren, Spencer J. Anthony-Cahill, Michael C. Griffith, Peter G. Schultz. 1989. 'A general method for site-specific incorporation of unnatural amino acids into proteins', *Science*, 244: 7.
27. Chun Jung Chen, Zhi-Jie Liu, John P. Rose and Bi-Cheng Wang. 1999. 'Low-salt crystallization of T7 RNA polymerase: a first step towards the transcription bubble complex', *Acta Crystallographica. Section D, Biological Crystallography*, 55: 5.
28. Coin, Irene. 2018. 'Application of non-canonical crosslinking amino acids to study protein–protein interactions in live cells', *Current Opinion in Chemical Biology*, 46: 8.
29. Daniel J. Klein, Peter B. Moore, and Thomas A. Steitz 2004. 'The contribution of metal ions to the structural stability of the large ribosomal subunit', *RNA*, 10: 15.
30. David Garenne, Matthew C. Haines, Eugenia F. Romantseva, Paul Freemont, Elizabeth A. Strychalski and Vincent Noireaux. 2021. 'Cell- free gene expression', *Nature Reviews Methods Primers*, 1: 18.
31. David Glueck, Anne Grethen, Manabendra Das, Ogochukwu Patricia Mmekaka, Eugenio Pérez Patallo, Annette Meister, Ritu Rajender, Stefan Kins, Markus Räsche, Julian Victor, Ci Chu, Manuel Etzkorn, Zoe Köck, Frank Bernhard, Jonathan Oyebamiji Babalola, Carolyn Vargas, and Sandro Keller. 2022. 'Electroneutral polymer nanodiscs enable interference-free probing of membrane proteins in a lipid-bilayer environment', *Small*, 18: 16.
32. Dmitry E. Agafonov, Yiwei Huang, Michael Grote, Mathias Sprinzl. 2005. 'Efficient suppression of the amber codon in E. coli in vitro translation system', *FEBS Letters*, 579: 5.
33. Dong-Myung Kim, James R. Swartz. 1999. 'Prolonging cell-free protein synthesis with a novel ATP regeneration system', *Biotechnology and Bioengineering*, 66: 10.
34. E. Hochuli, H. Dobeli and A. Schacher. 1987. 'New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues', *Journal of chromatography A*, 411: 8.
35. Edmund R.S. Kunji, Marilyn Harding, P. Jonathan G. Butler, Pearl Akamine. 2008. 'Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography', *Methods*, 46: 11.
36. Eiko Seki, Tatsuo Yanagisawa, and Shigeyuki Yokoyama. 2018. 'Cell-free protein synthesis for multiple site-specific incorporation of noncanonical amino acids using cell extracts from RF-1 deletion E.coli strains', *Methods in Molecular Biology*, 1728: 17.
37. Eric M. Tippmann, Wenshe Liu, Daniel Summerer, Antha V. Mack, and Peter G. Schultz. 2007. 'A Genetically Encoded Diazirine Photocrosslinker in Escherichia coli', *ChemBioChem*, 8: 5.
38. F. Olson, C.A. Hunt, F.C. Szoka, W.J. Vall and D. Papahadjopoulos. 1979. 'Preparation of liposomes of defined size distribution by extrusion through polycarbonate membranes', *Biochimica et Biophysica Acta*, 557: 15.

References

39. Fano, M. Demerec and U. 1945. 'Bacteriophage-resistant mutants in *Escherichia Coli*', *Genetics*, 30: 19.
40. Federico Katzen, Todd C. Peterson and Wieslaw Kudlicki. 2009. 'Membrane protein expression: no cells required', *Trends in Biotechnology*, 27: 6.
41. Gang Yu, Fusheng Chen, Georges Levesque, Masaki Nishimura, Dong-Mei Zhang, Lyne Levesque, Ekatarina Rgaeva, Donhong Xu, Yan Liang, Monika Duthie, Peter H. St George-Hyslop, and Paul E. Fraser. 1998. 'The presenilin 1 protein is a component of a high molecular weight intracellular complex that contains Beta-catenin', *The Journal of Biological Chemistry*, 273: 6.
42. Ge Yang, Kun Yu, Jan Kubicek, Joerg Labahn. 2018. 'Expression, purification, and preliminary characterization of human presenilin-2', *Process Biochemistry*, 64: 11.
43. George Umanah, Li-Yin Huang, Peter G. Schultz, Fred Naider and Jeffrey M. Becker. 2009. 'Incorporation of the Unnatural Amino Acid p-benzoyl-L-phenylalanine (Bpa) into a G Protein-coupled Receptor in its Native Context', *Advances in Experimental Medicine and Biology* 611: 3.
44. Gesteland, Raymond F. 1966. 'Isolation and characterization of ribonuclease I mutants of *Escherichia coli*', *Journal of molecular biology*, 16: 18.
45. Glenn D. Prestwich, György Dormán, John T. Elliott, Dale M. Marecak and Anu Chaudhary. 1997. 'Benzophenone Photoprobes for Phosphoinositides, Peptides and Drugs', *Photochemistry and Photobiology*, 65: 13.
46. Glenn F. Short, III, Serguei Y. Golovine, and Sidney M. Hecht. 1999. 'Effects of release factor 1 on in vitro protein translation and the elaboration of proteins containing unnatural amino acids', *Biochemistry*, 38: 12.
47. Gopal Thinakaran, Jean B. Regard, Christopher M. L. Bouton, Christie L. Harris, Donald L. Price, David R. Borchelt, and Sangram S. Sisodia. 1998. 'Stable association of presenilin derivatives and absence of presenilin interactions with APP', *Neurobiology of Disease*, 4: 16.
48. Grossmann, Kenneth D. Bloch and Barbara. 1995. 'Digestion of DNA with restriction endonucleases', *Current protocols in molecular biology*: 21.
49. Guy Cathala, Jean-Francois Savouret, Bernardita Mendez, Brian L. West, Michael Karin, Joseph A. Martial, John D. Baxter 1983. 'A method for isolation of intact, translationally active ribonucleic acid', *DNA*, 2: 7.
50. Gyorgy Dorman, Glenn D. Prestwich. 1994. 'Benzophenone photophores in biochemistry', *Biochemistry*, 33: 13.
51. Hamdi Nsairat, Dima Khater, Usama Sayed, Fadwa Odeh, Abeer Al Bawab, Walhan Alshaer. 2022. 'Liposomes: structure, composition, types, and clinical applications', *Heliyon*, 8: 15.
52. Harwood, Adrian I. 1994. *Protocols for gene analysis* (Humana Press).
- Howard G. Barth, Barry E. Boyes, and Christian Jackson. 1994. 'Size Exclusion Chromatography', *Analytical Chemistry*, 66: 26.

References

53. Huey-Lang Yang, Lionel Ivashkiv, Hui-Zhu Chen, Geoffrey Zubay, and Michael Cashel. 1980. 'Cell-free coupled transcription-translation system for investigation of linear DNA segments', *Proc Natl Acad Sci USA*, 77: 5.
54. Irem Avcilar-Kucukgoze, Howard Gamper, Ya-Ming Hou, and Anna Kashina. 2020. 'Purification and Use of tRNA for Enzymatic Post-translational Addition of Amino Acids to Proteins', *STAR Protocols*: 16.
55. J. D. Bain, Charles G. Glabe, Thomas A. Dix, and A. Richard Chamberlin. 1989. 'Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide', *American Chemical Society*, 111: 2.
56. J. D. Bain, Christopher Switzer, A. Richard Chamberlin and Steven A. Benner. 1992. 'Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code', *nature*, 356: 3.
57. Jackson, Hugh R. B. Pelham and Richard J. 1976. 'An efficient mRNA-dependent translation system from reticulocyte lysates', *Eur J Biochem*, 67: 10.
58. James C. Kauer, Susan Erickson-Viitanen, Henry R. Wolfe, Jr., and William F. DeGrado. 1986. 'p-Benzoyl-L-phenylalanine, A new photoreactive amino acid', *The Journal of Biological Chemistry*, 261: 6.
59. James Zawada, James Swartz. 2006. 'Effects of growth rate on cell extract performance in Cell-Free Protein Synthesis', *Biotechnology and Bioengineering*, 94: 7.
60. Janhavi A. Kolhe, Neethu L. Babu, and Brain C. Freeman. 2023. 'Protocol for establishing a protein interactome based on close physical proximity to a target protein within live budding yeast ', *STAR Protocols*, 4: 16.
61. Jason W. Chin, Andrew B. Martin, David S. King, Lei Wang, and Peter G. Schultz. 2002a. 'Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*', *PNAS*, 99 5.
62. Jason W. Chin, Stephen W. Santoro, Andrew B. Martin, David S. King, Lei Wang, and Peter G. Schultz. 2002b. 'Addition of p-Azido-L-phenylalanine to the genetic code of *Escherichia coli*', *American Chemical Society*, 124: 2.
63. Jason W. Chin, T. Ashton Cropp, J. Christopher Anderson, Mridul Mukherji, Zhiwen Zhang, Peter G. Schultz. 2003. 'An expanded eukaryotic genetic code', *Science*, 301: 5.
64. Jesús Oria-Hernández, Nallely Cabrera, Ruy Pérez-Montfort, and Leticia Ramírez-Silva. 2005. 'Pyruvate Kinase Revisited the activating effect of K⁺', *The Journal of Biological Chemistry*, 280: 6.
65. Ji Luo, Rajendra Uprety, Yuta Naro, Chungjung Chou, Duy P. Nguyen, Jason W. Chin, and Alexander Deiters. 2014. 'Genetically encoded optochemical probes for simultaneous fluorescence reporting and light activation of protein function with two-photon excitation', *Journal of the american chemical society*, 136: 8.
66. Jiro Adachi, Kazushige Katsura, Eiko Seki, Chie Takemoto, Mikako Shirouzu, Takaho Terada, Takahito Mukai, Kensaku Sakamoto, and Shigeyuki Yokoyama. 2019. 'Cell-free protein synthesis using S30 extracts from

References

- Escherichia coli RFzero strains for efficient incorporation of non-natural amino acids', *Int J Mol Sci*, 20: 12.
67. John W. Littlefield, Elizabeth B. Keller, Jerome Gross, and Paul C. Zamecnik. 1955. 'Studies on cytoplasmic ribonucleoprotein particles from the liver of the rat', *J Biol Chem*, 271: 14.
68. Josep Cladera, Jean-Louis Rigaud, Joaquim Villaverde and Mireia Dunach. 1997. 'Liposome solubilization and membrane protein reconstitution using Chaps and Chapso', *Eur. J. Biochem*, 243: 7.
69. Kairat Madin, Tatsuya Sawasaki, Tomio Ogasawara, and Yaeta Endo. 2002. 'A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: Plants apparently contain a suicide system directed at ribosomes', *Proc Natl Acad Sci USA*, 97: 6.
70. Kamaruzaman, Gloria Yi Wei Tseu and Khairul Azfar. 2023. 'A review of different types of liposomes and their advancements as a form of gene therapy treatment for breast cancer', *Molecules*, 28: 31.
71. Karin V. Loscha, Anthony J. Herlt, Ruhui Qi, Thomas Huber, Kiyoshi Ozawa, and Gottfried Otting. 2012. 'Multiple-site labeling of proteins with unnatural amino acids', *Angew Chem Int Ed Engl*, 51: 4.
72. Kim, Jong Woo Lee and Hyounghman. 1992. 'Fragmentation of dimyristoylphosphatidylcholine vesicles by apomyoglobin', *Archives of Biochemistry and Biophysics*, 297: 8.
73. Kiselev, Domenico Lombardo and Mikhail A. 2022. 'Methods of liposomes preparation: formation and control factors of versatile nanocarriers for biomedical and nanomedicine application', *Pharmaceutics*, 14: 49.
74. L. A. Ryabova, S. A. Ortlepp and V. I. Baranov. 1989. 'Preparative synthesis of globin in a continuous cell-free translation system from rabbit reticulocytes', *Nucleic acids research*, 17: 1.
75. LaBaer, Shane Miersch and Joshua. 2011. 'Nucleic acid programmable protein arrays: versatile tools for array-based functional protein studies', *Curr Protoc Protein Sci*, 27: 26.
76. Labrou, Nikolaos E. 2014. *Protein Downstream Processing* (Humana Press). Lance E. Steward, Cynthia S. Collins, Marcella A. Gilmore, Justin E. Carlson, J. B. Alexander Ross, and A. Richard Chamberlin. 1997. 'In vitro site-specific incorporation of fluorescent probes into β -galactosidase', *Journal of the american chemical society*, 119: 6.
77. Lei Wang, Ansgar Brock, Brad Herbrich, Peter G. Schultz. 2001. 'Expanding the genetic code of Escherichia coli', *Science*, 292: 4.
78. Lipmann, J. Mager and Fritz. 1958. 'Amino acid incorporation and the reversion of its initial phase with cell-free tetrahymena preparations', *Biochemistry*, 44: 5.
79. Loretta Eggenreich, Carolyn Vargas, Cenek Kolar and Sandro Keller. 2023. 'Lipid exchange among electroneutral Sulfo-DIBMA nanodiscs is independent of ion concentration', *Biological chemistry*, 404: 11.

References

80. Lubin, Herbert L. Ennis and Martin. 1965. 'Pre-ribosomal particles formed in potassium-depleted cells: Studies on degradation and stabilization', *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*, 95.
81. Lyman, Ilya Levental and Ed. 2023. 'Regulation of membrane protein structure and function by their lipid nano-environment', *Nature Reviews Molecular Cell Biology*, 24: 16.
82. Marilia Barros, William J Houlihan, Chelsea J Paresi, Matthew Brendel, Kevin D Ryneerson, Chang-wook Lee, Olga Prikhodko, Cristina Cregger, Geoffrey Chang, Steven L Wagner, M Lane Gilchrist, Yue-Ming Li. 2020. ' γ -Secretase partitioning into lipid bilayers remodels membrane microdomains after direct insertion', *Langmuir*, 36: 22.
83. Mark A. McLean, Ilia G. Denisov, Yelena V. Grinkova, Stephen G. Sligar. 2020. 'Dark, Ultra-Dark and Ultra-Bright Nanodiscs for membrane protein investigations', *Anal Biochem*, 607: 22.
84. Mark R. Fleissner, Eric M. Brustad, Tama'sKa'lai, Christian Altenbach, Duilio Cascio, Francis B. Peters, Ka'Ima'n Hideg, Sebastian Peuker, Peter G. Schultz, and Wayne L. Hubbell. 2009. 'Site-directed spin labeling of a genetically encoded unnatural amino acid', *PNAS*, 106: 6.
85. Martin Mehnert, Thomas Sommer, and Ernst Jarosch. 2013. 'Der1 promotes movement of misfolded proteins through the endoplasmic reticulum membran', *Nature cell biology*, 16: 30.
86. Michele A. McTigue, DeWight R. Williams and John A. Tainer. 1995. 'Crystal Structures of a Schistosomal Drug and Vaccine Target: Glutathione S-Transferase from *Schistosoma japonica* and its Complex with the Leading Antischistosomal Drug Praziquantel', *J. Mol. Biol.*, 246: 7.
87. Min Suk Kang, Seung-Hoon Baek, Yoon Sun Chun, A. Zenobia Moore, Natalie Landman, Diego Berman, Hyun Ok Yang, Maho Morishima-Kawashima, Satoko Osawa, Satoru Funamoto, Yasuo Ihara, Gilbert Di Paolo, Jeong Hill Park, Sungkwon Chung, and Tae-Wan Kim. 2013. 'Modulation of lipid kinase PI4KII α activity and lipid raft association of presenilin 1 underlies γ -secretase inhibition by ginsenoside (20S)-Rg3', *The Journal of Biological Chemistry*, 288: 15.
88. Monika Suchanek, Anna Radzikowska & Christoph Thiele. 2005. 'Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells', *Nature methods* 2: 7.
89. MT Rahman, MS Uddin, R Sultana, A Moue, M Setu. 2013. 'Polymerase chain reaction (PCR): a short review', *Anwer Khan Modern Medical College Journal*, 4: 7.
90. Naoto Watanabe, Taisuke Tomita, Chihiro Sato, Toshio Kitamura, Yuichi Morohashi, and Takeshi Iwatsubo. 2005. 'Pen-2 is incorporated into the γ -secretase complex through binding to transmembrane domain 4 of presenilin 1', *The Journal of Biological Chemistry*, 280: 9.
91. Nicola J. Harris, Eamonn Reading, Kenichi Ataka, Lucjan Grzegorzewski, Kalypso Charalambous, Xia Liu, Ramona Schlesinger, Joachim Heberle &

References

- Paula J. Booth. 2017. 'Structure formation during translocon-unassisted cotranslational membrane protein folding', *Scientific reports*, 7: 15.
92. Nicola J. Harris, Grant A. Pellowe, Laura R. Blackholly, Samuel Gulaidi-Breen, Heather E. Findlay and Paula J. Booth. 2022. 'Methods to study folding of alpha-helical membrane proteins in lipids', *Open Biology*, 12: 14.
93. Nirenberg, J. Heinrich Matthaei and Marshall W. 1961. 'Characteristics and stabilization of DNAase-sensitive protein synthesis in *E. coli* extracts', *Proc Natl Acad Sci USA*, 47: 9.
94. Nobumasa Takasugi, Taisuke Tomita, Ikuo Hayashi, Makiko Tsuruoka, Manabu Niimura, Yasuko Takahashi, Gopal Thinakaran & Takeshi Iwatsubo. 2003. 'The role of presenilin cofactors in the γ -secretase complex', *nature*, 422: 4.
95. Oliver Holmes, Swetha Paturi, Dennis J. Selkoe, and Michael S. Wolfe. 2014. 'PEN2 is essential for γ -secretase complex stability and trafficking but partially dispensable for endoproteolysis', *Biochemistry*, 53: 14.
96. Parichehre Davanloo, Alan H. Rosenberg, John J. Dunn, and F. William Studier. 1984. 'Cloning and expression of the gene for bacteriophage T7 RNA polymerase', *Proc. Natl. Acad. Sci.USA*, 81: 5.
97. Patrick C. Fraering, Matthew J. LaVoie, Wenjuan Ye, Beth L. Ostaszewski, W. Taylor Kimberly, Dennis J. Selkoe, and Michael S. Wolfe. 2004. 'Detergent-dependent dissociation of active γ -Secretase reveals an interaction between Pen-2 and PS1-NTF and offers a model for subunit organization within the complex', *Biochemistry*, 43: 11.
98. R. E. Galardy, L. C. Craig, and M. P. Printz 1973. 'Benzophenone triplet: a new photochemical probe of biological ligand-receptor interactions', *Nature New Biology*, 242: 2.
99. Ralf-Bernhardt Rues, Fang Dong, Volker Dötsch, Frank Bernhard. 2018. 'Systematic optimization of cell-free synthesized human endothelin B receptor folding', *Methods*, 147: 11.
100. Reuel, Jared L. Dopp and Nigel F. 2018. 'Process optimization for scalable *E. coli* extract preparation for cell-free protein synthesis', *Biochemical Engineering Journal*, 138: 8.
101. Richard S. Zitomer, Joel G. Flaks. 1972. 'Magnesium Dependence and Equilibrium of the *Escherichia coli* Ribosomal Subunit Association', *Journal of molecular biology*, 71: 17.
102. Richardson, Stanley Tabor and Charles C. 1985. 'A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes', *PNAS*, 82: 5.
103. Ring, Michael Chamberlin and Janet. 1973. 'Characterization of T7-specific ribonucleic acid polymerase ', *The Journal of Biological Chemistry*, 248: 10.
104. Roman Levin, Zoe Köck, Janosch Martin, René Zangl, Theresa Gewering, Leah Schüler, Arne Moeller, Volker Dötsch, Nina Morgner, Frank Bernhard.

References

2022. 'Cotranslational assembly of membrane protein/nanoparticles in cell-free systems', *Biochim Biophys Acta Biomembr*, 1864: 16.
105. Rosanne Wouters, Christine Michiels, Ragna Sannerud, Bertrand Kleizen, Katleen Dillen, Wendy Vermeire, Abril Escamilla Ayala, David Demedts, Randy Schekman, and Wim Annaert. 2021. 'Assembly of γ -secretase occurs through stable dimers after exit from the endoplasmic reticulum', *The Journal of Cell Biology*, 220: 26.
106. Roseman, Donna K. Fox and Saul. 1986. 'Isolation and characterization of homogeneous acetate kinase from *Salmonella typhimurium* and *Escherichia coli*', *The Journal of Biological Chemistry*, 261: 11.
107. S.R. Tonge, B.J. Tighe. 2001. 'Responsive hydrophobically associating polymers: a review of structure and properties', *Advanced Drug Delivery Reviews*, 53: 14.
108. Samuel B. Weiss, George Acs, and Fritz Lipmann. 1958. 'Amino acid incorporation in pigeon pancreas fractions', *Proc Natl Acad Sci USA*, 44: 8.
109. Sarah E. Walker, Jon Lorsch. 2013. 'RNA Purification - Precipitation Methods', *Methods Enzymol*, 530: 7.
110. Satoshi Mikami, Mamiko Masutani, Nahum Sonenberg, Shigeyuki Yokoyama, Hiroaki Imataka 2006. 'An efficient mammalian cell-free translation system supplemented with translation factors', *Protein Expression and Purification*, 46: 10.
111. Schultz, Han Xiao and Peter G. 2016. *At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code* (cold spring harbor laboratory press).
112. Sebastien S. Hebert, Chantal Godin, Georges Levesque. 2003. 'Oligomerization of human presenilin-1 fragments', *FEBS Letters*, 550: 5.
113. Shinsuke Sando, Atsushi Ogawa, Teruyuki Nishi, Masayoshi Hayami and Yasuhiro Aoyama. 2007. 'In vitro selection of RNA aptamer against *Escherichia coli* release factor 1', *Bioorganic & Medicinal Chemistry Letters*, 17: 5.
114. Shixian Lin, Dan He, Teng Long, Shuai Zhang, Rong Meng, and Peng R. Chen. 2014. 'Genetically Encoded Cleavable Protein Photo-Cross-Linker', *Journal of the american chemical society*, 136: 4.
115. Shixin Ye, Ekaterina Zaitseva, Gianluigi Caltabiano, Gebhard F. X. Schertler, Thomas P. Sakmar, Xavier Deupi & Reiner Vogel. 2010. 'Tracking G-protein-coupled receptor activation using genetically encoded infrared probes', *Nature*, 464: 5.
116. Shixin Ye, Morgane Riou, Stéphanie Carvalho, and Pierre Paoletti. 2013. 'Expanding the Genetic Code in *Xenopus laevis* Oocytes', *ChemBioChem*, 14: 6.
117. Shixin Ye, Thomas Huber, Reiner Vogel, and Thomas P Sakmar. 2009. 'FTIR analysis of GPCR activation using azido probes', *Nature chemical biology*, 5: 9.
118. Shujia Zhua, Morgane Riou, C. Andrea Yao, Stéphanie Carvalho, Pamela C. Rodriguez, Olivier Bensaude, Pierre Paoletti and Shixin Ye. 2014.

References

- 'Genetically encoding a light switch in an ionotropic glutamate receptor reveals subunit-specific interfaces', *PNAS*, 111: 6.
119. Stefan Scheidelaar, Martijn C. Koorengevel, Cornelius A. van Walree, Juan J. Dominguez, Jonas M. Dörr and J. Antoinette Killian. 2016. 'Effect of Polymer Composition and pH on Membrane Solubilization by Styrene-Maleic Acid Copolymers', *Biophysical Journal*, 111: 13.
120. Susan E. Cellitti, David H. Jones, Leanna Lagpacan, Xueshi Hao, Qiong Zhang, Huiyong Hu, Scott M. Brittain, Achim Brinker, Jeremy Caldwell, Badry Bursulaya, Glen Spraggon, Ansgar Brock, Youngha Ryu, Tetsuo Uno, Peter G. Schultz, and Bernhard H. Geierstanger. 2008. 'In vivo incorporation of unnatural amino acids to probe structure, dynamics and ligand binding in a large protein by Nuclear Magnetic Resonance spectroscopy', *Journal of the american chemical society*, 130: 27.
121. Szabolcs Fekete, Alain Beck, Jean-Luc Veuthey, Davy Guillarme. 2014. 'Theory and practice of size exclusion chromatography for the analysis of protein aggregates', *Journal of Pharmaceutical and Biomedical Analysis*, 101: 13.
122. T. Dawn Parks, Kerstin K. Leuther, Eric D. Howard, Stephen A. Johnston, and William G. Dougherty. 1994. 'Release of Proteins and Peptides from Fusion Proteins using a Recombinant Plant Virus Proteinase', *Analytical Biochemistry*, 216: 5.
123. Tae-Wan Kim, Dong-Myung Kim, Cha-Yong Choi 2006. 'Rapid production of milligram quantities of proteins in a batch cell-free protein synthesis system', *Journal of Biotechnology*, 124: 373.
124. Takanori Kigawa, Takashi Yabuki, Natsuko Matsuda, Takayoshi Matsuda, Rie Nakajima, Akiko Tanaka & Shigeyuki Yokoyama. 2004. 'Preparation of Escherichia coli cell extract for highly productive cell-free protein expression', *Journal of Structural and Functional Genomics*, 5: 6.
125. Tangpo Yang, Xiao-Meng Li, Xiucong Bao, Yi Man Eva Fung & Xiang David Li. 2016. 'Photo-lysine captures proteins that bind lysine post-translational modifications', *Nature Chemical Biology*, 12: 6.
126. Tao, Chengcheng. 2022. 'The presenilin-1-PEN-2 complex, the minimal catalytic subunit of the γ -secretase': 186.
127. Timothy H. Bayburt, Yelena V. Grinkova, and Stephen G. Sligar. 2002. 'Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins', *Nano Letters*, 2: 4.
128. Timothy J. Knowles, Rachael Finka, Corinne Smith, Yu-Pin Lin, Tim Dafforn, and Michael Overduin. 2009. 'Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer', *Journal of the American Chemical Society*, 131: 2.
129. Wieslaw Kudlicki, Gisela Kramer, and Boyd Hardesty. 1992. 'High Efficiency Cell-Free Synthesis of Proteins: Refinement of the Coupled Transcription/ Translation System', *Analytical Biochemistry*, 206: 5.

References

130. Yasmin Aydin, Irene Coin. 2023a. 'Genetically encoded crosslinkers to address protein–protein interactions', *Protein Science*, 32: 16.
131. Yasmin Aydin, Thore Böttke, Jordy Homing Lam, Stefan Ernicke, Anna Fortmann, Maik Tretbar, Barbara Zarzycka, Vsevolod V. Gurevich, Vsevolod Katritch & Irene Coin. 2023b. 'Structural details of a Class B GPCR-arrestin complex revealed by genetically encoded crosslinkers in living cells', *Nature Communications*, 14: 13.
132. Yoshihiro Shimizu, Akio Inoue, Yukihide Tomari, Tsutomu Suzuki, Takashi Yokogawa, Kazuya Nishikawa, and Takuya Ueda. 2001. 'Cell-free translation reconstituted with purified components', *Nature Biotechnology*, 19: 5.
133. Yoshihiro Shimizu, Takashi Kanamori, Takuya Ueda. 2005. 'Protein synthesis by pure translation systems', *Methods*, 36: 6.
134. Yu, Kun. 2016. 'Characterization of the subunits of the r-secretase complex'.
135. Yutetsu Kuruma, Takuya Ueda. 2015. 'The PURE system for the cell-free synthesis of membrane proteins', *Nature Protocols* 13:28–44, 10: 17.
136. Zachary A. Manzer, Ekaterina Selivanovitch, Alexis R. Ostwalt, and Susan Daniel. 2023. 'Membrane protein synthesis: no cells required', *Trends in Biochemical Sciences*, 48: 13.
137. Zheng Xiang, Haiyan Ren, Ying S Hu, Irene Coin, Jing Wei, Hu Cang & Lei Wang. 2013. 'Adding an unnatural covalent bond to proteins through proximity-enhanced bioreactivity ', *Nature Methods*, 10: 5.
138. Zubay, Huizhu Chen, and Geoffrey. 1983. 'Prokaryotic coupled transcription-translation', *Methods Enzymol*, 101: 17.

Appendix I: DNA and protein sequences

>Presenilin 1_WT_nucleotides

ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCG
 CGCGGCAGCCATATGACCGAACTGCCTGCACCGCTGAGCTATTTTCAGA
 ATGCACAGATGAGCGAAGATAACCATCTGAGCAATACCGTTCGTAGCCA
 GAATGATAATCGTGAACGTCAAGAACACAATGATCGTCGTAGCCTGGGT
 CATCCGGAACCGCTGAGTAATGGTCGTCCGCAGGGTAATAGCCGTCAG
 GTTGTTGAACAGGATGAAGAGGAAGATGAAGAACTGACCCTGAAATATG
 GTGCCAAACATGTGATTATGCTGTTTGTTCGGTTACCCTGTGTATGGTT
 GTTGTTGTGGCAACCATTAAAAGCGTGAGCTTTTATACCCGTAAAGATGG
 CCAGCTGATTTATACCCCGTTTACCGAAGATACCGAAACCGTTGGTCAG
 CGTGCACTGCATAGTATTCTGAATGCAGCAATTATGATTAGCGTGATTGT
 GGTGATGACCATTCTGCTGGTTGTTCTGTATAAATACCGCTGCTATAAAG
 TGATTCATGCCTGGCTGATTATTAGCAGCCTGCTGCTGCTGTTTTTCTTC
 AGCTTTATCTATCTGGGCGAAGTGTTCAAACCTATAATGTTGCCGTTGA
 TTATATCACCGTTGCACTGCTGATTTGGAATTTTGGTGTTGTTGGCATGA
 TTAGCATCCATTGGAAAGGTCCGCTGCGTCTGCAGCAGGCATATCTGAT
 TATGATTTTCAGCACTGATGGCCCTGGTGTTTCATCAAATATCTGCCGGAAT
 GGACCGCATGGCTGATTCTGGCAGTTATTAGCGTTTATGATCTGGTTGCA
 GTTCTGTGTCCGAAAGGCCCTCTGCGTATGCTGGTTGAAACCGCACAAAG
 AACGTAATGAAACCCTGTTTCCGGCACTGATTTATTCAAGCACCATGGTT
 TGGCTGGTTAATATGGCAGAAGGTGATCCGGAAGCACAGCGTCGTGTTA
 GCAAAAATAGCAAATACAATGCAGAAAGCACCGAACGTGAAAGCCAGGA
 TACCGTTGCAGAAAATGATGATGGTGGTTTTAGCGAAGAATGGGAAGCC
 CAGCGTGATAGCCATCTGGGTCCGCATCGTAGCACACCGGAAAGCCGT
 GCAGCAGTTCAAGAACTGAGCAGCTCAATCCTGGCAGGCGAAGATCCTG
 AAGAACGTGGTGTTAACTGGGTCTGGGTGATTTTATCTTTTATAGCGTT
 CTGGTTGGTAAAGCAAGCGCAACCGCAAGCGGTGATTGGAATACCACCA
 TTGCATGTTTTGTTGCCATTCTGATTGGTCTGTGTCTGACATTACTGCTGC
 TGGCCATTTTCAAAAAGCACTGCCTGCCCTGCCGATTAGCATTACCTTT
 GGTCTGGTTTTTTTACTTCGCAACCGATTATCTGGTTTCAGCCGTTTATGGA
 TCAACTGGCATTTCACCAGTTTTTACATCTAA

>Presenilin 1_WT_Protein

MGSSHHHHHHSSGLVPRGSHMTELPAPLSYFQNAQMSEDNHLSENTRVSQ
 NDNRRERQEHNDRRSLGHPEPLSNGRPQGNSRQVVEQDEEEDEELTLKYG
 AKHVIMLFVPVTLCMVVV/VATIKSVSFYTRKDGLIYTPFTEDTETVGQRALH
 SILNAAIMISVIVMTILLVVLKYRCYKVIHAWLISSLLLLFFFSFIYLGVEFKT
 YNVAVDYITVALLIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL
 EWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMVWL
 VNMAEGDPEAQR RVSKNSKYNAESTERESQDTVAENDDGGFSEWEAQR
 DSHLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKA

SATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVFYFATDYL
VQPFMDQLAFHQFYI

>Presenilin 1_N190X_nucleotides

ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCG
CGCGGCAGCCATATGACCGAACTGCCTGCACCGCTGAGCTATTTTCAGA
ATGCACAGATGAGCGAAGATAACCATCTGAGCAATACCGTTCGTAGCCA
GAATGATAATCGTGAACGTCAAGAACACAATGATCGTCGTAGCCTGGGT
CATCCGGAACCGCTGAGTAATGGTCGTCCGCAGGGTAATAGCCGTCAG
GTTGTTGAACAGGATGAAGAGGAAGATGAAGAACTGACCCTGAAATATG
GTGCCAAACATGTGATTATGCTGTTTGTTCGGTTACCCTGTGTATGGTT
GTTGTTGTGGCAACCATTAAAAGCGTGAGCTTTTATACCCGTAAAGATGG
CCAGCTGATTTATACCCCGTTTACCGAAGATACCGAAACCGTTGGTCAG
CGTGCACTGCATAGTATTCTGAATGCAGCAATTATGATTAGCGTGATTGT
GGTGATGACCATTCTGCTGGTTGTTCTGTATAAATACCGCTGCTATAAAG
TGATTCATGCCTGGCTGATTATTAGCAGCCTGCTGCTGCTGTTTTTCTTC
AGCTTTATCTATCTGGGCGAAGTGTTCAAACCTATATGGTTGCCGTTGA
TTATATCACCGTTGCACTGCTGATTTGGAATTTTGGTGTTGTTGGCATGA
TTAGCATCCATTGGAAAGGTCCGCTGCGTCTGCAGCAGGCATATCTGAT
TATGATTTTCAGCACTGATGGCCCTGGTGTTTCATCAAATATCTGCCGGAAT
GGACCGCATGGCTGATTCTGGCAGTTATTAGCGTTTATGATCTGGTTGCA
GTTCTGTGTCCGAAAGGCCCTCTGCGTATGCTGGTTGAAACCGCACAAAG
AACGTAATGAAACCCTGTTTCCGGCACTGATTTATTCAAGCACCATGGTT
TGGCTGGTTAATATGGCAGAAGGTGATCCGGAAGCACAGCGTCGTGTTA
GCAAAAATAGCAAATACAATGCAGAAAGCACCGAACGTGAAAGCCAGGA
TACCGTTGCAGAAAATGATGATGGTGGTTTTAGCGAAGAATGGGAAGCC
CAGCGTGATAGCCATCTGGGTCCGCATCGTAGCACACCGGAAAGCCGT
GCAGCAGTTCAAGAACTGAGCAGCTCAATCCTGGCAGGCGAAGATCCTG
AAGAACGTGGTGTTAACTGGGTCTGGGTGATTTTATCTTTTATAGCGTT
CTGGTTGGTAAAGCAAGCGCAACCGCAAGCGGTGATTGGAATACCACCA
TTGCATGTTTTGTTGCCATTCTGATTGGTCTGTGTCTGACATTACTGCTGC
TGGCCATTTTCAAAAAGCACTGCCTGCCCTGCCGATTAGCATTACCTTT
GGTCTGGTTTTTTTACTTCGCAACCGATTATCTGGTTTCAGCCGTTTATGGA
TCAACTGGCATTTCACCAGTTTTACATCTAA

>Presenilin 1_N190X_Protein

MGSSHHHHHHSSGLVPRGSHMTELPAPLSYFQNAQMSEDNHLSTVRSQ
NDNRERQEHNDRSLGHPEPLSNGRPQGNSRQVVEQDEEEDEELTLKYG
AKHVIMLFVPVTLCMVVV/VATIKSVSFYTRKDGLIYTPFTEDTETVGQRALH
SILNAAIMISVIVMTILLVVLKYRCYKVIHAWLISSLLLLFFFSFIYLGVEFKT
YBpAVAVDYITVALLIWNFGVVGMIHGWKGPLRLQQAYLIMISALMALVFIKY
LPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMV
WLVNMAEGDPEAQRVSKNSKYNAESTERESQDTVAENDDGGFSEWEA
QRDShLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVLVG

KASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKALPALPISITFGLVIFYFATD
YLVQPFMDQLAFHQFYI

>Flag-PEN2_nucleotides

ATGGACTACAAAGACGATGACGACAAGGCAATTGAAGGTCGTAATCTGG
AACGTGTTAGCAACGAAGAAAACTGAATCTGTGCCGCAAATATTACCTG
GGTGGTTTTGCATTTCTGCCGTTTCTGTGGCTGGTTAACATCTTTTGGTTT
TTTCGTGAAGCATTCTGGTTCCGGCATATACCGAACAGAGCCAGATTAA
AGGTTATGTTTGGCGTAGCGCAGTTGGTTTTCTGTTTTGGGTATTGTTC
TGACCAGCTGGATTACCATCTTTCAGATTTATCGTCCGCGTTGGGGTGCA
CTGGGTGATTATCTGAGCTTTACCATTCCGCTGGGCACCCCGTAA

>Flag-PEN2_Protein

MDYKDDDDKAIEGRNLERSNEEKLNLCKRYYLGGFAFLPFLWLNVNIFWFF
REAFLVPAYTEQSQIKGYVWRSavgflfwvvltswitifqiyrprwgalgd
YLSFTIPLGTP

>NCT-rho_nucleotides

ATGAATAGCGTTGAACGCAAAATCTATATTCCGCTGAATAAAACCGCACC
GTGTGTTCTGCTGCTGAATGCAACCCATCAGATTGGTTGTCAGAGCAGC
ATTAGCGGTGATACCGGTGTTATTCATGTTGTGGAAAAAGAAGAGGATCT
GCAGTGGGTTCTGACCGATGGTCCGAATCCGCCTTATATGGTTCTGCTG
GAAAGCAAACATTTTACCCGTGATCTGATGGAAAACTGAAAGGTCGTAC
CAGCCGTATTGCAGGTCTGGCAGTTAGCCTGACCAAACCGAGTCCGGCA
AGCGGTTTTAGCCCGAGCGTTCAGTGTCCGAATGATGGTTTTGGTGTTTA
TAGCAATAGCTACGGTCCGGAATTTGCACATTGTCGTGAAATTCAGTGGA
ATAGCCTGGGTAATGGTCTGGCCTATGAAGATTTTAGCTTTCCGATTTTC
CTGCTGGAAGATGAGAATGAAACCAAAGTGATCAAACAGTGCTATCAGG
ATCATAATCTGAGCCAGAATGGTAGCGCACCGACCTTTCCGCTGTGTGC
AATGCAGCTGTTTAGCCACATGCATGCAGTTATTAGCACCGCAACCTGTA
TGCGTCGTAGCAGCATTGAGAGCACCTTTAGCATTAAATCCGGAATTGTT
TGTGATCCGCTGAGCGATTATAATGTTTGGAGCATGCTGAAACCGATTAA
TACCACCGGCACCCTGAAACCGGATGATCGTGTTGTTGTTGCAGCAACC
CGTCTGGATAGCCGTAGCTTTTTTTTGAATGTTGCACCGGGTGCAGAAA
GCGCAGTTGCAAGCTTTGTTACCCAGCTGGCAGCAGCAGAAGCACTGCA
AAAAGCACCGGATGTTACCACCCTGCCTCGTAATGTGATGTTTGTTTTTT
TTCAGGGCGAAACCTTCGATTATATTGGTAGCAGCCGTATGGTGTACGAT
ATGGAAAAAGGTAAATTTCCGGTGCAGCTGGAAAATGTTGATAGCTTTGT
TGAAGTGGGTCAGGTTGCACTGCGTACCAGTCTGGAAGTGTGGATGCAT
ACCGATCCGGTTAGCCAGAAAAATGAAAGCGTTCGTAATCAGGTTGAAG
ATCTGCTGGCAACCCTGGAAAAAGCGGTGCGGGTGTTCGGGCAGTTAT
TCTGCGTCGTCCGAATCAGAGCCAGCCGCTGCCTCCGAGCAGCCTGCA
GCGTTTTCTGCGTGACGTAATATTAGTGGTGTGTTCTGGCAGATCATA
GCGGTGCATTTACAATAAATACTACCAGAGCATCTATGACACCGCAGAA

AATATCAATGTTAGCTATCCGGAATGGCTGAGTCCGGAAGAAGATCTGAA
TTTTGTTACCGATACCGCAAAAGCACTGGCAGATGTTGCAACCGTTCTGG
GTCGTGCACTGTATGAACTGGCAGGCGGTACAAATTTTAGCGATACCGT
TCAGGCAGATCCGCAGACCGTTACCCGTCTGCTGTATGGTTTTCTGATTA
AAGCAAATAACAGCTGGTTCCAGAGCATTCTGCGCCAGGATCTGCGTAG
CTATCTGGGTGATGGTCCGCTGCAGCACTATATTGCAGTTAGCAGCCCCG
ACCAATACCACCTATGTTGTTCAGTATGCACTGGCAAATCTGACCGGGCAC
CGTTGTTAATCTGACCCGTGAACAGTGTCAGGATCCGAGCAAAGTTCCG
AGCGAAAATAAAGATCTGTATGAGTATAGCTGGGTTCAGGGTCCTCTGC
ATAGCAATGAAACGGATCGTCTGCCTCGTTGTGTTCTAGTACCGCACG
TCTGGCACGTGCGCTGTCACCGGCATTTGAACTGAGCCAGTGGTCAAGC
ACCGAATATAGCACCTGGACCGAAAGCCGTTGGAAAGATATTCGTGCCC
GTATTTTTCTGATCGCAAGCAAAGAACTGGAAGTATTACCTGACCGTG
GGTTTTGGTATTCTGATTTTTAGCCTGATTGTGACCTATTGCATTAACGCA
AAAGCCGATGTTCTGTTTATTGCACCGCGTGAACCGGGTGCCGTTAGCT
ATGGCTCCTCCGGCACCGAGACTTCCCAGGTGGCGCCAGCTTGATGA

>NCT-rho_Protein

MNSVERKIYIPLNKTAPCVRLNATHQIGCQSSISGDTGVIHVVEKEEDLQWV
LTDGPNPPYMVLLLESKHFTRDLMELKLGRTSRIAGLAVSLTKPSPASGFSPS
VQCPNDGFGVYSNSYGPEFAHCREIQWNSLGNGLAYEDFSFPIFLEDENE
TKVIKQCYQDHNLSQNGSAPTFLCAMQLFSMHAVISTATCMRRSSIQSTF
SINPEIVCDPLSDYNVWSMLKPINTTGTLKPD DRV VVAATRLDSRSFFWNVA
PGAESAVASFVTQLAAAEALQKAPDVTTLP RNVMFVFFQGETFDYIGSSRM
VYDMEKGKFPVQLENVDSFVELGQVALRTSLELWMHTDPVSQKNESVRNQ
VEDLLATLEKSGAGVPAVILRRPNQSQPLPPSSLQRFLRARNISGVVLADHS
GAFHNKYYYQSIYDTAENINVSYPEWLSPEEDLN FVTD TAKALADVATVLGRA
LYELAGGTNFSDTVQADPQTVTRLLYGFLIKANNSWFQSILRQDLRSYLG DG
PLQHYIAVSSPTNTTYVVQYALANLTGT VVNL TREQCQDPSKVPSENKDL YE
YSWVQGPLHSNETDRLPRCVRSTARLARALSPAFELSQWSSTEYSTWTES
RWKDIRARIFLIASKELELITLVGFGILIFSLIVTYCINAKADVLF IAPREPGAVS
YGSSGTETSQVAPA

>ApH1-HA_nucleotides

ATGGGTGCAGCAGTTTTTTTTGGTTGTACCTTTGTTGCATTTGGTCCGGC
ATTTGCACTGTTTCTGATTACCGTTGCCGGTGATCCGCTGCGTGTTATTA
TTCTGGTTGCGGGTGCATTTTTTGGCTGGTTAGCCTGCTGCTGGCAAG
CGTTGTTTGGTTTATCCTGGTTCATGTTACCGATCGTAGTGATGCACGTC
TGCAGTATGGTCTGCTGATTTTTGGTGCAGCCGTTAGCGTTCTGCTGCAA
GAAGTTTTTCGTTTCGCATATTACAAACTGCTGAAAAAAGCAGATGAAGG
TCTGGCAAGCCTGAGCGAAGATGGTCGTAGCCCGATTAGCATTCTGTCAG
ATGGCCTATGTTAGCGGTCTGAGCTTTGGTATTATTAGCGGTGTTTTTAG
CGTGATTAACATTCTGGCAGATGCACTGGGTCCGGGTGTTGTTGGTATT
CATGGTGATAGCCCGTATTATTTCTGACCAGCGCATTTCTGACCGCAG

CAATTATCCTGCTGCATACCTTTTGGGGTGTTGTGTTTTTTGATGCATGTG
AACGTCGTCGTTATTGGGCACTGGGCCTGGTTGTGGGTAGCCATCTGCT
GACCAGTGGTCTGACCTTTCTGAATCCGTGGTATGAAGCAAGTCTGCTG
CCGATTTATGCAGTTACCGTTAGCATGGGTCTGTGGGCATTTATTACCGC
AGGCGGTAGCCTGCGTAGCATTAGCGTAGTCTGCTGTGTCGTCGTCAA
GAGGATAGCCGTGTTATGGTTTATAGCGCACTGCGTATTCCGCCTGAAG
ATGGCTCCTCCGGCTACCCATACGATGTTCCAGATTACGCTTAATAG

>ApH1-HA_Protein

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVILVAGAFFWLVSLLLASVWW
FILVHVTDRSDARLQYGLLIFGAAVSVLLQEVFRFAYYKLLKKADEGLASLSE
DGRSPISIRQMAYVSGLSFGIISGVFSVINILADALGPGVVGIHGDSPYYFLTS
AFLTAAIILLHTFWGVVFFDACERRRYWALGLVVGSHLLTSGLTFLNPWYEA
SLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCRRQEDSRVMVYSALRIPPED
GSSGYPYDVPDYA

>His-BpA-RS_nucleotides

ATGCACCACCACCACCACCACGAGAATTTGTATTTTCAGGGTGACGAATT
TGAAATGATAAAGAGAAACACATCTGAAATTATCAGCGAGGAAGAGTTAA
GAGAGGTTTTAAAAAAGATGAAAAATCTGCTGGTATAGGTTTTGAACCA
AGTGGTAAAATACATTTAGGGCATTATCTCCAAATAAAAAAGATGATTGAT
TTACAAAATGCTGGATTTGATATAATTATATTGTTGGCTGATTACACGCC
TATTTAAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATAGGAGATTA
TAACAAAAAAGTTTTTTGAAGCAATGGGGTTAAAGGCAAAATATCTTTATG
GAAGTCCTTTCCAGCTTGATAAGGATTATACACTGAATGTCTATAGATTG
GCTTTAAAAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACCTTATAGC
AAGAGAGGATGAAAATCCAAAGGTTGCTGAAGTTATCTATCCAATAATGC
AGGTTAATACGAGTCATTATCTGGGCGTTGATGTTGCAGTTGGAGGGAT
GGAGCAGAGAAAAATACACATGTTAGCAAGGGAGCTTTTACCAAAAAAG
GTTGTTTGTATTACAAACCCTGTCTTAACGGGTTTGGATGGAGAAGGAAA
GATGAGTTCTTCAAAGGGAATTTTATAGCTGTTGATGACTCTCCAGAAG
AGATTAGGGCTAAGATAAAGAAAGCATACTGCCAGCTGGAGTTGTTGA
AGGAAATCCAATAATGGAGATAGCTAAATACTTCCTTGAATATCCTTTAAC
CATAAAAAGGCCAGAAAAATTTGGTGGAGATTTGACAGTTAATAGCTATG
AGGAGTTAGAGAGTTTATTTAAAAATAAGGAATTGCATCCAATGGATTTAA
AAAATGCTGTAGCTGAAGAAGCTTATAAAGATTTTAGAGCCAATTAGAAAG
AGATTATAA

>His-BpA-RS_Protein

MHHHHHHENLYFQGDEFEMIKRNTSEIISEEELREVLKKDEKSAGIGFEP
SG KIH LGHYLQIKKMIDLQNAGFDIILLADLHAYLNQKGELDEIRKIGDYNKKVFE
AMGLKAKYLYGSPFQLDKDYTLNVYRLALKTTTLKRARRSMELIAREDENPKV
AEVIYPIMQVNTSHYLGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTG
LDGEGKMSSSKGNFIAVDDSP EEIRAKIKKAYCPAGVVEGNPIMEIAKYFLEY

Appendix

PLTIKRPEKFGGDLTVNSYEELESLFKNKELHPMDLKNVAEELIKILEPIRKRL

>His-AzF-RS_nucleotides

ATGCACCACCACCACCACGAGAATTTGTATTTTCAGGGTATGGACG
AGTTCGAAATGATTAAACGCAACACCAGCGAAATTATCTCTGAAGAAGAG
CTGCGCGAGGTGCTGAAGAAAGACGAGAAGAGCGCGACTATTGGCTTT
GAGCCGTCCGGTAAAATTCACCTGGGTCACTACCTGCAAATCAAGAAGA
TGATTGATCTGCAAAACGCTGGTTTTGACATCATTATCCTGCTGGCGGAC
CTGCACGCCTACCTGAATCAAAAGGGCGAGCTGGATGAGATTCGCAAGA
TCGGCGACTACAATAAGAAAGTCTTCGAAGCCATGGGTTTGAAGGCTAA
ATACGTCTACGGTAGCAATTTTCAGCTGGATAAGGATTACACGTTGAATG
TGTACCGTCTGGCGCTGAAAACCACGCTGAAACGCGCCCGTCTGTTCCAT
GGAGCTGATTGCGCGCGAGGATGAGAATCCAAAAGTTGCTGAGGTTATT
TACCCTATTATGCAAGTTAATCCGTTGCACTACCAGGGTGTTGATGTTGC
CGTCGGTGGTATGGAGCAACGCAAAATTCACATGCTGGCACGTGAACTG
CTGCCGAAAAAGGTTGTCTGTATTCATAATCCGGTCCTGACCGGCCTGG
ATGGCGAGGGTAAAATGAGCAGCAGCAAGGGTAACTTTATTGCAGTTGA
CGATAGCCCGGAAGAAATCCGTGCGAAGATCAAGAAAGCGTACTGCCC
GGCAGGCGTGTTGAGGGTAACCCGATCATGGAAATCGCCAAGTATTTT
CTGGAATACCCACTGACGATTAAGCGCCCGGAGAAATTTGGCGGCGACC
TGACCGTCAACAGCTACGAGGAGCTGGAAAGCTTGTTTAAGAACAAAGA
ACTGCATCCGATGCGCCTGAAAAACGCCGTGGCGGAAGAGCTGATTAAG
ATTCTGGAACCAATTCGCAAACGTCTGTAA

>His-AzF-RS_Protein

MHHHHHHENLYFQGMDEFEMIKRNTSEIIEEELREVLKKDEKSATIGFEP
SKIHLGHYLQIKKMIDLQNAGFDIILLADLHAYLNQKGELDEIRKIGDYNKKV
EAMGLKAKYVYGSNFQLDKDYTLNVYRLALKTTTLKRARRSMELIAREDENPK
VAEVIYPIMQVNPLHYQGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLT
GLDGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGNPIMEIAKYFLE
YPLTIKRPEKFGGDLTVNSYEELESLFKNKELHPMRLKNVAEELIKILEPIRK
RL

>BpA-tRNA_nucleotides

TGCACGGCTAACTAAGCGGCCTGCTGACTTTCTCGCCGATCAAAAGGCA
TTTTGCTATTAAGGGATTGACGAGGGCGTATCTGCGCAGTAAGATGCGC
CCCGCATTCCGGCGGTAGTTCAGCAGGGCAGAACGGCGGACTCTAAAT
CCGCATGGCAGGGGTTCAAATCCCCTCCGCCGGACCAAATTCGAAAAGC
CTGCTCAACGAGCAGGCTTTTTTGCATG

>T7 RNAP_nucleotides

ATGAACACGATTAACATCGCTAAGAACGACTTCTCTGACATCGAACTGGC
TGCTATCCCGTTCAACACTCTGGCTGACCATTACGGTGAGCGTTTAGCTC

Appendix

GCGAACAGTTGGCCCTTGAGCATGAGTCTTACGAGATGGGTGAAGCACG
CTTCCGCAAGATGTTTGAGCGTCAACTTAAAGCTGGTGAGGTTGCGGAT
AACGCTGCCGCCAAGCCTCTCATCACTACCCTACTCCCTAAGATGATTG
CACGCATCAACGACTGGTTTGAGGAAGTGAAAGCTAAGCGCGGCAAGC
GCCCCGACAGCCTTCCAGTTCCTGCAAGAAATCAAGCCGGAAGCCGTAGC
GTACATCACCATTAAGACCACTCTGGCTTGCCTAACCAGTGCTGACAATA
CAACCGTTCAAGGCTGTAGCAAGCGCAATCGGTGCGGCCATTGAGGACG
AGGCTCGCTTCGGTCGTATCCGTGACCTTGAAGCTAAGCACTTCAAGAA
AAACGTTGAGGAACAACCTCAACAAGCGCGTAGGGCACGTCTACAAGAAA
GCATTTATGCAAGTTGTGCGAGGCTGACATGCTCTCTAAGGGTCTACTCG
GTGGCGAGGCGTGGTCTTCGTGGCATAAGGAAGACTCTATTCATGTAGG
AGTACGCTGCATCGAGATGCTCATTGAGTCAACCGGAATGGTTAGCTTA
CACCGCCAAAATGCTGGCGTAGTAGGTCAAGACTCTGAGACTATCGAAC
TCGCACCTGAATACGCTGAGGCTATCGCAACCCGTGCAGGTGCGCTGG
CTGGCATCTCTCCGATGTTCCAACCTTGCGTAGTTCCTCCTAAGCCGTG
GACTGGCATTACTGGTGGTGGCTATTGGGCTAACGGTCGTCGTCCTCTG
GCGCTGGTGCCTACTCACAGTAAGAAAGCACTGATGCGCTACGAAGACG
TTTACATGCCTGAGGTGTACAAAGCGATTAACATTGCGCAAAACACCGCA
TGGAATCAACAAGAAAGTCCTAGCGGTGCGCAACGTAATCACCAAGT
GGAAGCATTGTCCGGTCGAGGACATCCCTGCGATTGAGCGTGAAGAACT
CCCGATGAAACCGGAAGACATCGACATGAATCCTGAGGCTCTACCGCG
TGGAACGTGCTGCCGCTGCTGTGTACCGCAAGGACAAGGCTCGCAAG
TCTCGCCGTATCAGCCTTGAGTTCATGCTTGAGCAAGCCAATAAGTTTGC
TAACCATAAGGCCATCTGGTTCCCTTACAACATGGACTGGCGCGGTCTG
GTTTACGCTGTGTCAATGTTCAACCCGCAAGGTAACGATATGACCAAAG
GACTGCTTACGCTGGCGAAAGGTAAACCAATCGGTAAGGAAGGTTACTA
CTGGCTGAAAATCCACGGTGCAAACCTGTGCGGGTGTGCGATAAGGTTCCG
TTCCCTGAGCGCATCAAGTTCATTGAGGAAAACACGAGAACATCATGG
CTTGCGCTAAGTCTCCACTGGAGAACACTTGGTGGGCTGAGCAAGATTG
TCCGTTCTGCTTCCCTTGCCTTCTGCTTTGAGTACGCTGGGGTACAGCAC
CACGGCCTGAGCTATAACTGCTCCCTTCCGCTGGCGTTTGACGGGTCTT
GCTCTGGCATCCAGCACTTCTCCGCGATGCTCCGAGATGAGGTAGGTG
GTCGCGCGGTTAACTTGCTTCCCTAGTGAAACCGTTCAGGACATCTACGG
GATTGTTGCTAAGAAAGTCAACGAGATTCTACAAGCAGACGCAATCAATG
GGACCGATAACGAAGTAGTTACCGTGACCGATGAGAACACTGGTGAAAT
CTCTGAGAAAGTCAAGCTGGGCACTAAGGCACTGGCTGGTCAATGGCTG
GCTTACGGTGTTACTCGCAGTGTGACTAAGCGTTCAGTCATGACGCTGG
CTTACGGGTCCAAAGAGTTCGGCTTCCGTCAACAAGTGCTGGAAGATAC
CATTACGCCAGCTATTGATTCCGGCAAGGGTCTGATGTTCACTCAGCCG
AATCAGGCTGCTGGATACATGGCTAAGCTGATTTGGGAATCTGTGAGCG
TGACGGTGGTAGCTGCGGTTGAAGCAATGAACTGGCTTAAGTCTGCTGC
TAAGCTGCTGGCTGCTGAGGTCAAAGATAAGAAGACTGGAGAGATTCTT
CGCAAGCGTTGCGCTGTGCATTGGGTAACCTCCTGATGGTTTCCCTGTGT
GGCAGGAATACAAGAAGCCTATTACAGACGCGCTTGAACCTGATGTTCCCT

Appendix

CGGTCAGTTCCGCTTACAGCCTACCATTAAACACCAACAAAGATAGCGAG
ATTGATGCACACAAACAGGAGTCTGGTATCGCTCCTAACTTTGTACACAG
CCAAGACGGTAGCCACCTTCGTAAGACTGTAGTGTGGGCACACGAGAAG
TACGGAATCGAATCTTTTGCAGTATTACGACTCCTTCGGTACCATTCC
GGCTGACGCTGCGAACCTGTTCAAAGCAGTGCGCGAAACTATGGTTGAC
ACATATGAGTCTTGTGATGTACTGGCTGATTTCTACGACCAGTTCGCTGA
CCAGTTGCACGAGTCTCAATTGGACAAAATGCCAGCACTTCCGGCTAAA
GGTAACTTGAACCTCCGTGACATCTTAGAGTCGGACTTCGCGTTCGCGT
AA

>T7 RNAP_Protein

MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRK
MFERQLKAGEVADNAAAKPLITLLPKMIARINDWFEEVKAKRGKRPTAFQF
LQEIKPEAVAYITIKTTLACLT SADNTTVQAVASAIGRAIEDEARFGRIRDLEAK
HFKKNVEEQNLNKRVGHVYKKAFMQVVEADMLSKGLLGGEAWSSWHKEDSI
HVGVRCEIEMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRAGALA
GISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVY
MPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKPE
DIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIFW
PYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYWCLKIHGAN
CAGVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEY
AGVQHHGLSYNCSLPLAFDGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQ
DIYGIVAKKVNEILQADAINGTDNEVVTVTDENTGEISEKVKLGTKALAGQWL
AYGVTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQ
AAGYMAKLIWESVSVTVVAAVEAMNWLKSAKLLAAEVKDKKTGEILRKRC
AVHWVTPDGFVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQ
ESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKA
VRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILES
DFAFA

>TEV_nucleotides

GGAGAAAGCTTGTTTAAGGGGCGCGTGATTACAACCCGATATCGAGCA
CCATTTGTCATTTGACGAATGAATCTGATGGGCACACACATCGTTGTAT
GGTATTGGATTTGGTCCCTTCATCATTACAAACAAGCACTTGTTTAGAAG
AAATAATGGAACACTGTTGGTCCAATCACTACATGGTGTATTCAAGGTCA
AGAACACCACGACTTTGCAACAACACCTCATTGATGGGAGGGACATGAT
AATTATTCGCATGCCTAAGGATTTCCCACCATTTCTCAAAAGCTGAAATT
TAGAGAGCCACAAAGGGAAGAGCGCATATGTCTTGTGACAACCAACTTC
CAAATAAGAGCATGTCTAGCATGGTGTGACAGACTAGTTGCACATTCCC
TTCATCTGATGGCATATTCTGGAAGCATTGGATTCAAACCAAGGATGGGC
AGTGTGGCAGTCCATTAGTATCAACTAGAGATGGGTTTATTGTTGGTATA
CACTCAGCATCGAATTTACCAACACAAACAATTATTTACAAGCGTGCC
GAAAAACTTCATGGAATTGTTGACAAATCAGGAGGCGCAGCAGTGGGTT
AGTGGTTGGCGATTAAATGCTGACTCAGTATTGTGGGGGGGCCATAAAG

Appendix

TTTTCATGGTGAAACCTGAAGAGCCTTTTCAGCCAGTTAAGGAAGCGACT
CAACTCATGAATCGTCGTCGCCGTCGCTAA

>TEV_Protein

GESLFKGPRDYNPISSSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFRRNNG
TLLVQSLHGVFKVKNTTTLQQHLIDGRDMIIRMPKDFPPFPQKLKFREPQRE
ERICLVTTNFQTKSMSSMVSDTSCTFPSSDGIFWKHWIQTGDGQCGSPLVS
TRDGFIVGIHSASNFTNTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADS
VLWGGHKVFMVKPEEPFQPVKEATQLMN

Appendix II: Primers

Table II.1 Primers used for cloning

Primer Name	Primer Sequence
Flag-PEN2-F	AAACATATGGACTACAAAGACGATGACGACAAGGCAATTGAAGGTCGTAATC TG
Flag-PEN2-R	CCGCTCGAGTTACGGGGTGCCCAG
NCT-rho-F	AAACATATGAATCTGGAACGTGTTAGCAACG
NCT-rho-R	AAACTCGAGTCATCAAGCTGGCGCCACCTGGGAAGTCTCGGTGCCGGAGGA GCCATAGCTAACGGCACCCG
ApH1-HA-F	AAACATATGGGTGCAGCAGTTTTTTTTTG
ApH1-HA-R	AAACTCGAGCTATTAAGCGTAATCTGGAACATCGTATGGGTAGCCGGAGGA GCCATCTTCAGGCGGAATACGC
BpA-RS-F	GAGAATTTGTATTTTCAGGGTGACGAATTTGAAATGATAAAGAGAAACACAT
BpA-RS-R	ACCCTGAAAATACAAATTCTCGTGGTGGTGGTGGTGGTGCATATGTATATC
AzF-RS-F	CACCACCACACCACCACGAGAATTTGTATTTTCAGGGTATGGACGAGTTCGA AATGATTAAACGCAACACC
AzF-RS-R	CTCGAGTTACAGACGTTTGCGAATTGGTTCCAGAATCTTAATC
PS1-N190X-F	CAAAACCTATTAGGTTGCCGTTGATTATATCACCG
PS1-N190X-R	CAACCTAATAGGTTTTGAACACTTCGCCC

Acknowledgements

After finishing my master's degree, I started to think about studying abroad for a PhD degree. In three years of preparation, I had the chance to apply for the scholarship from the China Scholarship Council (CSC) with the help of my supervisor, Prof. Dr. Jörg Labahn, my former colleagues, Dr. Ge Yang, Dr. Chengcheng Tao, and my previous employer, TsingkeBiotechnologyCo. Luckily, I received the scholarship from CSC; therefore, I started the journey to study in Germany.

To my supervisor, Prof. Dr. Jörg Labahn, I deeply appreciate that you offered me the chance to study at CSSB in the first one and a half years and at Forschungszentrum Jülich in the last 3 years, where I had opportunities to attend symposiums, share advanced techniques with colleagues in the routine scientific meeting, and learn new things from others; you had enough patience with my project to answer all my questions no matter they were scientific or from daily life; you guided me in the right track to finish my project; you trusted I can complete my project.

To my friend, Dr. Taras Balandin, I sincerely thank you for investing lots of time in discussing different issues that I had in my experiments, providing the device Crosslinker CL-1 to help me conduct the photocrosslinking assay, and giving me an introduction on how to use ÄKTA.

To my colleagues, Dr. Ge Yang, Dr ChengchengTao, Dr. Nishika Sabharwal, Dr. Aziz Tumeh, and Dr. Liang Ma, thanks for helping me in my scientific and daily life.

To my family, I appreciate your unconditional love and support.

Erklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und ohne unerlaubte Hilfe verfasst habe. Die vorliegende Arbeit wurde weder in der jetzigen oder in ähnlicher Form bei einer anderen Institution eingereicht. Es wurden zuvor keine Promotionsversuche unternommen.

Juelich, 07.10.2025

Place, Date

Yajing Xiao

Name