THE UNIVERSITY
of EDINBURGH

Light-activatable apoptosis via genetic code expansion as an *in-vivo* single-cell ablation tool in *Caenorhabditis elegans*

Zhiyan Xi

THE UNIVERSITY *of* EDINBURGH

Doctor of Philosophy in in Biomedical Sciences (Integrative Physiology)

The University of Edinburgh

2021

Abstract

Natural proteins are biopolymers built from a limited variety of canonical amino acids that are encoded by corresponding triplet codons. Genetic code expansion via amber suppression enables me to install (incorporate) in a target protein a “designer” amino acid beyond canonical amino acids, at the site of a pre-introduced amber stop codon. This is through expression in the host cell of an orthogonal pair, consisting of an aminoacyl-tRNA synthetase and an amber-suppressing tRNA ($tRNA_{CUA}$) evolved for the non-canonical amino acids (NCAA). Site-specific incorporation of NCAA endows target proteins with new properties, enabling protein measurement and/or manipulation in ways that are otherwise impossible.

Photo-caged cysteine (PCCys) as a useful NCAA has not been used in any animal before. By using a PCCRS/$tRNA^{Pyl}_{CUA}$ pair evolved from a PylRS/$tRNA^{Pyl}_{CUA}$ pair, I introduced (PCCys) into protein synthesis of multicellular model species *Caenorhabditis elegans* (*C. elegans*). I demonstrated this incorporation of PCCys by expressing a fluorescent reporter either throughout the nematode or in two different neuronal classes.

I used site-specific PCCys incorporation to develop a light-activatable caspase for precisely ablating cells (especially neurons) in living worms. Cell

ablation has been widely adopted in studies on *C. elegans* cell lineage and cell functions. Common ablation methods include high-powered laser ablation, genetic ablation and optogenetic ablation. However, they are unable to ablate single cells in fully developed worms. Caspase is a core executor of apoptosis of both *C. elegans* and human cells. I designed and engineered a photo-caged caspase from human Caspase-3 by replacing its catalytic cysteine with PCCys. 365-nm UVA illumination removes the caging group of PCCys in the caged caspase, thereby activating the caspase to induce apoptosis of the cell(s) targeted. I succeeded in using global UV illumination to activate respective apoptosis events of oxygen-sensing neurons, touch receptor neurons and muscular cells in adult worms. Also, I demonstrate that individual adult neurons can be selectively targeted and efficiently killed with the use of a microscope-mounted 365-nm laser. With this better spatiotemporal control than other ablation methods, our approach is likely to facilitate future *C. elegans* studies with unprecedented specificity and precision.

Lay summary

Cells are basic units of life. To study a cell's role in a living body, one way is to kill the cell and then to see how the living body changes. By applying such cell killings in a kind of transparent worm called *Caenorhabditis elegans*, knowledge on many shared mechanisms of cell activities among most animals (humans included) have been acquired. For example, how a group of cells cooperate in an animal's nervous system to generate behaviour. However, currently available tools are unable to cleanly kill individual cells in the already developed nervous system of an adult worm thus their usages are insufficient to reveal division of a behavioural task among adult nerve cells. Therefore, I planned to develop a better way of cell killing by specifically inducing a cell's suicide program, on the basis of a technique of controlling a protein of interest by light in a living worm.

As executors of various cell activities, numerous kinds of proteins are assembled from a chain or chains of basic units called amino acids by synthesis machinery. During the programmed suicide of a cell, a specialised type of protein-cutting protein named caspase cuts a linkage within specific sequences of amino acids, so to break down vital proteins to cell maintenance. Therefore, switching on caspases would help me kill a specific cell target. By modifying the synthesis machinery of worm cells, I installed into caspase a

man-made amino acid that contains a bulky part (caging group) being removable by UV light and let them be produced only in certain groups of worm cells. When worms were protected from UV, the caging group prevented caspases from cutting other proteins and the cells were alive. After I shone UV to specific cells of worms, the caging group within the caspase in these cells would drop, allowing the caspase to execute the killing program. I did experiments on three different types of cells and all succeeded. Therefore, this method is promising for killing any cell of interest in the worms.

Signed declaration

I am conscious of and understand the University's policy on plagiarism and I declare that this thesis is written on my own with due references provided on all supporting resources, as an original report of my research. I declare that almost all the experimental work contained herein is my own, with collaborative contributions explicitly acknowledged in the text, and that analyses and interpretations of the data are conducted solely by myself. I certify that this work has not been submitted for any other degree or professional qualification.

Zhiyan Xi

Deanery of Biomedical Sciences,

College of Medicine and Veterinary Medicine

October 2021

Acknowledgements

First of all, I am very grateful to my supervisor, Sebastian Greiss. When I was in the first month of my master program, I attended a presentation by Sebastian about genetically encoding non-canonical amino acids, which was the first time I met him and learned about the concept. The excitement I felt about the research theme is still vivid whenever I recall the scene. I am privileged to have the opportunity to engage in this genetic code expansion project in my pursuit of PhD, under the supervision of Sebastian. His insights into synthetic biology and *C. elegans* research and his optimism towards anything encountered have enlightened me to the research work along my scientific journey. Without his patient guidance and warm encouragement, I would not be able to work through all the difficulties, especially since the outbreak of this ongoing global pandemic.

There are many members of Greiss lab who have kindly helped me. I would like to express the warmest thanks to Lloyd Davis, for teaching me various experimental techniques and for providing a helping hand whenever I asked for. My special thanks go to Ailish Tynan for her meticulousness scoring of worms, and to Kieran Baxter for his assistance in several bombardment experiments. My thank you also goes to Andromachi Xypnitou for teaching me microinjection and for managing the stock of frozen strains during her stay, to

Jack O'Shea for exchanging thoughts and opinions with, to Angeliki Goutou for preparing stock plates, to Sammy Banhste for having a part of her master project under my guidance.

I would also like to thank many colleagues and alumni of Busch lab and Doitsidou Lab, whom I share lab area, experimental equipment and reagents, snacks, ideas, and happiness with. Thanks to Emanuel Busch and Maria Doitsidou for discussing with me about scientific questions and for providing me worm strains and plasmids. Thanks to Qiaochu Li for exchanging ideas about almost every aspect of our doctoral lives and for chatting with me during out-of-hour lab work. Thanks to Feng Xue for hosting home parties with delicious food and fun games, and for creating artistic works that have added zest to my time in the office. Thanks to Daniel Marcu, Elisabeth Fischer, Eugenia Goya, Lourdes Riquelme-Dominguez, Simon Warburtonpitt, and other colleagues for presenting research progress and thoughts on our “Worm people” meetings.

My thankfulness goes to my thesis committee, Andrew Jarman and Maria Doitsidou, for giving me insightful advice on the progress of my research project. I am thankful to the College of Medicine and Veterinary and the University of Edinburgh, for admitting me to this PhD project, and for funding

my research and life in the lovely city, Edinburgh, for three years. I would also like to thank Mayank Dutia for recommending me for the two scholarships I received.

I would not be able to accustom myself to living abroad so quickly, without the companion of my friends Ruoyu Ma, Yixi Chen, Yuting Lu, Naiming Chen and Qiaochu Li. They are always willing to give me mental and behavioural support when I was upset or in trouble, and it is such a pleasure to have spent my leisure time with them.

I would like to express special thankfulness to Jiaqi Ren, my best friend, for his insightful articles about methodological analysis, for giving suggestions on lay summary and optics-related contents of this thesis, and for encouraging me throughout the writing process.

Finally, my utmost gratefulness goes to my beloved parents, for backing me up as always, even across eight time zones. Without their support, I would not have been able to complete any work.

Abbreviations and Acronyms

aaRS	Aminoacyl-tRNA synthetase
aa-tRNA	Aminoacyl-tRNA
Ala-(Dansyl)Ala	Alanyl-2-amino-3-(5-(dimethylamino)napthalene-1-sulfonamido)propanoic acid dipeptide
AMP	Adenosine 5’-monosphosphate
ANOVA	Analysis of variance
Apaf-1	Apoptotic protease activating factor-1
ARTS	Apoptosis-related protein in the TGF-β signalling pathway
Asp	Asparate
ATP	Adenosine 5’-triphosphate
AU	Arbitrary unit
BAK	Bcl-2 antagonist/killer
BAX	Bcl-2-associated X protein
Bcl-2/BCL-2	B-cell lymphoma 2
BH3	Bcl-2 homology region 3
Boc-Lys	*Nε*-(*tert*-butyloxycarbonyl)-L-lysine
bp	Base pair
CARD	Caspase recruitment domain

Caspase	Cysteinyl aspartate specific proteinase
C. elegans	*Caenorhabditis elegans*
CeOpt	Codon optimised to *C. elegans*
C15	$tRNA^{C15}_{CUA}$
CFP	Cyan fluorescent protein
ChR2	Channelrhodopsin-2
CRF1R	Class B GPCR corticotropin releasing factor type 1 receptor
CRISPR	Clustered regularly interspaced short palindromic repeats
Cys	Cysteine
CZ	C-terminal leucine-zipper
(Dansyl)Ala	2-amino-3-(5-(dimethylamino)napththalene-1-sulfonamido) propanoic acid
DIC	Differential interference contrast
DMNB	4,5-dimethoxy-2-nitrobenzyl
DMNPE	3,4-dimethoxy-6-nitrosoacetophenone
DNA	Deoxyribonucleic acid
dsRNA	Double-stranded RNA
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin-protein enzyme

ECL	Enhanced chemiluminescent
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetraacetic acid
EF-Tu	Elongation factor-thermo unstable
ER	Endoplasmic reticulum
eRF1	Eukaryotic translation termination factor 1
F1	Filial 1
F2	Filial 2
F2A	2A element of FMDV
F3	Filial 3
FMDV	Foot-and-mouth disease virus
FPKM	Fragments per kilobase of exon model per million mapped fragments
FRET	Fluorescence resonance energy transfer
FRT	Flp-recombinase target
FUDR	Floxuridine/5-fluorodeoxyuridine
GC3AI	Superfolder GFP-based caspase-3-like protease activity indicator
GFP	Green fluorescent protein
GlnRS	Glutaminyl-tRNA synthetase
Gly	Glycine

HA	Hemagglutinin
HE	Hematoxylin-eosin
HL	Happy linker
HRP	Horseradish peroxidase
Hu	Human
hygB	Hygromycin B
HygR	Hygromycin B resistance gene
IAP	Inhibitor of apoptosis
IgG	Imunoglobulin G
INM	Inner nuclear membrane
IGF	Insulin/insulin-like growth factor
KASH	Klarsicht, ANC-1, and syne/nesprin homology
kb	Kilo base pair
kD	Kilodalton
L1/2/3/4	*C. elegans* larval stage1/2/3/4
LINC	Linker of nucleoskeleton and cytoskeleton
LOV	Light-oxygen-voltage
LPS	Lipopolysaccharide
LS	Long subunit
Lys	Lysine
M15	$tRNA^{M15}_{CUA}$

Mb	*Methanosarcina barkeri*
mCherry	Red fluorescent protein variant from *Discosoma*
MiniSOG	Mini singlet oxygen generator
mito	Mitochondria-localised
Mj	*Methanocaldococcus jannaschii*
MLS	Membrane localised signal
Mm	*Methanosarcina mazei*
mM	Millimolar
N/A	Not applicable
NA	Numerical aperture
nAchR	Nicotinic acetylcholine receptor
NaN	Not a number
NB	*Ortho*-nitrobenzyl or 2-nitrobenzyl
NCAA	Non-canonical amino acid
ncRNA	Non-coding RNAs
NES	Nuclear export signal
NGM	Nematode growth medium
NLS	Nuclear localisation sequence
NMD	Nonsense-mediated decay
nonOpt	Codon unoptimised
NP	Nonylphenol

NPE	1-(2-nitrophenyl)ethyl
Npu	*Nostoc punctiforme*
ns	No significance
NZ	N-terminal leucine-zipper
ONM	Outer nuclear membrane
ORF	Open reading frame
P0	Parental generation
PARP	Poly ADP-ribose polymerase
PATC	Periodic A_n/T_n cluster
PBS	Phosphate buffered solution
PCAA	Photo-caged amino acid
PCCys	*S*-[(*R*,*S*)-1-{4',5'-(methylenedioxy)-2'-nitrophenyl}ethyl]-L-cysteine
PCCRS	Photo-caged cysteine-tRNA synthetase
PCLys	*N*-ε-({(*R*,*S*)-1-[4',5'-(methylenedioxy)-2'-nitrophenyl] ethoxy}carbonyl)-L-lysine
PCKRS	Photo-caged lysine-tRNA synthetase
PCR	Polymerase chain reaction
PH	Pleckstrin homology
piRNA	Piwi-interacting RNA
PKIα	Protein kinase inhibitor (form alpha)

Pol II	RNA Polymerase II
Pol III	RNA Polymerase III
Pro	Proline
Pyl	Pyrrolysine
PylRS	Pyrrolysyl-tRNA synthetase
PylT	$tRNA^{Pyl}_{CUA}$
redox	Reduction-oxidation
RF	Release factor
RFP	Red fluorescent protein
RMCE	Recombinase-mediated cassette exchange
RNA	Ribonucleic acid
RNAse	Ribonuclease
RNAseq	RNA sequencing
ROS	Reactive oxygen species
SAGE	Serial analysis of gene expression
S. cerevisiae	*Saccharomyces cerevisiae*
Sec	Selenocysteine
SECIS	Selenocysteine insertion sequence
Ser	Serine
SOD	Superoxide dismutase
SS	Short subunit

SUN	Sad1 and UNC-84 domain
TEV	Tobacco etch virus
TevS	TEV cleavage site
TF	Transcription factor
Trp	Tryptophan
Tyr	Tyrosine
TyrRS	Tyrosyl-tRNA synthetase
UAS	Upstream activating sequence
UM	Upstream motif
uORF	Upstream open reading frame
UPR	Unfolded protein response
UTR	Untranslated region
UV	Ultraviolet light
UVA	Long-wavelength UV
VC3AI	Venus-based caspase-3-like protease activity indicator
VNC	Ventral nerve cord
WB	Western blot
wt	Wild type
XIAP	X-linked inhibitor of apoptosis
YA	Young adult
YFP	Yellow fluorescent proteins

Z-Lys	*Nε*-(benzyloxycarbonyl)-L-lysine

Contents

Chapter 1: introduction

1.1 Genetic code expansion

1.1.1 Central dogma of molecular biology

Living organisms can be distinguished from the non-living substances by their characteristic capabilities of metabolism and reproduction. From a thermodynamic perspective, both metabolic and reproductive activities of an organism act against a spontaneous tendency of constituent substances of the organism to reach a thermodynamic equilibrium state within an isolated system, i.e., the state with maximal atomistic disorder (more usually described as maximal entropy)[1]. Specifically, an organism through metabolic processes absorb and integrate into itself extrinsic substances of negative entropy (mainly from food), while excreting wastes of positive entropy into its environment. And establishment of these metabolic mechanisms are guided by genome. Meanwhile, reproductive cells from primordial germ cells to gametes are of relatively low entropy within an organism and the information embedded in their genome guides in the offspring a *de novo* establishment of similar metabolic mechanisms to the parent[2].

Therefore, accurate transmission and translation of genomic information are essential aspects of reproductive and metabolic activities for the maintenance of a low-entropy state of individual organisms, of their offspring and eventually of their population. The flow of information through bio-molecular carriers is summarised in central dogma of molecular biology: DNAs can be replicated for inheritance, or transcribed for passing the information to messenger RNAs (mRNA) that can subsequently instruct the production of proteins for executing metabolic processes[3].

The role of messenger RNA in the central dogma is pivotal. A mRNA molecule is transcribed from the sequential nucleotide permutation of a DNA molecule and subsequently encodes a protein with a corresponding permutation of twenty amino acids[3]. The genetic information embedded in mRNA nucleotides is decoded by a set of rules that is, apart from very few exceptions, universal to all organisms: from the start of a gene coding region, each non-overlapping nucleotide triplet (codon) uniquely maps to one amino acid or signals termination of translation[4-6]. Two steps constitute the codon decipherment during this protein synthesis (or translation) process, which is relayed by transfer RNAs (tRNAs)[7]. First, a tRNA is charged with its cognate amino acid by a corresponding aminoacyl-tRNA synthetase (aaRS)[8]. Then the aminoacylated tRNA gets recruited to a working ribosome, according to

the match between its anticodon nucleotides and the codon on standby of the ribosome-bound mRNA in translation[9]. Therefore, the cognate amino acid is carried into the ribosomal catalytic site and subsequently forms a peptide bond with the growing nascent polypeptide, which is what “incorporation” refers to in this thesis.

1.1.2 Transfer RNAs (tRNAs) and aminoacyl-tRNA synthetases

1.1.2.1 Transfer RNAs

A study on a yeast alanine tRNA first proposed a four-lobed clover model of tRNA secondary structure[7]. In this model, intramolecular hydrogen bonds between complementary bases stabilise the tRNA “clover” stems, resulting in three loops and one open stem (present in Figure 1.1)[7]. The acceptor stem forms from paired nucleotides of 5’ and 3’ ends of a tRNA molecule. Particularly, the 3’ end terminates in a highly conserved nucleotide sequence CCA for tRNA aminoacylation. Opposite to the acceptor stem is the anticodon loop. As the name suggests, a nucleotide triplet at the loop centre (position 35-37) constitutes the anticodon complementary to the mRNA codon and hence the tRNA is matched up with the correspondent codon to the cognate

amino acid to be charged. The remaining two loops contain minor bases guanine and uridine-derived pseudouridine (ψ) and dihidrouridine (D), therefore are named as TψC loop and D loop, respectively. Later on, Kim *et al.* modelled the tertiary structure of a yeast phenylalanine tRNA in three dimensions[10]. Base-stacking interactions and hydrogen bonds help the tRNA fold further from a two-dimensional "clover" to a three--dimensional L shape, where both anticodon loop and acceptor stem remain accessible for intermolecular reactions.

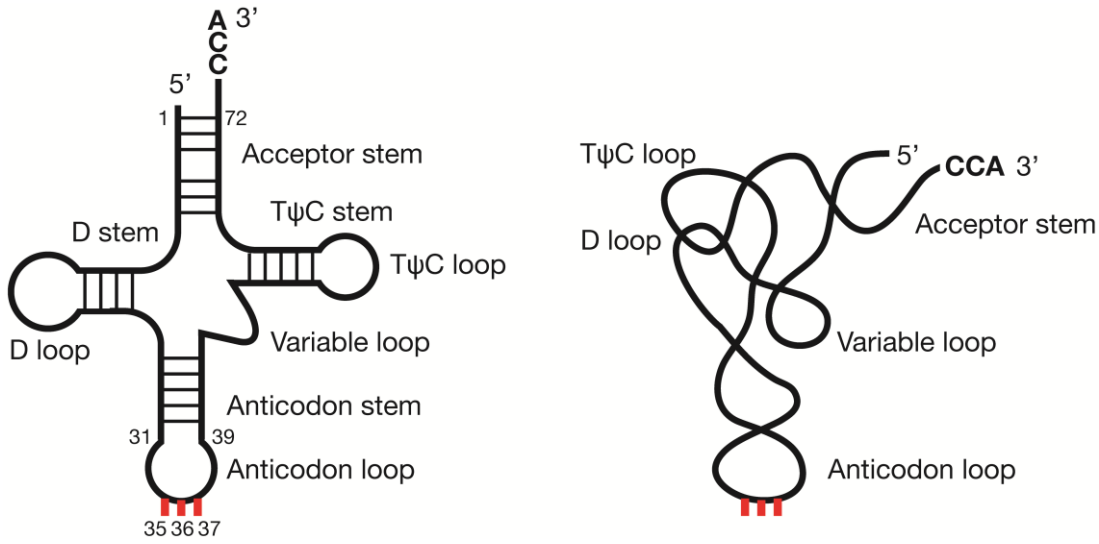

Figure 1.1 Schematic of a transfer RNA.

Nucleotide pairing of a tRNA facilitates the nucleic acid to form a secondary structure (left side) that consists of four stems (in clockwise are acceptor stem, TψC stem, anticodon stem, and D stem) and four loops (in clockwise are TψC loop, variable loop, anticodon loop and D loop)[7]. Further folding generates the tertiary structure

on the right side[10]. Nucleotide numbering of tRNA starts from the 5' end and the anticodon nucleotides (in red) take up the positions 35 to 37.

High fidelity of the decoding of mRNA sequences to proteins requires each tRNA involved to be charged only with a cognate amino acid in aminoacyl-tRNA synthesis. Despite their structural similarities, tRNAs each harbour identity elements composed of variable nucleotides that help aaRS to specifically identify their cognate tRNAs[11]. Depending on their participation in either promoting specific aminoacylation or inhibiting a mismatch, these elements are termed determinants or anti-determinants, respectively[11]. Determinants outnumber anti-determinants and cluster at positions directly involved in tRNA recruitment, e.g., the five base pairs at the end of acceptor stem, seven sites of the anticodon stem, and a discriminatory base at position 73[11, 12].

1.1.2.2 Aminoacylation of tRNA by aminoacyl-tRNA synthetase

An aminoacyl-tRNA synthetase mediates two-step aminoacyl-tRNA synthesis (Figure 1.2A&B). The first step is the activation of a specific amino acid[8, 13]: with the chemical energy from ATP hydrolysis, an aaRS transforms a cognate amino acid into an aminoacyl adenylate (aa-AMP) intermediate that remains

non-covalently bound to aaRS active site. Next, the aaRS binds a cognate tRNA and via esterification transfers the bound aa-AMP to either the 2' or the 3' hydroxy group of the 3’-ribose of the ending adenosine of tRNA, generating an aminoacyl-tRNA (aa-tRNA). The aa-tRNA is able to enter the A site of an mRNA-bound ribosome upon forming a complex with EF-Tu (elongation factor-thermo unstable)[9]. In the facility of the ribosomal peptidyl transferase centre, the amino moiety of aa-tRNA attacks the carbonyl carbon of the peptidyl-tRNA in the ribosomal P-site to form a peptide bond, thus the amino acid gets incorporated into the growing polypeptide[9] (Figure 1.2C).

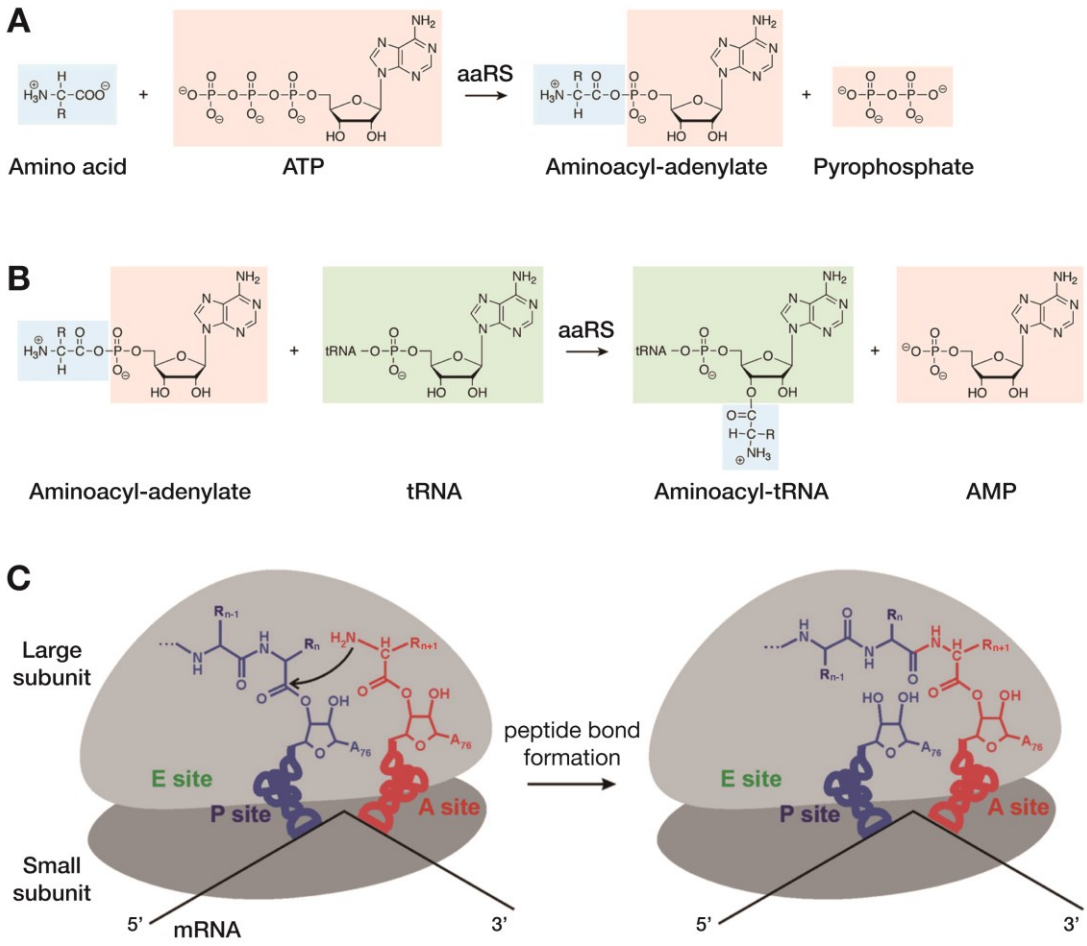

Figure 1.2 Schematic of aminoacyl-tRNA synthesis

(A) Adapted from Beringer *et al.*[9]. The first step forms an aminoacyl-adenylate (aa-AMP), via an aaRS-catalysed condensation between the α-carboxylate of the amino acid and the α-phosphate of an adenosine triphosphate (ATP) molecule. **(B)** Adapted from Beringer *et al.*[9]. The second step requires a cognate tRNA to bind the aaRS. Then transesterification takes place between the α-carboxylate of aa-AMP and either 2'-OH or (the illustrated) 3'-OH group of the 3'-ribose of A76 of tRNA to form an aminoacyl-tRNA. **(C)** Adapted from Francklyn *et al.*[14]. Schematic of peptide bond formation between an aminoacyl-tRNA and a peptidyl-tRNA in ribosome. Peptidyl transferase centre, located in the large subunit (light grey) of

ribosome, catalyses the reaction: the amino group of aminoacyl-tRNA in the A site (red) attacks the carbonyl carbon of the peptidyl-tRNA in the P site (blue), elongating the peptide of the peptidyl-tRNA (A site) by one amino acid as well as a deacylated tRNA (P site).

To ensure the specificity of the tripartite tRNA aminoacylation, aaRS needs to specifically recognise not only its cognate tRNA but also its cognate amino acid. The first step of amino acid recognition works through an induced-fit mechanism of the binding pocket that strictly excludes all amino acids larger than the cognate amino acid[15]. Up to this stage, smaller and similar amino acids to the cognate one are not sufficiently selected against. These non-cognate aa-AMPs get trapped and hydrolysed in the second step by “fine sieve” - a special editing domain of aaRS, thereby leaving only the cognate aa-AMP to be subsequently transferred to tRNA[16, 17].

1.1.3 Exceptions to the common codon assignments

Canonical codon assignment maps 64 nucleotide triplets to 20 canonical amino acids as well as a signal of termination[5]. The same set of decoding rules guides the decoding of mRNA nucleotide permutations into an amino

acid sequence in organismal synthesis of almost all proteins. This universality was hypothesised to have resulted from a “frozen accident”: accompanied with persistent evolution of the collective proteome of all organisms, the complexity of the decoding rule system reached a threshold where ambiguity in protein encoding could no longer be tolerated[18, 19]. In other words, any reassignment or addition to the established codons would cause harmful amino acid substitutions in too many proteins for the organism to survive, henceforth the codon assignment as it is today became invariable[18, 19].

This hypothesised ultimacy in codon assignment, however, was later found violated in some species of prokaryotes, archaea and eukaryotes[20]. Ten deviations from the canonical assignment have so far been revealed in nuclear codes, while for mitochondrial codes there are 16 deviations discovered[20]. Alterations include codon vacancy, reassignment and expansion (illustrated in Figure 1.3). To begin with, reassignments of UGA for tryptophan and of AUA from isoleucine- to methionine- encoding were discovered in yeast and human mitochondrial genomes[21, 22]. A single cell contains many mitochondria, and these semi-autonomous organelles each encode a limited number of proteins that are confined in the organelle[23-26]. It is therefore possible that a codon reassignment sporadically occurred in one mitochondrion without interfering the protein synthetic events in the

cytosol or other mitochondria by canonical decoding rules[20]. Hence such a reassignment was tolerated and inherited from then on. Notably, all 10 different exceptions found in nuclear codes are contained within the set of 16 mitochondria code variations[20]. Two compatible possibilities may explain this subordinate relationship: on one hand, communal codon variations may form through a conserved mechanism of codon reassignment prior to the first endosymbiosis of mitochondria in ancient eukaryotic cells[23]. On the other hand, the 6 unique mitochondrial codons possibly occurred in the adaption of ancient mitochondria for endosymbiosis[20, 27].

Codon reassignments and the natural degeneracy of canonical codons share molecular mechanisms as summarised in “wobble hypothesis”[6, 27]. In this theory, Crick purposed that pairing of the third base of a triplet codon with the first base of an anticodon (tRNA position 35) is not restrict, allowing a tRNA of one amino acid to decode several different codons[6]. One example is guanine-uracil pairing. Also, the existence of a minor base at position 35 of tRNA permits additional pairing possibilities. Particularly, post-transcriptional deamination of an adenosine at this position by adenosine deaminase generates an inosine whose hypoxanthine base (both are denoted as “I”)[28] can pair with three (A, C, U) of the four major bases[6]. Thus the translation of all the 61 sense codons needs only 32 different species of tRNAs. Similarly

for the cases of codon variation in mitochondria, there are idiosyncratic modifications on the tRNA position-35 nucleotide that promote the codon degeneracy[20, 29, 30]. As few as 22 tRNA species can undertake the amino acid encoding task of a mitochondrion[31]. Therefore, such codon reassignments benefits the host organelle by reducing the number of different tRNA species required for protein translation and these mitochondria gain a reproductive advantage over others and hence prevailed through evolution[20]. Unlike the assumption of “frozen accident” theory, proteome of most organisms tolerate mild flexibility in codon assignments possibly as neutral fluctuations with evolutionary potential[20].

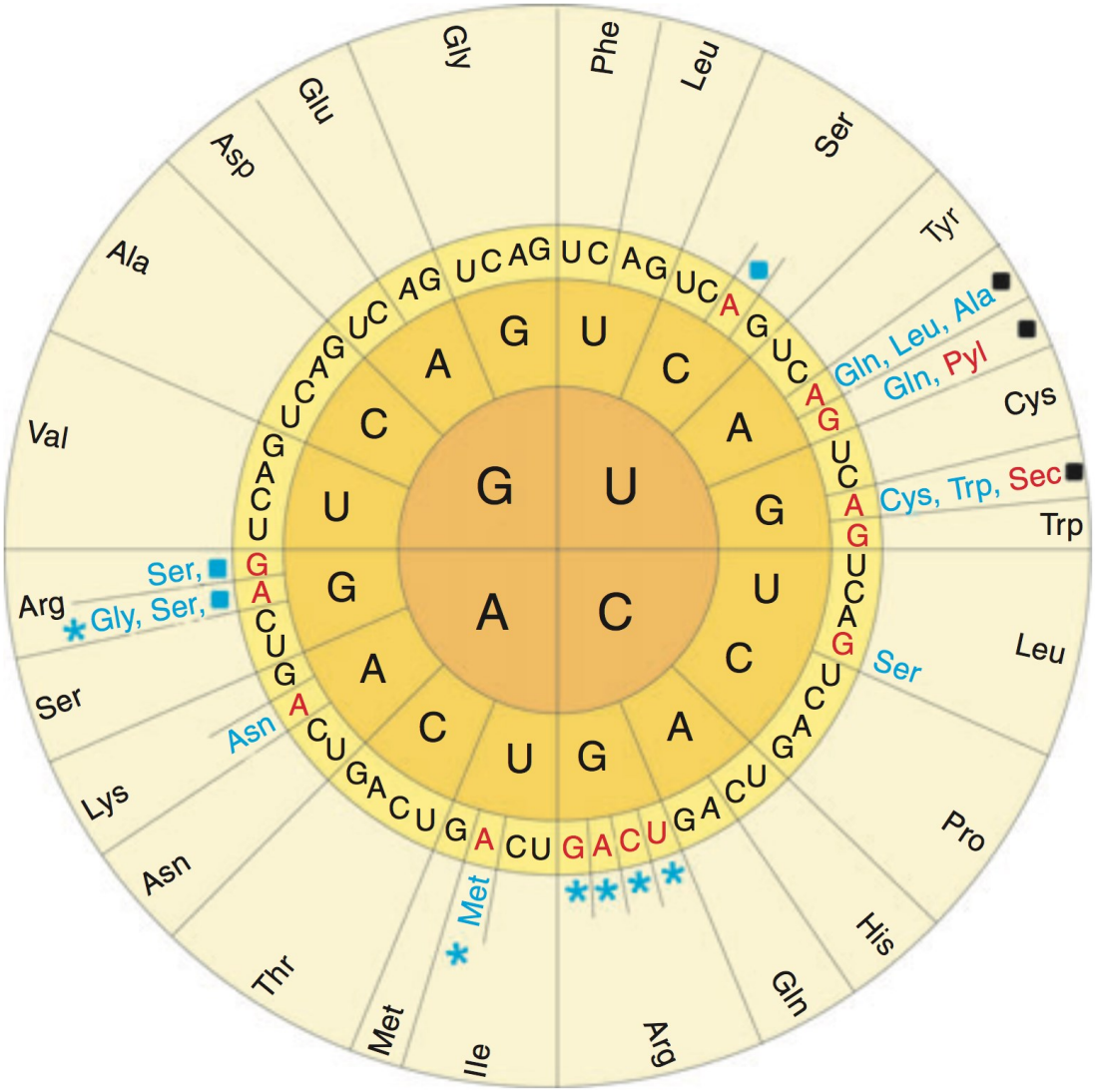

Figure 1.3 The many-to-one mapping of codons to amino acids.

Duplicated from Ambrogelly *et al.*[20]. A codon is radially illustrated: its first nucleotide base is in the innermost circle, the second base is in an outer layer and the last is in the outmost yellow-shaded circle. 61 Sence codons are canonically assigned to the amino acids (as three-letter abbreviations) on at the outmost circle and three stop codons (UAA/ochre, UAG/amber, UGA/opal) signal the stoppage of protein translation (shown as black squares). Red bases indicate the codon reassignments to either a standard amino acid (in blue) or a non-canonical amino

acid (in red, Sec for selenocysteine and Pyl for pyrrolysine). Blue asterisks indicate codon vacancies in some organelles or organisms.

Another subgroup of codon changes is the switch of a stop/nonsence codon (namely UAA/ochre, UAG/amber, UGA/opal)[32-34] to a sense codon. Termination of protein translation upon a stop codon is usually mediated by its correspondent release factor(s)[35]. A tRNA with its anticodon able to pair with a nonsense codon can suppress this termination process and encode an amino acid, which is termed nonsense suppression[36]. Nonsense-suppressing tRNAs, or nonsense-suppressors, were firstly identified in bacteriophage T4[36]. Then more tRNAs of this kind were found in prokaryotes like *E. coli*[37] and even eukaryotic species including yeast *Saccharomyces cerevisiae*[38] and *C. elegans*[39].

All the aforementioned codon alterations exclusively match a “word” (codon) with an already existed “meaning” (namely, to add one canonical amino acid or to terminate). As 64 codons use up all permutations of 4 different bases in triplets and have been all assigned, it seemed that the twenty canonical amino acids were the only available building blocks of protein translation in nature, with or without reassigning codons. The assumption has been negated by a series of discoveries that two non-canonical amino acids

(NCAAs), selenocysteine (Sec)[40, 41] and pyrrolysine (Pyl)[42, 43], can be naturally encoded.

The natural encoding of selenocysteine was first found in *E. coli*[41] and mice[40]. Two steps are needed to charge an opal (UGA)-suppressor tRNA with selenocyteine: a seryl-$tRNA^{Ser}$ synthetase acylates the suppressor tRNA with a serine and then the serine is converted into a selenocysteine[44, 45]. No confusion of opal decoding gets introduced into protein translation, though, with the addition of this "21st amino acid": apart from the few selenoproteins, opal codons appearing in the genome are still deciphered as the signal of translation termination[46]. To break the default termination of protein translation at a UGA codon, adequate amount of selenocysteine-charged tRNAs (Sec-$tRNA_{UCA}$) has to be recruited to the vicinity of a working ribosome to compete with release factors[47]. This is mediated by SelB, a protein recognises a selenocysteine insertion sequence (SECIS) in the mRNA[48]. This signal element lies only in the 3' proximity of the sense UGAs among all opal codons, so the incorporation of selenocysteine will not interfere with normal terminations[48].

Pyrrolysine, a derivative of lysine, is first discovered as "the twenty-second amino acid" in the natural synthesis of the enzyme methylamine

methyltransferase of archaea species *Methanosarcina barkeri* (*M. barkeri* or *Mb*)[42, 43]. Pyrrolysine is directly charged onto a cognate amber suppressor tRNA ($tRNA^{Pyl}_{CUA}$) by pyrrolysyl-tRNA synthetase (PylRS)[43]. Contrary to selenocysteine incorporation that needs signalling elements SelB and SECIS, pyrrolysine encoding is amber-directed and can be independent of any specialised factor or recognition site as a contextual requisite[49-51]. The straightforward amber decoding suggests that it is possible to site-specifically install new amino acids into proteins. Next section will introduce how this possibility can be turned into a reality.

1.1.4 Genetically encoding non-canonical amino acids (NCAA)

Proteins are translated as biopolymers of a limited variety of amino acids. Nevertheless, the natural occurrence of code expansion indicates a possibility of manipulating proteins at the monomer level via site-specific installation (incorporation) of “designer” amino acid. Proteins of interest can be endowed with properties of the incorporated NCAAs, consequently becoming measurable and/or controllable with unprecedented spatiotemporal resolution[52].

The successful incorporation of a NCAA via codon expansion requires the fulfilment of following conditions: 1) a vacant codon pre-installed in the coding sequence of interest; 2) a tRNA that exclusively decodes the vacant codon and is acylated exclusively with the NCAA; 3) an aaRS that only charges the NCAA onto the cognate tRNA as is described by Condition 2. The conditions combined ensure the orthogonality of NCAA incorporation, meaning that the aaRS/tRNA pair is the sufficient and exclusive prerequisite for accurately translating the vacant codon to the NCAA. Also, an aaRS/tRNA pair being orthogonal means that neither the tRNA nor the aaRS cross-reacts with endogenous genetic components, including codons, tRNAs and aaRSs.

Initial attempts towards experimental genetic code expansion turned amber stop codon (UAG) into a sense codon for various canonical amino acids[53, 54]. Amber suppressor tRNAs were engineered from the cognate tRNA(s) of a target amino acid and then tested for amino-acid encoding with an acceptable efficiency[53, 54]. The amber codon was chosen by two reasons: 1) in *E. coli* it is the least abundant one stop codon of all three[55]; 2) amber suppression can be naturally tolerated, as amber suppressor tRNAs have been found in various species[38, 39]. The selected $tRNA_{CUA}$ variants succeeded in incorporating canonical amino acids at the inserted amber codons of target proteins expressed *in vitro* and in *E. coli* cells[53, 54].

The next step taken was modifying a canonical aminoacyl-tRNA synthetase for charging the engineered $tRNA_{CUA}$ with NCAA[56]. At the time, the crystal structures of both *E.coli* glutaminyl-tRNA synthetase (GlnRS) and its $tRNA^{Gln}$ had been revealed[57], so they were chosen as the scaffold of an orthogonal aaRS/$tRNA_{CUA}$ pair[56]. The rational engineering consisted of two steps: the $tRNA^{Gln}$ was first modified for supressing the amber codon whilst being orthogonal to all endogenous aaRSs (GlnRS include) and then GlnRS was engineered to recognise the modified $tRNA^{Gln}$ for aminoacylation[56]. However, it turned out that the resultant synthetase only inefficiently acylated the engineered tRNA, possibly due to its cross-reaction with endogenous $tRNA^{Gln}$ in *E. coli*[56]. Later, a *Saccharomyces cerevisiae* GlnRS/$tRNA^{Gln}$ pair was engineered for decoding bacterial amber stop codon, for it shares less similarity with bacterial aaRS[58]. It was shown in *E. coli* that the modified GlnRS/$tRNA^{Gln}_{CUA}$ pair can be charged with various analogues of glutamine and glutamic acid, but no incorporation in protein translation was observed[58]. It suggests that the tRNA variant, though with its anticodon replaced by CUA, remains inefficient in decoding amber codon.

It was discovered that a tyrosyl-tRNA synthetase (TyrRS) of archaeal species *Methanocaldococcus jannaschii* (*Mj*) lacks anticodon-binding domain and hence it still charges an anticodon-mutant derivative of its cognate tRNA with

little efficiency drop[59]. Furthermore, the eukaryotic-like tRNA recognition element of *Mj* TyrRS does not fit with prokaryotic tRNAs[59]. By mutating *Mj* $tRNA^{Tyr}$ to an amber-pairing variant, Wang *et al.* demonstrated the experimental genetic code expansion for NCAA incorporation first in an organism (*E. coli*)[60].

However, it was later found that bacterial endogenous synthetases were also able to aminoacylate this *Mj* $tRNA^{Tyr}_{CUA}$[60]. To improve the orthogonality of engineered tRNA/synthetases, Wang *et al.* implemented an engineering approach called directed evolution, whose protocol alternated between two selection steps (negative- and positive-selection)[61]. The cognate amino acid to the tRNA was supplemented throughout the entire procedure. In the negative selection, candidate $tRNA^{Tyr}_{CUA}$ variants were co-expressed with a lethal toxin gene with a premature amber codon[61]. If a particular tRNA variant gets acylated by endogenous synthetase, then the amber in the toxic gene will be read through in translation and a full-length toxin will be produced to kill the host cells of the tRNA candidate. Candidates that survived the first selection step were then put through a positive selection, in which they are expressed together with *Mj* TyrRS and an ampicillin- resistant ß-lactamase gene inserted with a premature amber codon[61]. Here requires the tRNA variant to be acylated by *Mj* TyrRS for suppressing the premature amber in

the resistance gene for the carrier bacteria to survive antibiotic selecting. Thus this selection step picks variants of $tRNA^{Tyr}_{CUA}$ that can be charged by the synthetase. In the practice of the protocol, several alternating rounds of the negative and positive selections were taken for obtaining a $tRNA_{CUA}$ candidate with efficient suppression and enhanced orthogonality for the host context[62].

A similar procedure has been set for evolving the orthogonal synthetase for a specific NCAA[62]. First, various mutations are introduced into the amino-acid binding pocket of a natural synthetase to generate a library of candidates. Then each candidate was expressed in pair with the previously selected $tRNA_{CUA}$ in the absence or presence of the NCAA for negative or positive selection, respectively. The evolution procedure has succeeded in generating new *Mj* TyrRSs for incorporating many different non-canonical amino acids, most of which are phenylalanine derivatives[62-64].

However, *Mj* $tRNA^{Tyr}$ harbours similar major tyrosine identity determinants to those of eukaryotic $tRNA^{Tyr}$, which impairs the orthogonality of *Mj* TyrRS/$tRNA^{Tyr}$ pair in eukaryotes and therefore limits its implementation in genetic code expansion[65]. To solve the problem, an *E. coli* TyrRS/$tRNA^{Tyr}$ pair was used for evolving orthogonal aaRS/tRNA pairs for expanding

eukaryotic genetic codes[66]. An amber-decoding TyrRS/$tRNA_{CUA}^{Tyr}$ pair was created from this approach and its incorporation of NCAAs in *S. cerevisiae* is the first demonstration of genetic code expansion in eukaryotic cells[66]. Shortly afterwards, by using the same evolution strategy, an amber-suppressing *E. coli* Leucyl-tRNA synthetase (LeuRS)/$tRNA^{Leu}_{CUA}$ pair was made and was shown able to encode new amino acids in yeast cells with high orthogonality[67].

The above two aaRS/tRNA pairs have a bacterial origin, so they are not orthogonal in *E. coli* and have to be further evolved in yeasts for incorporating new NCAAs. The yeast cells have low transformation rates[68], so the process is inefficient and the selection will be confined within a relatively small library of aaRS/tRNA candidates. One solution is to find a natural aaRS/tRNA pair being orthogonal to endogenous pairs of both prokaryotes and eukaryotes, therefore directed evolution of this pair for specific NCAA recognition can be efficiently conducted in *E. coli* hosts. Also, the resultant encoding system can be directly applied in the eukaryotic context. Archaeal aaRS/tRNA pairs are candidates for this, as archaea separated in primordial stage of evolution from prokaryotes and then from eukaryotic species[69].

The aaRS/tRNA involved in the natural decoding of amber codon for pyrrolysine in *M. barkeri* [43], *Mb* PylRS/tRNA$^{Pyl}_{CUA}$, was experimentally expressed in *E. coli* and was shown able to incorporate the 22nd amino acid[70] as well as its analogues[71]. By using a bacteria-based directed evolution protocol, Neumann *et al.* successfully evolved orthogonal *Mb* PylRS that encodes *N*-ε-acetyl-L-lysine in *E. coli* [72]. The work is a milestone of the engineering of PylRS for artificially expanding the substrate scope of PylRS/tRNA$^{Pyl}_{CUA}$ beyond its cognate pyrrolysine and pyrrolysine analogues. Shortly afterwards, lysine derivatives also got incorporated by evolved PylRS/tRNA$^{Pyl}_{CUA}$ variants in mammalian cells[73], as the first NCAA incorporation in eukaryotic cells by using orthogonal aaRS/tRNA pairs evolved in bacteria. It was found that some other lysine derivatives can also be charged onto the suppressor tRNA by the original *Mb* PylRS, possibly due to a permissive binding mechanism of the synthetase[74, 75].

Structural analysis of the PylRS of a related species, *Methanosarcina mazei* (*M. mazei* or *Mm*), has revealed that the synthetase binding pocket recognises pyrrolysine mainly through its hydrophobic interaction with the pyrrole ring of pyrrolysine[76]. This interaction seems sufficient for a substrate to be recognised by PylRS, as pyrrolysine analogues with cyclopentane rings can also be incorporated into a polypeptide via the PylRS/tRNA$^{Pyl}_{CUA}$ pair[77].

Importantly, it was first found in a similar bacterial aaRS/tRNA pair[78] and later confirmed by the counterpart pair of *M. mazei* [79] that the compact core of $tRNA^{Pyl}_{CUA}$ rather than the anticodon determines the tRNA interaction with PylRS. With this structural feature, the PylRS/$tRNA^{Pyl}_{CUA}$ system not only has high orthogonality to the endogenous aaRS/tRNA pairs of prokaryotes and eukaryotes, but also allows efficient anticodon alterations for decoding possibly other stop codons and even quadruplet codons[80, 81].

To express genes of PylRS/$tRNA^{Pyl}_{CUA}$ in the eukaryotic context, the below difference of transcription machineries of eukaryotes from prokaryotes should be considered. RNA Polymerase III (Pol III) transcribes eukaryotic tRNA genes, whose recruitment relies on genetic elements Box A and Box B that reside within the internal promoters of the tRNA genes[82, 83]. In contrast, archaean tRNAs lack such elements[73]. To improve the expression efficiency of the archaean tRNA in mammalian cells, a U6 promoter of mammalian Pol III gene was added to 5' of the tRNA coding region[73]. One way used for expressing *Mm* $tRNA^{Pyl}_{CUA}$ in a yeast cell is to replace the downstream coding region of a bicistronic host tRNA gene with an archaean tRNA sequence, so the eukaryotic internal A- and B- box promoters were eventually attached to the suppressor tRNA[84].

Since its first application in single-cell systems like *E. coli* and cultured cells[60, 66], genetic code expansion has facilitated NCAA incorporation and the development of many NCAA-based protein analytic methods[52]. And the list of NCAAs that are available for *in-vivo* incorporation keeps growing since the first successful incorporation of *o*-methyl-L-tyrosine[52, 60]. The NCAAs can be categorized into four major groups as below.

1) Label NCAAs, including fluorescent types for imaging[85, 86], halogen- or metal-labelled types for crystallography[87, 88], isotope-containing types for nuclear magnetic resonance[89], redox NCAAs for electron transfer[90], spin-labelled NCAAs for electron paramagnetic resonance[91], *etc.*. The NCAA incorporation renders correspondent protein measurements high spatial precision while NCAA has little interference with the structure of interest[92].

2) Cross-linkers for interactome mapping[93] and through an intramolecular linkage for control of protein activity[94]. These NCAAs are useful for discovering transient and weak interactions between proteins of interest[92].

3) bio-orthogonal linkers that can perform multiple functions depending on which type of ligand (e.g. labels, protein inhibitors/activators) is later attached[95]. Three aspects constitute bio-orthogonality[96]: i) the NCAA should be innocuous *per se*; ii) the reagent should react selectively with its

ligand *in vivo* and iii) the reaction should generate a covalent bond without reactive or cytotoxic byproducts.

4) NCAAs for *in-vivo* control of protein activity, whose side chain contains “caging” groups that can be either removed optically (photo-caged)[67] or chemically[97, 98]. This enables target protein tuning with high spatiotemporal resolution and controllability[92].

Successful genetic code expansion in bacteria and eukaryotic cells with engineered archaeal PylRS/ $tRNA^{Pyl}_{CUA}$ pairs suggests a possibility of expanding the codons of multicellular organisms. Expansion of animal’s genetic code first succeeded via amber suppression in nematode *Caenorhabditis elegans*[99], which verifies the potential of code expansion on the level of metazoan species. Since then, NCAA incorporation has been demonstrated in other model species like the fruit fly *Drosophila melanogaster*[100],zebrafish and mouse[101]. Next sections will introduce how an *Mm* pyrrolysine-encoding tRNA/synthetase pair gets adapted for incorporating NCAA in *C. elegans*.

1.2 Model organism *Caenorhabditis elegans* and biotechnological tool development

1.2.1 The general biology of *Caenorhabditis elegans* and basic methodology

Early in 1960s, Sydney Brenner began to contemplate solving questions on the development and nervous system of metazoan organisms by implementing the frontier knowledge about molecular biology at the time[102]. The model species for this research purpose had to meet several criteria[103]: i) it should be easy to maintain in laboratory and be amenable to genetics; ii) it has a relatively simple nervous system; iii) it should have similarities in the molecular and cellular processes with other higher animals. *Caenorhabditis elegans* (*C. elegans*), a soil-based non-parasitic species of nematode worm, was found a good candidate. A wild isolate (from Bristol, UK) named N2 was cultured as the laboratory wild-type strain with the conditions later set as a standard[103].

Nowadays, *C. elegans* is one of the most important model organisms. The species has many advantages. First, the animal can be easily maintained in the laboratory with *E. coli* as food. As a small animal of approximately 1 mm length in adulthood, *C. elegans* has a relatively short growth cycle (Figure 1.4):

the development from an egg to an adult takes around 3 days at 25℃. The maturation comprises 4 larval stages (L1 to L4) with 4 molts[104]. They are of two sexual forms, hermaphrodite and male, and sexual dimorphism starts to display from the fourth larval stage (L4) when worms are still sexually immature[105]. The germline of a hermaphrodite becomes transiently male in L4 stage and produces sperm before its oogenesis in adulthood[106]. Hence, even in the absence of males, a hermaphrodite can use the previously made sperm to fertilise eggs it produces as an adult to produce progeny[107]. One hermaphrodite usually gives birth to 200-300 eggs throughout adulthood and the vast majority (> 99%) of the progeny are hermaphrodites[105]. These traits altogether facilitate a rapid generation of an isogenic population of worms. Importantly, crossing hermaphrodites and males of selected genetic background can easily generate double mutants of desired genotypes.

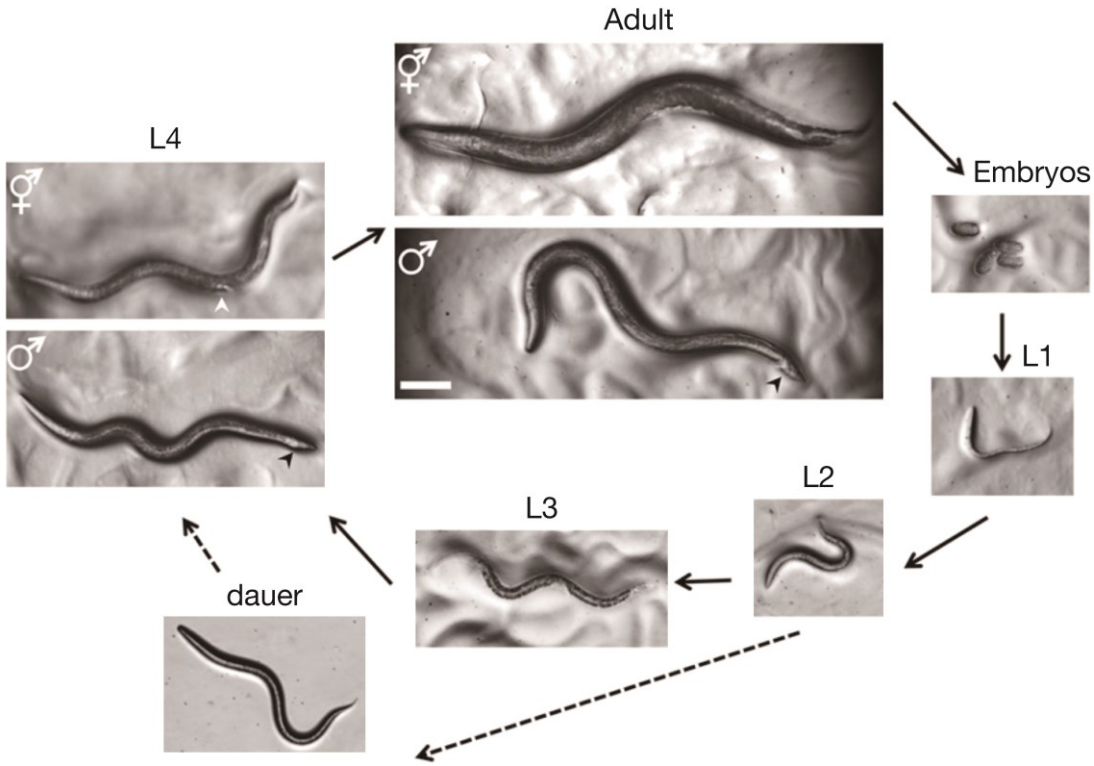

Figure 1.4 Life cycle of *C. elegans*.

Adapted from Corsi *et al.*[104]. The growth of *C. elegans* from an egg to an adult goes through 4 larval stages, named L1 to L4. From the L4 stage, the sexual dimorphism between hermaphrodites (♀) and males(♂) becomes apparent: a hermaphrodite L4 has a half circle-like vulva at the ventral mid of its body (white arrowhead) and a tapering tail, while a male has a characteristic wider tail (black arrowhead) that will be fan-shape when matures in adulthood (black arrowhead). Dauers are the larvae with a developmental arrest in a harsh environment and have the thinnest girth among all larval stages[107]. Scale bar = 100 μm.

The second advantage of the species is its transparency. Individual cells and even some subcellular structures like cell nuclei within a worm can be directly

observed using DIC optics[104]. And a protein of interest can be easily visualised in intact worms with fluorescent tags such as GFP for live imaging or strain screening[108]. Also, fluorescently labelled antibodies can be used to detect proteins, and *in-situ* hybridisation with labelled oligonucleotides can easily detect and visualise target mRNAs, in whole worms[109].

Additionally, every individual worm has an invariant number of somatic cells whose complete lineages have been mapped[110-114], which has facilitated studies on worm development as well as individual cell function. Also, *C. elegans* has a relatively simple nervous system: an adult male worm has 387 neurons[115] while a hermaphrodite has only 302 neurons, with their relations reconstructed as wiring diagrams[116, 117]. It is therefore convenient to use *C. elegans* for studying relationships between neuronal circuits and behaviour.

Additionally, the nematode species has a high similarity in molecular and cellular processes with other metazoan animals. For instance, more than 38% protein-coding genes of *C. elegans* have respective predicted human homologues[118] and reversely, at least 60% human genes have *C. elegans* orthologues[119]. Many discoveries of human diseases came from studying conserved worm genes, including the association of presenilin with familial

Alzheimer's disease[120-122], the link of insulin signalling pathway with Type-II diabetes[123, 124], and the role of serotonergic signalling in one of the antidepressant mechanisms[125].

The ease of genetic manipulation is another important merit of *C. elegans* as a model animal. One large category is forward genetic approaches. They start from artificially induction of random worm mutants for a genotype screen, and then deduce the natural role(s) of a mutated gene(s) through comparing the mutant phenotype with wild type worms[126]. Mutagenesis of worms can be conducted with chemical mutagens like ethylmethane sulphonate[103], short wavelength ultraviolet light illumination[127], or transposon, a mobile genetic component that can jump from its host position within a genome to another random site[128, 129]. Mutant genes that contribute to the phenotype can be rapidly located by strategies based on whole-genome sequencing of the mutant worms, with the known genome of wild-type isolates for reference[130, 131]. With the genes located and sequenced, gene editing tools can then be used to generate specific mutant strains for a better understanding of gene functions, which is the strategy of "reverse genetics"[126]. Reverse genetic approaches are commonly combined with exogenous DNA transference means like microinjection[132, 133] or biolistic bombardment[134, 135]. More recently, transgenic worms can be generated from direct genomic editing by

using DNA insertion tool based on Mos1 transposon[136, 137] or CRISPR-Cas9 system[138-142]. In addition, gene silencing from RNA interference (RNAi) was first described in *C. elegans*, and then developed as a common way of knocking down the activity of target genes[143-146].

There are also quantitative ways of profiling gene expression levels at the genomic scale, such as SAGE (serial analysis of gene expression)[147] and DNA microarray[148]. More importantly, data systematically collected thereby have been assembled into online databases available to all researchers of *C. elegans*[149, 150]. The ever-growing list of such publicly accessible knowledge bases repays to *C. elegans*-based biotechnological studies and the success of genetic code expansion of this multicellular invertebrate is one vivid illustration.

1.2.2 Adaption of genetic code expansion in *C. elegans*

Previous studies have found in *C. elegans* that several mutant variants of a $tRNA^{Trp}$ have *in vivo* amber suppression activities meanwhile worms of the mutant strains were viable[39, 151, 152]. This suggests a possibility of using amber suppression for encoding non-canonical amino acids (NCAA) in worms.

As anticipated, *C. elegans* has become the first animal of NCAA incorporation[99]. Also, Greiss *et al.* demonstrated that an orthogonal tRNA/synthetase pair (PylRS/ $tRNA^{Pyl}_{CUA}$) evolved in bacterial can encode NCAAs in an intracellular context of multicellular organisms[99]. More importantly, code expansion in worms may allow the development of tools for nearly real-time *in-vivo* protein control of biological processes that can only be studied in a multicellular context, such as development, intercellular communication and neuron-based behaviour.

In the first report of genetic code expansion in *C. elegans*, a fluorescent reporter of amber suppression was constructed as P*rps-0*::GFP::amber::mCherry::HA::NLS::*unc-54*[99]. The ribosomal protein promoter of *rps-0* expresses the construct throughout the worm[153], which was also used for ubiquitously expressing the PylRS[99]. The reporter was a fusion of two proteins of different fluorescence, GFP and mCherry, linked by an amber stop codon as the site for incorporation. "::*unc-54"* stands for the 3' untranslated region (3' UTR) from gene *unc-54* that helps stabilising mRNAs of the construct[99]. Among all three types of upstream motifs (UM1-3) of *C. elegans* non-coding RNAs (ncRNA), UM3 is exclusively upstream to stem-bulge ncRNAs that show structural straits (terminated in oligo-T and uncapped) of RNA polymerase III transcription[154, 155]. With

UM3-containing CeN74-1 promoter added to the 5' of its sequence, the archaeal $tRNA^{Pyl}_{CUA}$ gene becomes able to recruit worm endogenous RNA polymerase III [99, 154, 155]. And the 3' region of *sup-7*, the gene of worm amber suppressor $tRNA^{Trp}_{CUA}$, was used for termination this transcription[99].

The above constructs was firstly transferred into worms in a *lin-15B(n765)* background so that transgenic offspring could be selected by wild-type *lin-15B* as the marker[156]. The transformed genetic components exist in each worm cell as a repetitive extrachromosomal array, in which multiple copies of all co-bombarded plasmids are mostly head-to-tail and randomly ligated[157]. The heritage of extrachromosomal arrays does not follow a Mendelian segregation in cell division hence shows low mitotic stability: only a small fraction (20-30%) of the offspring of an original transformant carry the transgenes and hence the population would lose all transgenes within next few generations[99, 157]. Furthermore, the repetitive sequence of extrachromosomal arrays induces strong silencing in the germline[158] and sometimes the arrays are suppressed in somatic tissues or ectopically expressed[133]. Altogether these factors increase the variability of transgene expression in different cells and among the different worms among the progeny.

To improve transmission stability of the extrachromosomal transgenes, a culture strategy as the follows has been adopted[159]: synthetase PylRS was constructed into a new vector with gene HygR encoding a phosphotranspherase that inactivates eukaryotic antibiotic hygromycin B (hygB) by phosphorylation[160]. Meanwhile, hygB solution was added to nematode growth medium (NGM) plates for selecting and maintaining the transgenic population. Transgenic strains were therefore kept under the persistent pressure of antibiotic selection.

Greiss *et al.* also reported surprisingly weak signal of the transgenic marker GFP in transformants made from the *lin-15B(n765)* background, in the absence of NCAA[99]. One underlying cause is the post-transcriptional degradation of mRNAs with a premature stop codon through an mRNA surveillance pathway called nonsense-mediated decay (NMD)[161-163]. Functional NMD mechanism of transgenic worms reduces the amount of the GFP reporter transcripts, which may lead to inefficient incorporation as well[99]. NMD is an important and conserved part of post-transcriptional mRNA surveillance of eukaryotes[161]. As protein SMG-2 is the key to this degradation mechanism in *C. elegans*, a strain with the loss-of-function allele *smg-2(e2008)* was then used[161, 164]. As was anticipated, transgenic worms with NMD deficiency hence expressed normal level of GFP[99]. And

Western blot of these transgenic worms confirmed expression of the orthogonal pair of tRNA/synthetase[70].

With the aforementioned adaption of orthogonal components to worm cells made, Greiss *et al.* demonstrated incorporation of two different NCAAs, *N*-ε-tert- butoxycarbonyl-L-lysine and *N*-ε-propargyloxycarbonyl-L-lysine into the reporter polypeptides[99]. Each NCAA was dissolved into melted NGM agar for culture plates and was then taken up by worms. After 24-48 hours of NCAA feeding, transgenic worms of several strains began to show observable red fluorescence of the incorporated GFP::amber::mCherry fusion protein, whose NCAA-dependent presence was also validated by Western blot.

Incorporation of NCAAs via amber suppression was also attained by Parrish *et al.* in worms with orthogonal pairs $tRNA_{CUA}$/LeuRS and $tRNA_{CUA}$/TyrRS[165]. In addition to the mentioned means of adaptions to a multicellular host (e.g. using external Pol III promoter to express orthogonal tRNAs and silencing NMD pathway), Parrish *et al.* increased the copy number of tRNA genes in transformed plasmid and integrated the orthogonal pair transgenes into the worm genome[165]. Resultant transgenic strains show invariantly high NCAA incorporation and can be stably maintained[165].

Benefiting from the above optimisation work, genetic code expansion has been recently implemented in developing a genetic editing tool for *C. elegans*[166, 167]. Through the incorporation of a photo-caged NCAA into Cre active site, Davis *et al.* developed a light-activateable version of the DNA recombinase for worm research[167-169]. The engineered Cre recombinase can be selectively switched on in individual worm cells by light illumination, which was shown as expression of an RFP-tagged optogenetic channel in a single neuron of living worms, so that the effects were displayed as both red fluorescence within the cell and the corresponding behaviour change of the animal[169]. This study further dissected a neural circuit and the related behaviour at single-cell level, by cell-specifically expressing channelrhodopsins in individual neurons[169-171].

In brief, *C. elegans* has been proved a good experimental platform for further improvements of genetic code expansion in multicellular organisms, and resultant implements of NCAA incorporation in return promotes the invention of advanced tools for studying worm biological questions with unprecedented precision and controllability.

1.2.3 Development of cell ablation tools in *C. elegans*

In order to determine the function of a cell of interest within an organism, a widely adopted approach is ablation, where the cell is killed and subsequently observed phenotypes are recorded for analysis. The essence of ablation research is deducing the subject function by chronologically comparing the respective systematic states in the presence and absence of the subject whilst keeping the irrelevant variables unchanged. It has been extensively used in various research fields such as neuroanatomy and genetics, and *C. elegans* study not long after the first laboratory cultivation of the species[103, 172]. Success of this tactic in *C. elegans* studies is largely attributed to the non-renewability of worm somatic cells and also the persistent engagement of researchers in improving the precision and controllability of ablation methods[110-113].

The first and still common ablation approach for worms is to target cells by high-powered laser microbeam[172]. In operation of the method, pulses from laser generator mounted on a microscope are used to hit the nucleus of a target cell in anaesthetised L1 larvae. Two types of laser, with different properties, are adopted for ablation[172]: Nanosecond-pulses were the earliest and later on the femtosecond laser was applied in some cases[173].

Different pulse durations lead to different killing mechanisms[174]. Nanosecond irradiation generates a seed free electron that starts an ionisation avalanche to form a plasma[175]. Then the plasma emits shock waves whose cavitation effect is strong enough to rupture the cytoplasmic membrane of the target to cause cell death[175-177]. On the contrary, when using femtosecond pulses, far fewer free electrons are created from multi-photon ionisation and then these low-energy electrons resonate with and break the chemical bonds of biomolecules[174].

Methods of another technical route are collectively termed genetic ablation, namely the cell-specific expression of a cytotoxic or pro-apoptotic gene. For example, ectopically expressing a gain-of-function allele of degenerins, like *mec-4*(*u231*) and *deg-1*(*gf*) mediates a degeneration of neurons as well as necrotic deaths of hypodermal and muscle cells[178, 179]. Over-expression of a caspase protein like CED-3 can also ablate target cells[180]. Compared with degenerins, the caspase renders the elimination process more rapid with less damage to the surroundings.

However, such methods are highly dependent on the promoter selected and this brings three major disadvantages: 1) only 44 neuronal groups[181] have their unique promoters and only 46 neuronal groups can be specifically

labelled even with two-promoter combination[180], let alone to selectively pick individual neuron or cell of other tissue types; 2) the temporal control of removing target neurons is in most cases not feasible, as it is limited by the intrinsic timing of the chosen promoter(s); 3) when using a relatively weak promoter, the experimental outcome may be functional impairments rather than complete elimination of a target, allowing some cellular activities to remain[179].

In order to achieve temporal control over the initiation of cell death, photo-sensitisers such as KillerRed and MiniSOG have been developed for *in-vivo* application[182, 183]. MiniSOG and KillerRed are fluorescent proteins that produce reactive oxygen species (ROS) upon suitable illumination[184] and this form of light-inducible cytotoxicity is therefore called optogenetic ablation. Photo-sensitisers were initially invented for specifically destroying proteins of interest within cells, which can be attained with fewer radicals than the amount needed for ablating a cell[185]. Upon absorbing excitation photos, a ground-state photo-sensitiser gets excited into a singlet state with high energy, and then the latter immediately converts into a triplet state. Subsequent photochemical reactions of the metastable triplet-state sensitiser with oxygen (as a receptor of electron or energy) lead to ROS production[186]. Two main approaches have been explored for improving the killing efficiency

of optogenetic ablation methods: optimising photo-sensitiser's subcellular localisation and increasing photo-sensitiser's quantum yield.

KillerRed was built on the basis of anm2CP, a chromoprotein discovered in a phototoxicity screen of GFP homologues in hydrozoan *Anthomedusa*[185]. Cytosolic expression of the photo-sensitiser can kill human 293T cells upon green light illumination. Localising KillerRed to mitochondria (denoted as mito-KillerRed) substantially increases the cell-killing rate and dramatically reduces the required illumination intensity. Under irradiation of wavelength between 540-580 nm, the chromophore of KillerRed mainly generates superoxide anion radicals ($O_2^{\bullet -}$) that can be subsequently converted into hydrogen peroxide (H_2O_2) and O_2 by superoxide dismutase (SOD), through an electron transfer mediated photo-sensitising mechanism[187].

Subsequently, the more efficient mito-KillerRed was tested by Shibuya *et al.* in *C. elegans*[183]. Worm Body-wall muscle that expressed mito-KillerRed underwent non-apoptotic death after 1-hr green light irradiation and such ablation effects can be attained in worms at various developmental stages[183]. However, mitochondrial expression of the photo-sensitiser alone was already toxic, manifested as thinner muscular mitochondria and thinner mitochondrial tubules, higher stress response of the muscle, retarded

development and therefore a reduced body size and impaired locomotion of the transgenic worms[183]. Such deficits did not appear in the transgenic worms expressing the same mito-KillerRed at a lower level and after short irradiation only mitochondrial fragmentations rather than muscle deaths happened[188]. These two pieces of evidence suggest a potential toxicity of KillerRed overexpression.

To solve the basal toxicity of KillerRed, a more recent work by Kobayashi *et al.* localised the protein onto plasma membrane by a MLS (Membrane Localised Signal)[189]. By expressing the MLS-KillerRed in neuron AWA with a cell-specific promoter, Kobayashi *et al*. ablated AWA without damaging a neighbouring neuron AWC within the illumination area[189]. Besides, no pre-illumination deficit was reported[189]. Although its application is in a different cell type from previous studies (neuronal vs. muscular) for a direct comparison, MLS-KillerRed is still a practical KillerRed version in the toolkit of optogenetic ablation. Indeed, KillerRed efficiency varies in ablating different targets, even different cells of the same neuronal class[188]. The underlying cause could be the various efficiencies of SOD in its conversion of superoxide to hydrogen peroxide in different cells[188, 190]. Therefore, silencing this ROS-detoxifying pathway could be considered for future improvement on the optogenetic ablation[188, 190].

The intrinsic dimerisation tendency of KillerRed was found a drawback in its usage in specifically destroying target protein[185]. A monomeric version of KillerRed called SuperNova has therefore been developed[191]. The new photo-sensitiser has its fluorescence excitation maximum at 579 nm (orange), a higher quantum yield of singlet oxygen and less superoxide production than the original KillerRed[191]. In the absence of light, SuperNova expression in worm neurons did not bring about the cellular damage accompanied with cytosolic KillerRed expression in the same type of cells[191]. Practicing optogenetic ablation via SuperNova is simpler too, as very short duration of illumination is needed. The disadvantage of the method, however, is that its killing effects was obtained by strong irradiation of 1 W/cm^2 (rather than mW/cm^2) magnitude[191]. Hence this SuperNova version need further efficiency improvement and is currently more of a basis for better optogenetic ablation tools of the KillerRed series.

In contrast to KillerRed and its variants, developing MiniSOG (Mini Singlet Oxygen Generator) from *Arabidopsis* Phototropin 2 was intended for sample labelling for both light and electron microscopy[192]. The photochemical reaction of the flavin mononucleotide (FMN) chromophore of MiniSOG upon blue irradiation generates a mixture of singlet oxygen (via Type II mechanism) as well as superoxide anion (by Type I reaction)[192, 193]. Most molecules

including FMN stably exist as singlet ground state, in which all electrons are in pair[194]. When absorbing irradiation, a ground-state molecule may form an excited state, either singlet (all electrons are in pairs and have opposite spins) or triplet (the excited electron spins in parallel to the ground-state electron)[195]. The most common allotrope of oxygen (O_2 or dioxygen), however, is in a triplet ground state (written as $^3O_2$), where oxygen atoms each have a single electron and the two single electrons are unpaired (in different spins) in the molecule[194]. Upon excitation, one single electron of the two oxygen atom transits to a higher energy level with its spin unchanged, so is still paired with the other single electron. This high-energy state is called singlet oxygen ($^1O_2$)[196].

As illustrated in Figure 1.5A, upon light excitation, ground-state FMN is first converted to the high-energy singlet excited state (^{1}FMN) and then to a triplet excited state (^{3}FMN) through intersystem crossing[197]. Via an intramolecular electron transfer, the excited flavin of ^{3}FMN becomes a radical anion ($FMN^{\bullet-}$) that subsequently transfers an electron to a molecule of $^3O_2$ to produce a superoxide molecule ($O_2^{\bullet-}$) while restores FMN to its ground state[197]. This reaction pathway is regarded as Type I mechanism of ROS generation. A competitive mechanism (or Type II reaction) for ground-state FMN

regeneration from its triplet state, is a direct transfer of energy to $^3O_2$, which yields singlet oxygen $^1O_2$ as the reactive species[197].

As one of reactive oxygen species, singlet oxygen is toxic from its reactivity with a range of biomolecules, including nucleic acids, lipids and proteins[198]. For instance, singlet oxygen can oxidise a tryptophan residue in a protein (Figure 1.5B)[199, 200]: first, $^1O_2$ reacts at Trp C3 to form an intermediate with either a hydroperoxide or a dioxetane across the C2-C3 double bond; the intermediate subsequent decomposes to *N*-formylkynurenine upon a breakage between C2 and C3 bond, and *N*-formylkynurenine can undergo further hydrolysis to produce kynurenine. The oxidation process *per se* degrades structure of the protein at its tryptophan residue and its main products, *N*-formylkynurenine and kynurenine, are sensitive to photo-excitation and can generate more reactive species to transfer oxidant damage to other proteins[199, 200].

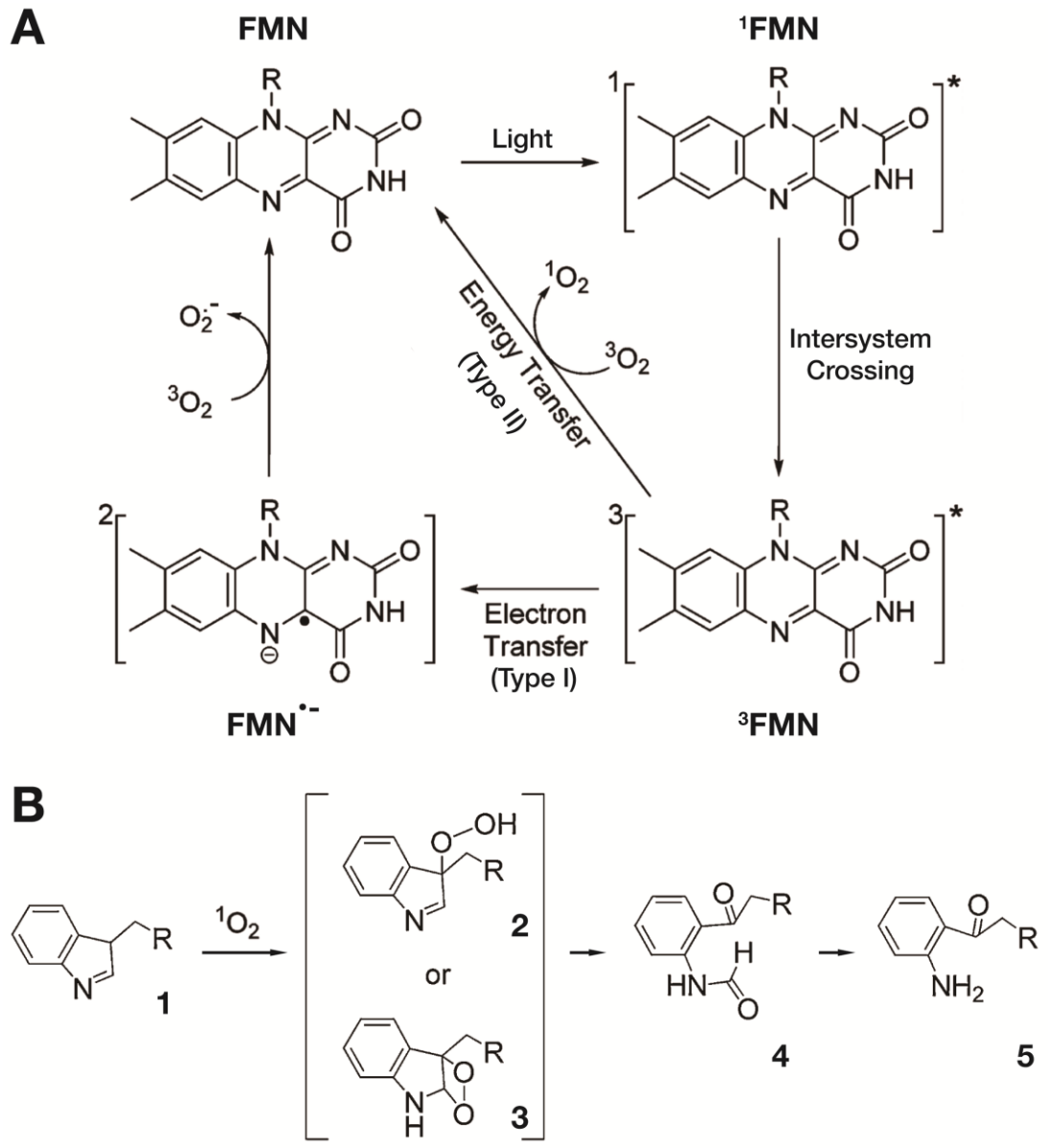

Figure 1.5 Mechanisms of singlet oxygen generation from MiniSOG and typical reaction of singlet oxygen with tryptophan.

(A) Adapted from Westberg *et al.*[197]. Scheme of the generation of reactive oxygen species such as singlet oxygen and superoxide upon FMN absorbing excitation photos. Light illumination excites ground-state FMN to a high-energy single electronic state (^{1}FMN) and then the latter nonradiatively converts to triplet state

(^{3}FMN). The regeneration of neural FMN in ground state from ^{3}FMN has two hypothetic pathways: in Type I reaction, single electron transfer of ^{3}FMN generates a radical anion ($FMN^{\bullet-}$) that transfers an electron to $^3O_2$ to produce a superoxide ion of oxygen; in Type II reaction, the energy emitted from electronic state conversion is transferred directly to $^3O_2$ to produce singlet oxygen. $^3O_2$ denotes the oxygen in ground electronic state, while $^1O_2$ represents singlet oxygen, an excited electronic state with the lowest energy. $O_2^{\bullet-}$ is superoxide anion. R represents a ribityl-5'-phosphate moiety that does not directly associate with ROS generation. **(B)** Adapted from Pattison *et al.*[200]. Tryptophan (**1**) in oxidation by $^1O_2$ generates either a C-3 hydroperoxide (**2**) or a dioxetane (**3**) intermediate. Both form can decompose to *N*-formylkynurenine (**4**) that can be further hydrolysed to kynurenine (**5**). Both **4** and **5** are able to produce reactive species upon photo-sensitisation. R represents $CH(CO_2H)NH_2$.

In *C. elegans*, cytosolic expression of MiniSOG has exhibited no cytotoxicity in the absence of its activating blue illumination, while the singlet oxygen produced by mitochondria-localised MiniSOG (mito-MiniSOG) was sufficient to induce autonomous deaths of the irradiated neurons[182]. The process does not rely on the function of worm caspase, suggesting a non-apoptotic death mechanism that awaits more research[182]. A drawback of MiniSOG is that it requires illumination of a long time: the threshold of an effective ablation

was 30 minutes for continuous irradiation, while a pulsed lighting procedure (0.5-s light, 1.5-s dark) for 15 minutes attained twice more effective killing[182]. The motives in reducing the required illumination time include to raise experimental throughput as well as to alleviate the disruptive effects of blue irradiation, for a prolonged exposure can cause acute injury and even worm lethality[201, 202].

Mitochondrial expression of MiniSOG shows no acute cytotoxicity in absence of the inducing illumination[182], but mito-MiniSOG expression in epidermis chronically affect the host worm that significantly reduces its lifespan[203]. The impact seems to associate with mitochondria, as worm strains with cytosolic MiniSOG shows a normal lifespan in the same study[203]. With MLS-KillerRed as a reference of solving the mitotoxicity, Xu *et al.* localised MiniSOG to the plasma membrane by attaching a PH (Pleckstrin Homology) domain that binds phosphoinositides[203]. In comparison with its mitochondrial counterpart, PH-MiniSOG displayed 5-6 times the efficiency with an even lower expression level, which shortens the required irradiation time to several minutes. Such improvements were confirmed by the deaths of neurons and muscle cells with consequent physiological changes[203]. Additionally, PH-MiniSOG generated ROS was found to mediate excessive lipid peroxidation at plasma membrane[203]. Such membrane damage was a

typical cause of the ferroptosis, a type of cell death however unreported so far in *C. elegans*[204]. In the meanwhile, a review into cell fates post the action of photo-sensitiser claimed that both the presence of superoxide ions and the loss of membrane integrity could contribute to a necrotic fate of the targeted cell[205]. Therefore, future research is needed to determine which type of non-apoptotic death mechanism takes place in the action of this membrane-localised MiniSOG.

In addition to optimising subcellular position of the ROS generator, rational modification of MiniSOG structure has also been shown to improve cell ablation effects[197]. By introducing a Q103L mutation intro MiniSOG, Westberg *et al.* reduced hydrogen bonds between the flavin mononucleotide chromophore and the encasement part of MiniSOG, subsequently reducing the frequency of electron transfer from ^{3}FMN to oxygen[197]. As the two ROS production pathways of MiniSOG compete for ^{3}FMN with each other (Figure 1.5A)[193], in more ^{3}FMNs of MiniSOG (Q103L) take the energy transfer reaction; hence the protein has a higher quantum efficiency of singlet oxygen generation[197]. Plasma membranous expression of MiniSOG (Q103L) in body-wall muscle of worms was able to induce acute paralysis after blue irradiation[203]. Significant ablation effects were also seen in experiments on

cholinergic neurons and epidermis with illumination as short as 20 seconds[203].

Along with the continuous development of various ablations tools as introduced above, many research groups have implemented the tools for studying biological problems of *C. elegans* and their cell ablation practice in return propelled this biotechnological field by identifying the weaknesses of a certain method. Here I evaluated all the introduced ablation methods by spatial precision, temporal control, undesired impairment before experiments or to the unintended cells and tissues, efficiency and easiness of practice. Directly destroying cells by laser microbeam ensures high spatial resolution, but such experiments have to be conducted in L1 larval stage (e.g., little flexibility of the timing)[206]. Also, in practice it is technically challenging and time-demanding[173]. In contrast, genetic ablation is easy to perform: once the correspondent transgenic line is obtained, an adequate number of worms can be fast generated for the follow-up experiment[178-180]. However, the spatiotemporal control of genetic ablation by a promoter-driven toxin has severely limited by the innate characteristics of the promoter(s), hence it only takes effect in a minority of all worm cells[180]. Replacing the toxin by a pro-apoptotic protein switches the death course from radical-induced cell deaths to apoptosis and hence reduces damage to neighbouring tissue, as

the apoptotic contents are enclosed within cell membrane and eventually engulfed, but still this method does not have a better temporal control[180]. On the contrary, optogenetic ablation allows users to choose the time point of ablation freely within the temporal coverage of the promoter[182, 183]. Besides, conducting it only requires a corresponding transgenic line and a field illumination source instead of laser generator. However, the working mechanism of this route, ROS-induced death (possibly necrosis[205]), leads to unwanted photochemical damage in the form correspondent to the different ROS type, localisation, etc. of the generator version[183, 203].

In summary, a better ablation method should combine as many aforementioned advantages of all currently available tools. And the goal is not intangible via genetically encoding of an NCAA with the suitable optical property into an apoptosis-promoting protein.

1.3 Thesis objectives

In this thesis I have developed site-specific incorporation of a new non-canonical amino acid, photo-caged cysteine (PCCys), and have applied the technique in engineering a caspase to be a cell ablation tool with

unprecedented spatiotemporal precision.

Chapter 2 will begin as a brief review of the photo-caged NCAAs that have been previously incorporated in *C. elegans* in literature. There I will first introduce instances of side chains commonly used for photo-caged NCAAs as well as the optimisation of NCAA delivery to worms. Then I will display trials of PCCys incorporation throughout the entire worm and in specific neuronal classes, followed by the practice and discussion of improving PCCys uptake by the nematode. The results altogether incept the application of PCCys via amber suppression in the tool development described in Chapter 3.

In Chapter 3, I will first introduce the core pathway of apoptosis of *C. elegans*. Then I review past attempts of caspase engineering for apoptosis-based cell ablation that illuminates me in my design of caspase modification. I will display the assessment of my engineered caspase by PCCys-incorporation experiments and light-uncaging experiments in two classes of neuronal cells as well as muscle cells. I will also demonstrate the dependence of the apoptosis induction on the combinational treatments of PCCys incorporation and optical uncaging in tandem. The killing efficiency and single-cell precision of my ablation method will be evaluated, along with a discussion of

its advantages versus existent cell-killing tools in *C. elegans*. Improvements on the orthogonality of the current code expansion system are also discussed.

The experimental material and methods of the result parts of Chapter 2 and Chapter 3 will be presented in Chapter 4 and then in Chapter 5 I will briefly introduce some future research directions of this work present in this thesis.

Chapter 2: Photo-caged cysteine (PCCys) incorporation in *C. elegans*

2.1 Introduction

2.1.1 Photo-caged amino acids (PCAA)

Photo-caged amino acids are usually derivatives of canonical amino acids with a photolabile “caging” moiety[207, 208]. The moiety that is able to mask (i.e. "cage”) the natural part of the Cα side chain and can be removed upon illumination of a specific wavelength range[207, 208]. The original side chain being masked temporarily loses the chemical reactivity as long as the caging group is present, though it is not actually enclosed by a “cage” moiety[167, 208]. These NCAAs are therefore termed “photo-caged”, in emphasis of the wavelength-specific photolability. When a photo-caged amino acid gets incorporated during translation, the resulting polypeptide will inherit the correspondent photochemical reactivity to the uncaging light[208]. The PCAA installation in the active site of the target protein may allow the protein to be activated precisely and rapidly from optical uncaging, which provides researchers with acute experimental perturbation on various biological processes of interest[208]. My experiments used one photo-caged amino acid, photo-caged cysteine (PCCys), of which more details will be introduced later in this section.

There are natural proteins whose activity is dependent on light of a certain wavelength; hence they can be optically controlled *in vitro* or *in vivo*. Typical examples of them are channelrhodopsin[170, 209-211] and phototropin[212-215]. Channelrhodopsin was first found in an algae species *Chlamydomonas reinhardtii* as a light-gated cation channel[209]. When bound with its co-factor all-*trans*-retinal (ATR), the variant channelrhodopsin-2 (ChR2) will rapidly react to blue light by opening to cations[209]. It has therefore been widely used as a light-controllable depolarizer of neurons[210, 211]. Phototropin was found in model plant species *Arabidopsis* as a photoreceptor required for the light-induced relocation of chloroplasts[212, 213]. Like ChR2, it also responds to blue light, but it is through the binding of its LOV (Light-Oxygen-Voltage) domain with an exogenous FMN chromophore[212]. FMN-bound LOV domain undergoes conformational changes when being illuminated[216]. A modified LOV version for protein control has a flavin covalently attached, so its transformation only requires blue light[217]. When in fusion with this LOV variant, a target protein for instance a kinase can be activatable by light[217]. As was introduced in Chapter 1.4, the singlet oxygen generation of FMN-LOV in illumination has been utilised in building MiniSOG for specifically destroying target proteins and cells[192, 214]. The optical control strategy basing on naturally occurring proteins, however, has its versatility limited by intrinsic characteristics of the

proteins. For instance, ion influx gating of the channelrhodopsin restricts its usage to the electrophysiological activation of cells. In comparison, optical control based on the incorporation of photo-caged amino acids can be applied to virtually any type of proteins, including kinases, proteases, nucleases, polymerases, receptors and ion channels[218].

In the early 1990s, photo-caged amino acids began to be used to trigger activities of proteins *in vitro*[219-221]. Two different means of PCAA application were performed at the time: one was to add the chemical simply into the culture solution[219, 221] and the other introduced PCAA during *in vitro* protein synthesis via charged tRNAs[220]. Amino acids including glycine and glutamate constitute a subgroup of neurotransmitters used by many neurons[222, 223], presenting as a situation where pcAAs can be applied, and in this case the solution-based PCAA supplement mimicked the functioning environment of neurotransmitters (i.e., extracellular, in synaptic cleft). A photolabile derivative of glycine (with a caging group attached to the carboxy) can be directly added into the medium for cultured neurons and its photolysis upon illumination enables a targeted and rapid activation of glycine receptors and corresponding neuronal response[219, 221]. Such usage of PCAA does not involve protein synthesis and the photo-caging strategy here is to mask the carboxyl group of canonical amino acids. In contrast, Mendel

et al. built a photo-caged aspartate (namely, aspartyl ß-nitrobenzyl ester or NB-Asp) by caging the side chain of aspartate, and ligated NB-Asp to an amber suppressor tRNA, thereby incorporating the PCAA into a bacteriophage T4 lysozyme (T4L)[224] in cell-free protein translation[220, 225]. This study first demonstrated that amber suppressor tRNAs pre-charged with an NCAA can incorporate the NCAA during *in vitro* protein synthesis[220]. It was followed by the first attempt of *in-vivo* PCAA incorporation for studying the ligand-binding mechanism of nicotinic acetylcholine receptor (nAchR)[226]. Philipson *et al.* injected into *xenopus* oocytes the nAchR mRNA with a premature amber codon and the amber suppressor tRNA charged with a caged cysteine (*ortho*-nitrobenzyl cysteine, NB-Cys)[226]. It demonstrates that NB-Cys can be incorporated into the receptor proteins synthesised *in vivo* via amber suppression[226]. The incorporation methodology, however, still limited the scope of PCAA application: microinjection only supplies the cell with a limited pool of PCAA-acylated tRNAs that are not replenished while the target protein are produced in translation persistently.

Further advancement from the above methods genetically expresses the amber suppressor tRNA as well as the adapted tRNA acylation machinery (i.e., aaRS) for continuously producing the PCAA-charged tRNAs and

consequently photo-caged polypeptides, which has become a main strategy of *in vivo* genetic code expansion[92]. As has been introduced in Chapter 1.1.4, genetic code expansion enables introduce photo-caged derivatives of cysteine, lysine, serine and tyrosine into specific sites during *in-vivo* polypeptide translation, which allows optical control of proteins and correspondent biological activities of bacteria[227], eukaryotic cells[67] and even living animals[166, 167, 228].

Similar to criteria for bio-orthogonal NCAAs (in Chapter 1.1.4)[95], photo-cleavable moiety candidates need to fulfil several criteria:

1) The amino acid is not harmful and should not react with other biomolecules within the host environment.

2) The photolysis of the caged compound should be efficient with the required illumination benign to biological activities.

3) The uncaging reaction should not produce by-products that interfere with the host physiological activities.

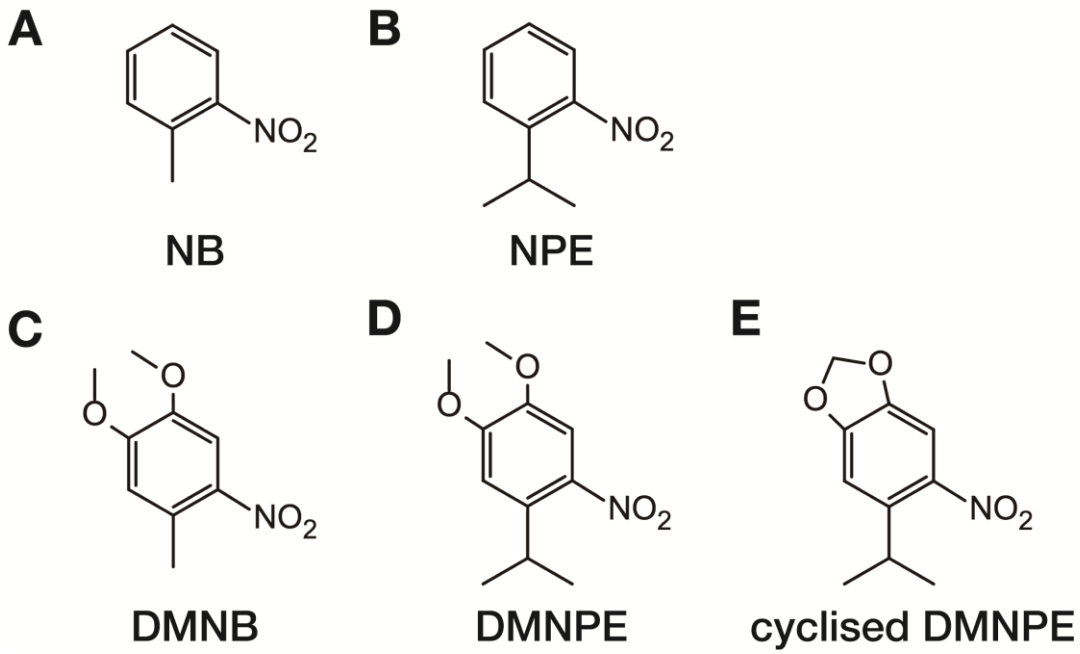

Figure 2.1 Common moieties for photo-caging amino acids.

(A) NB stands for *ortho*-nitrobenzyl, or 2-nitrobenzyl. **(B)** NPE means 1-(2-nitrophenyl)ethyl. **(C)** DMNB is for 4,5-dimethoxy-2-nitrobenzyl that has two methoxy groups added to the nitrobenzene. **(D)** DMNPE means 2-(4,5-dimethoxynitrophenyl)ethyl. Compared with NPE, it has two additional methoxy groups. **(E)** 1-{4',5'-(methylenedioxy)-2'-nitrophenyl}ethyl, in which the methoxy groups of DMNPE are cyclised. It is the caging group used in my experiments.

Figure 2.1 displays the structural formula of five moieties widely used for caging amino acids. The earliest and most common one is *ortho*-nitrobenzyl (or 2-nitrobenzyl, NB)[226], with its first application in making caged bio-molecules dates back to the 1970s: a caged-cAMP (namely, adenosine

3',5'-cyclophosphate-*o*-nitrobenzyl-‧ester‧) was developed for precisely activating cAMP-dependent protein kinases in cells[229]. A similarly designed mollecule, NB-caged cGMP, was later used to increase intracellular cGMP concentration[230]. Kaplan *et al.* used the same group for caging ATP by making a P^3-2-nitrobenzyl-ATP, in order to precisely release ATP with illumination[231]. However, the photolysis of the nitrobenzyl caged-ATP only had a low yield (< 25%) of free ATP, possibly because the photolytic byproduct (2-nitrosobenzaldehyde) reacts with the adenosine moiety of ATP[231]. There was also concern that this by-product reacts with cellular amines to generate toxic imines[232]. In search of more inert substitutions, Kaplan *et al.* found 1-(2-nitrophenyl)ethyl (NPE) a better candidate[231]. Compared with NB, NPE has an additional methyl moiety on the α-carbon. Upon illumination within a similar spectrum, this caging group presents higher quantum yield and faster kinetics in photolysis, with the generation of a less toxic product, nitrosoketone[233]..

However, the absorption maxima of both groups are shorter than 340 nm (e.g. 312 nm for nitrobenzyl cysteine)[226, 234, 235]. Ultraviolet light (UV) at such a short wavelength is detrimental to cells due to its photo-reactivity with many biomolecules like nucleic acids and disulphide bonds, so the intracellular application of uncaging favours photolabile groups by light of longer

wavelengths[236-238]. Adding two methoxy moieties to the benzene ring of NB generates a new group named 4,5-dimethoxy-2-nitrobenzyl (DMNB) with the absorption range shifted slightly to the red[231]. On receiving 350-nm light, it presents (around 10 folds) higher quantum yields in releasing cAMP than nitrobenzyl[229, 230, 239]. Similarly to the case of NB and NPE, a further increase in photolytic rate has been achieved from developing a 2-(4,5-dimethoxynitrophenyl)ethyl (DMNPE) from DMNB[233]. Also the byproduct of uncaging DMNPE, 3,4-dimethoxy-6-nitrosoacetophenone, is less reactive than that of either NB or DMNB[233].

The nitrobenzyl series of groups are available for caging cysteine residue in proteins. Examples include the incorporation of *ortho*-nitrobenzyl cysteine (NB-Cys) in human Caspase-3 expressed in yeasts[67] and the genetic encoding of DMNPE-caged cysteine (DMNPE-Cys) in engineered fluorescent proteins expressed in mammalian cells[240]. The caging group that I aimed to use for my experiments, 1-{4',5'-(methylenedioxy)-2'-nitrophenyl}ethyl, has the two methoxy moieties of DMNPE cyclised. The cysteine derivative with this caging group, *S*-[(*R*,*S*)-1-{4',5'-(methylenedioxy)-2'-nitrophenyl}ethyl]-L-cysteine (PCCys or photo-caged cysteine in the rest of the thesis), can be uncaged by 365-nm light, producing cysteine and nitrosoacetophenone. Figure 2.2 illustrated the

potential mechanism of this uncaging reaction. Both the incorporation and photolysis of PCCys have been examined in exogenously expressed proteins in human cells[241]. In the study, Nguyen *et al.* used an orthogonal pair *Mm* PCCRS/$tRNA_{CUA}$ for encoding PCCys in substitution of the catalytic cysteine of TEV (Tobacco Etch Virus) protease[241]. A TevS(TEV cleavage site)-linked CFP-YFP fusion was also pre-introduced so that a reduction in signal of FRET (Fluorescence Resonance Energy Transfer) effects implies the light-controlled restoration of TEV activity[241]. Such signal changes have been observed; hence PCCys have been incorporated into TEV and successfully uncaged. Additionally, no physiological abnormalities have been observed in these cells in all these experimental conditions, especially the 365-nm UVA illumination for PCCys uncaging, suggesting that the caged compound is physiologically compatible with the host cells[241].

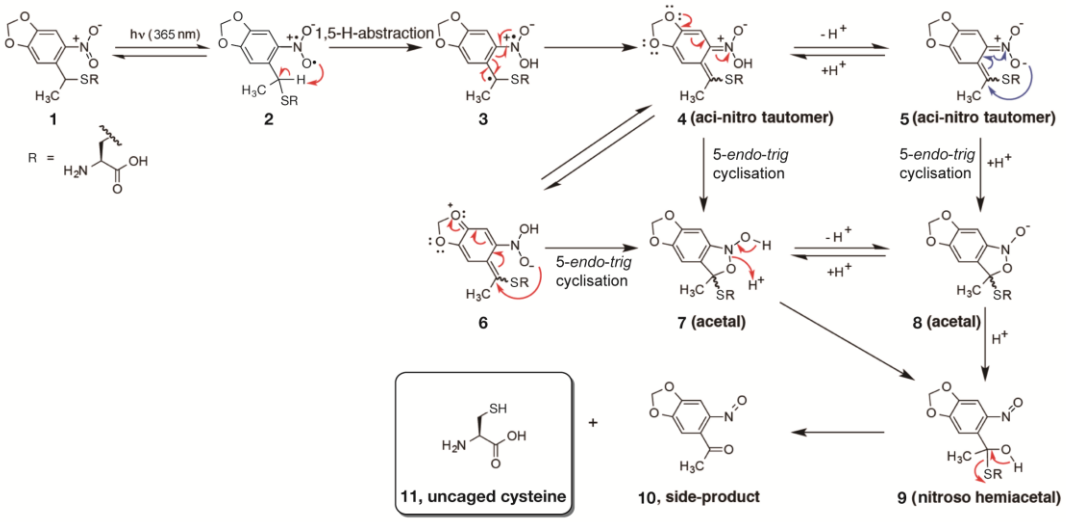

Figure 2.2 A potential mechanism of the photo-cleavage of PCCys upon 365-nm illumination.

Adapted from Radman *et al.*[167]. The photo-deprotection of PCCys **1** by 365-nm illumination generates the free cysteine and nitrosoacetophenone as a side product, which has been experimentally confirmed in aqueous biological medium at room temperature[167, 207, 232, 241]. Illumination of 365 nm light electronically excites the nitropiperonyl moiety of PCCys **1**, which initially generates intermediate **2** with fast reaction kinetics (typically in < 1 ns[207]). This intermediate **2** is unstable and forms aci-nitro tautomers **4** and **5** via intramolecular 1,5-hydrogen abstraction. The tautomeric intermediates (**4** and **5**) undergo 5-*endo*-*trig* cyclisation to generate cyclic acetal intermediates **7** and **8**, with intramolecular stabilisation gained from the resonance of their aromatic rings with oxygen atoms. Eventually these acetal intermediates have their rings open to generate nitroso hemiacetal intermediate **9**

which then breaks down to generate a side product nitrosoacetophenone **10** and the free cysteine **11**.

For these reasons, I considered PCCys a promising candidate for caging target proteins in *C. elegans*. This NCAA has not been adopted in any animal before and I set out to establish its use in *C. elegans*. Thus the next question for me to solve is how to deliver the amino acid into the cells of an animal like *C. elegans*.

2.1.2 NCAA delivery to multicellular *C. elegans*

Up to this point PCCys has not been encoded in *C. elegans*, so the conditions (e.g., genetic components, concentration) previously used for PCCys in cell culture and conditions used for other NCAAs in *C. elegans* are collectively a guide for me to establish a system for encoding PCCys. Particularly, a photo-caged lysine (later referred to as PCLys) named as *N*-ε-({(*R*,*S*)-1-[4',5'-(methylenedioxy)-2'-nitrophenyl]ethoxy}carbonyl)-L-lysine have been genetically incorporated via amber suppression in both mammalian cells and *C. elegans*[166, 167, 242]. PCLys bears structural similarities to PCCys[241, 242] and both PCLys and PCCys produce nitrosoacetophenone as the side

product upon optical uncaging[167, 241]. Not only have the incorporation and the uncaging of PCLys in target proteins worked, but also cultured cells and worms expressing the corresponding photo-caged lysine-tRNA synthetase (PCKRS)/$tRNA_{CUA}$ pair are in normal state throughout these experiments[166, 167, 242]. Such physiological compatibility of PCLys suggests that the similar encoding system and photolytic mechanism of PCCys can be tolerated by *C. elegans*.

The orthogonal pair that has been tested in human cells for tRNA acylation by PCCys, *Mm* PCCRS/$tRNA_{CUA}$[241], can be transformed into worms via either biolistic bombardment or microinjection[132-135]. However, as the codon preference differs between human and worm, I optimised the coding sequence of the synthetase accordingly before transformation experiments[243].

A major difference between delivering NCAA to cells *in vitro* and *in vivo* (more precisely, within multicellular organisms) is that the latter has the extra steps of uptake: the NCAA needs to be taken up first through skin or to be ingested and absorbed through gut, and then to be delivered to target tissues before crossing the plasma membrane of target cells. A hypothetical approach to bypass the delivery is to inject NCAA solution directly into the vicinity of target

cells. However in my project plan, this method was not the first choice in view of its corresponding acute damage to the animal, demand of time and technical difficulties[167]. And more appropriate means of PCCys delivery to worms is discussed as below.

Many previous studies have explored factors in drug uptake of *C. elegans*. Worms can be cultured either on solid NGM plates (1) or in liquid S medium (2) and the NCAA can be mixed with melted agar (a), added onto the agar surface (b), mixed with the liquid medium (c). Additionally, an NCAA whose structure highly divergent from its canonical counterparts (e.g., (Dansyl)Ala from alanine) is unable to be taken up through transmembranous amino acid transporters and is instead sequestered in the intestine of the worms upon delivery[165, 244]. When being built in a dipeptide with a canonical amino acid (e.g. Ala-(Dansyl)Ala), the NCAA becomes transportable into cells of interest[165]. When comparing choice 1 and 2 in the trials aimed at incorporating Ala-(Dansyl)Ala), worms fed on agar supplemented with the NCAA expressed more reporter proteins and grew healthier than worms of the same strain in liquid culture[165]. In contrast, a report on the incorporation of *ortho*-nitrobenzyl lysine (NB-Lys) suggests that dissolving the NCAA molecule in the liquid culture outperforms NCAA delivery by on-agar culture[245]. A subtle difference between the two on-agar cultures of the

studies is that the former group dissolved NCAA in agar to cast plates, while the latter group blended NB-Lys solution and *E. coli* suspension to feed worms in one go[165, 245].

The above observation brings a neglected factor into my sight - the interaction of NCAA with the bacteria used as food[167]. Investigation into the influence of bacterial metabolism on worm xenobiotic uptake found that the concentration of a chemical pre-mixed in NGM agar dropped much faster with living bacteria than with dead bacteria, so did the respective *in-vivo* concentrations of the chemical, regardless of solid or liquid culture of worms[246]. Such phenomena have been observed for both resveratrol (hydrophilic) and FUDR (hydrophobic)[246]. Also, adding chemicals (including resveratrol, FUDR and NCAA solution) freshly to worms on NGM plates showed no improvement[167, 246]. It has been shown that increasing the permeability of the worm cuticle can raise xenobiotic uptake of worms and subsequently improve NCAA incorporation[247]. Altogether these studies suggest that the xenobiotic absorption of NCAA can happen through the worm cuticle in addition to the intestine and that it takes longer than the possible bacterial uptake or degradation of the chemical. Following this, I decided to give PCCys to worms from early larval stages to ensure a longer period for NCAA uptake and to also test whether the detergent addition at the

degree of slight cuticle permeation would improve PCCys delivery and subsequently incorporation efficiency.

The experimental uptake results of several chemicals (including amino acid derivatives) have provided me with some clues about PCCys delivery, but I still decided to first ensure PCCys incorporation in *C. elegans* before exploring the optimal way to feed worms for higher yields of incorporated proteins. The progress is therefore present here in a step-wise way.

2.2 Results

2.2.1 Site-specific PCCys incorporation in *C. elegans* via amber codon suppression

Although in *C. elegans* the genetic encoding of photo-caged lysine has been realised[166, 167, 169], the incorporation of PCCys in multicellular organisms remains to be confirmed. Therefore, I first tried to incorporate PCCys into a fluorescent reporter throughout the worm.

2.2.1.1 Ubiquitous incorporation of PCCys

Studies succeeded in NCAA incorporation in worms adopted a similar genetic framework of their NCAA-encoding systems (Figure 2.3)[99, 167, 169] which is a reference for me to design my plasmids. The same orthogonal tRNA (pyrrolysyl $tRNA_{CUA}$) sequence with a ubiquitous P*rpr-1* promoter[165] added to its 5' (Figure 2.3B) can be directly used in my PCCys-encoding system. To strongly express the orthogonal PCCys-tRNA synthetase (PCCRS) developed by Nguyen *et al.*[241] in the expressing context of *C. elegans*, I optimised codons of the original PCCRS[243] and the optimised PCCRS sequence to replace the PCKRS coding sequence in the synthetase-expressing construct (Figure 2.3A). The original synthetase plasmid used the promoter of a ubiquitously expressed gene *rps-0* (P*rps-0*)[153]. It was later reported that the P*rps-0*-driven high expression of PCKRS correlates with a poor yield of transgenic worms[167]. Also, strong expression of transgenic genes (including HygR[160]) is a burden to worm growth, possibly leading to a selection for the progeny with low copy number of all transformed genes. Ubiquitous promoters P*sur-5* and P*wars-1* drive genes at lower expression levels[159, 248-250], so I used them in different permutations for expressing genes of PCCRS and HygR. There are totally four promoter permutations and each was tried in one experimental group (Table 2.1).

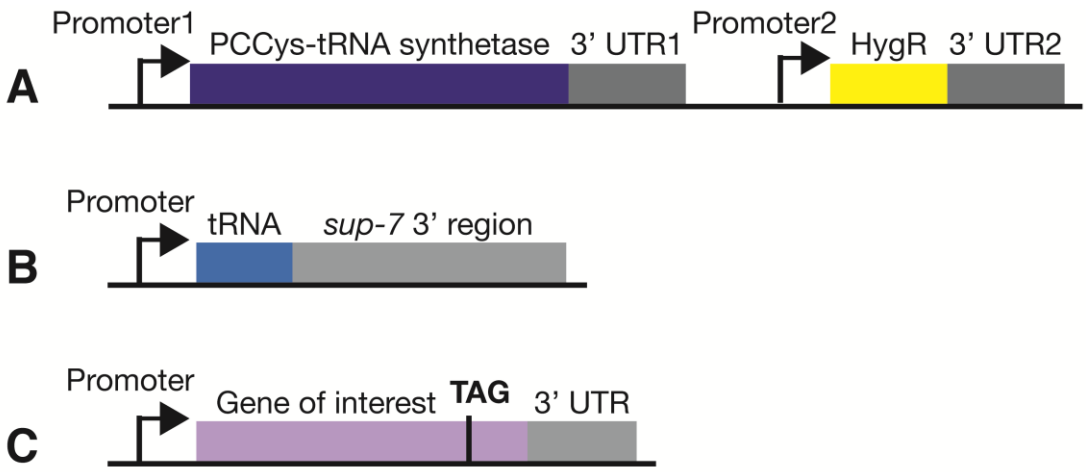

Figure 2.3 Genetic components for incorporating photo-caged cysteine at the amber stop codon in a target gene of *C. elegans*.

Adapted from Greiss *et al.* [99]. **(A)** The gene of PCCys-tRNA synthetase and the resistance gene to hygromycin B (HygR)[159, 160] are arranged in tandem on one plasmid. Promoter1 is chosen according to the cell type to target, while Promoter2 is for ubiquitous expression so the transgenic worms are resistant to hygromycin B. The latter 3' untranslated region (UTR2) is usually *unc-54* 3'UTR. **(B)** The coding construct of tRNA (*Mm* pyrrolysyl $tRNA_{CUA}$, unless indicated otherwise). The polymerase III promoter of *rpr-1*, which drives expression of the RNAse subunit of RNAseP[165], is most frequently used. And the terminator is the 3' region of *sup-7*, an amber suppressor tRNA gene[39, 151, 152, 251]. For multi-copy tRNA construct, the 3-element cassette is arranged repetitively in series for usually another 3 times. **(C)** In-frame TAG (amber stop codon) is introduced into the gene of interest to direct PCCys incorporation.

To visualise the incorporation of PCCys, I designed a fluorescent reporter in which GFP and an HA::NLS(nuclear localisation signal)-tagged mCherry are interlinked with an amber stop codon (Figure 2.4A). The HA (hemagglutinin) tag facilitates biochemical detection of this reporter protein, while the NLS of *egl-13* localises the fusion protein to cell nucleus for concentrating the fluorescent signal [167, 252]. When no PCCys gets incorporated at the amber site during translation of the reporter mRNA, protein synthesis will be terminated, yielding a truncate reporter (only GFP). Contrarily, PCCys incorporation at this amber codon will allow the translation to continue until it meets the stop codon of the mCherry part, consequently producing a GFP::PCCys::mCherry::HA fusion protein.. Promoter of *rps-0* promoter is also used for expressing this reporter[153]. The plasmid was co-transformed with the genetic components of the orthogonal PCCRS/tRNA pair into worms carrying a *smg-2(e2008)* mutant allele[161] by biolistic bombardment[134, 135]; hence transcripts of the reporter gene are unlikely to be degraded by NMD in these transgenic worms[161-163].

After bombardment, antibiotic (hygromycin B) selection and a screen for GFP fluorescence, I obtained 25 independent transgenic strains. All four experimental groups have ≥ 5 strains, and the group using plasmid P*wars-1*::PCCRS::P*sur-5*::*HygR* has the highest number of strains (Table

2.1). These transgenic worms were grown on NGM plates supplemented with 5mM PCCys and 0.1% Triton X-100 for 1-2 days and screened for red fluorescence of mCherry. Triton X-100 (or Triton) is a mild detergent and its addition has shown to increase the permeability of the *C. elegans* cuticle, followed by an improvement of the efficiency of PCLys incorporation[247].

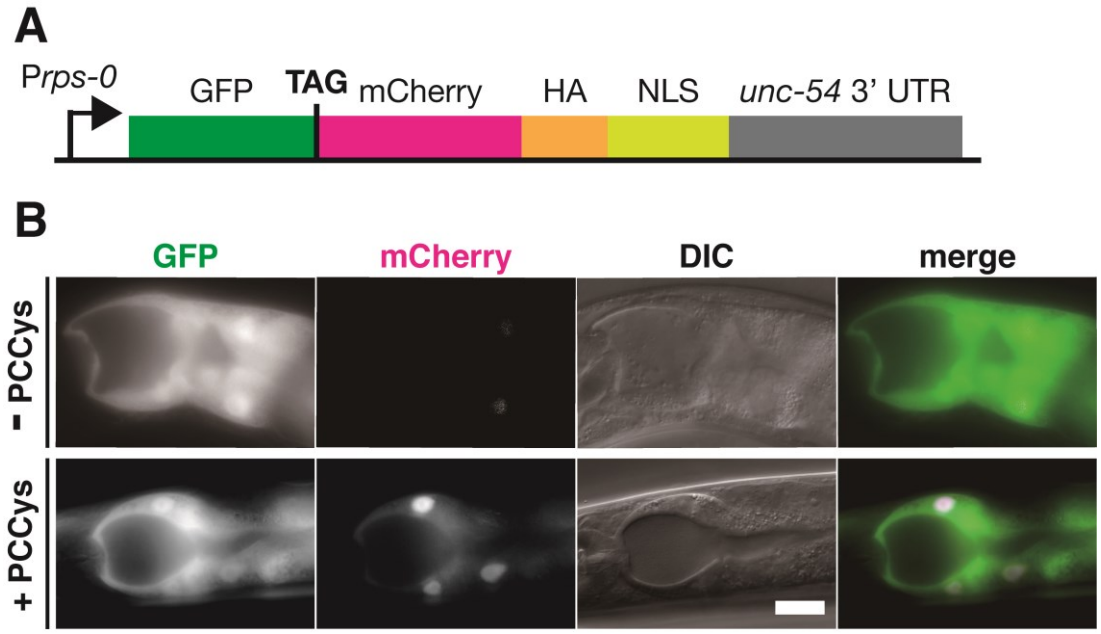

Figure 2.4 Incorporation of photo-caged cysteine site-specifically in a fluorescent reporter.

(A) Schematic of the incorporation-reporting plasmid that were co-bombarded with the constructs of the PCCys-tRNA synthetase/tRNA pair for worm strain BZ2.4.2. P*rps-0* is a ubiquitous reporter[153]. The amber stop codon TAG links the 3' of GFP and the 5' of mCherry. An HA tag is attached to the 3' of mCherry for blotting, while the nuclear localisation sequence (NLS) of *egl-13*[252] compartments mCherry

signal from that of GFP. 3' UTR of gene *unc-54* is used as the terminator and stabiliser of the mRNA[99]. **(B)** BZ2.4.2 worms show strong GFP signal (green) in the gut. No mCherry expression (magenta) is observed without PCCys (top panel), while signal of mCherry shows in the cell nucleus 48 hours after 5mM PCCys (0.1% Triton X-100) supply, indicating PCCys can be incorporated via amber suppression in *C. elegans*. For all 8 images, scale bar = 20 μm.

Worms cultured on the control NGM plates strongly express GFP in intestinal cells but no signal of mCherry can be observed. In comparison, worms grown on PCCys-NGM plates showed mCherry fluorescence localised to cell nuclei (Figure 2.4B). While red fluorescence was mainly present in the gut, it can also be observed in other tissues throughout the animal. Among all these transgenic strains, BZ2.4.2 showed the strongest signal of PCCys incorporation. So far I demonstrated by mCherry fluorescence that photo-caged cysteine could be incorporated into proteins expressed in *C. elegans* via amber suppression. This is confirmed by the results from Western blotting the lysates of BZ2.4.2 worms with anti-HA antibody (Figure 2.7 in the next section).

The different promoter permutations for driving PCCRS and HygR genes did not give an observable difference in the reporting fluorescence of transgenes

or PCCys incorporation in worms (Table 2.1). It seems that the incorporation efficiency does not differ much for P*sur-5* and P*wars-1*. PCCys incorporation efficiency can be quantified from the imaging data: first step is to calculate the integral of mCherry signal over the nuclear area of a target cell (denoted as [mCherry]); then to get the integral of GFP signal over the cytoplasmic area of the same cell (denoted as [GFP]), so the [mCherry]/[GFP] ratio represents a relative incorporation efficiency of the cell. The fluorescent images here, though, were taken from only a few sample worms and their signal ratios would have inadequate statistical power in representing the average efficiencies of PCCys incorporation of the strain.

Table 2.1 Summary of PCCys incorporation experiments of all promoter permutations designed.

"BZ no." stands for the serial number of bombardment experiments. "Incorporating strains" column shows the numbers of strains that have succeeded in incorporating PCCys. In the term R4, R stands for the red fluorescence of mCherry and 4 ranks the signal intensity on a scale of 1 to 5.

BZ No.	**PCCRS**	**HygR**	**Transgenic strains**	**Incorporating strains**	**Incorporation pattern**
2	P*wars-1*	P*sur-5*	9	1	nuclear R4, mainly in gut, some head neurons
3	P*sur-5*	P*sur-5*	6	2	nuclear R3-4, mainly in gut

4	P*wars-1*	P*wars-1*	5	1	nuclear R3-4, mainly in gut
5	P*sur-5*	P*wars-1*	5	1	nuclear R4, mainly in gut, some head neurons

2.2.1.2 Neuronal incorporation of PCCys

The previous section displays successful incorporation of PCCys throughout the body of worms. However with the ubiquitous promoters, it is hard to specifically determine in which cells PCCys got incorporated. Moreover, ubiquitous promoter P*rps-0* hardly expresses GFP in neuronal cells possibly due to the shut of *rps-0* in neurons during worm development[167] and PCLys incorporation levels vary among different tissue types of *C. elegans* (from Sebastian Greiss, personal communication). Therefore, I decided to particularly test PCCys incorporation in neuronal cells. In choosing promoters, I looked up the FPKM (Fragments per kilobase of exon model per million mapped) expression data from RNAseq in WormBase[150] for a group candidates with the below dynamic expression profiles: 1) can initiate gene transcription in a small group of neurons, of which each is anatomically discernable under a microscope; 2) altogether have various expression levels, so as to provide information about the subsequent effects on different incorporation levels; 3) stage-wise expression does not decrease significantly during worm development, especially the L4 and young adult period during

which the prospective uncaging experiments would be performed. Otherwise, the level of the protein-coding transcripts (with the premature amber stop for incorporation) would drop even more in my future trials of encoding PCCys in N2 strains with active nonsense-mediated mRNA decay. The candidates chosen finally were promoters of gene *gcy-32*, *dat-1*, *ncs-1*, *ttx-3*, *hen-1*, *ser-2*, *flp-18*, and *inx-7*[253-259].

Each of the candidate promoters was put at the 5' of both PCCRS gene and reporter gene in the constructs for each transformation experiment, so the design of all plasmids, except the plasmid of the orthogonal tRNA, is similar to the constructs for ubiquitous incorporation (Figure 2.5). Nearly half of all *C. elegans* tRNA genes are found embedded in the introns of protein-coding genes and the expression profile of the "host" Pol II promoter contributes to the tissue-specific expression of these "guest" tRNAs[260]. A dual-promoter tRNA (Pol II promoter of *glr-1* and Pol III P*rpr-1* in tandem arrangement) was therefore applied to drive tRNA expression for encoding PCLys specifically in glutamatergic neurons and resultant incorporation has been observed in all target neurons, which had not been shown when using P*rpr-1* alone (from Lloyd Davis, personal communication). Thus I built tRNA plasmids in the same way: the upstream promoter was selected from the above candidate list, followed by the global promoter P*rpr-1*.

I generated transgenic strains (termed as “reporting strain” below) for each candidate promoter by bombarding the plasmids of PCCRS, $tRNA_{CUA}$ and fluorescent reporter into *smg-2* worms.

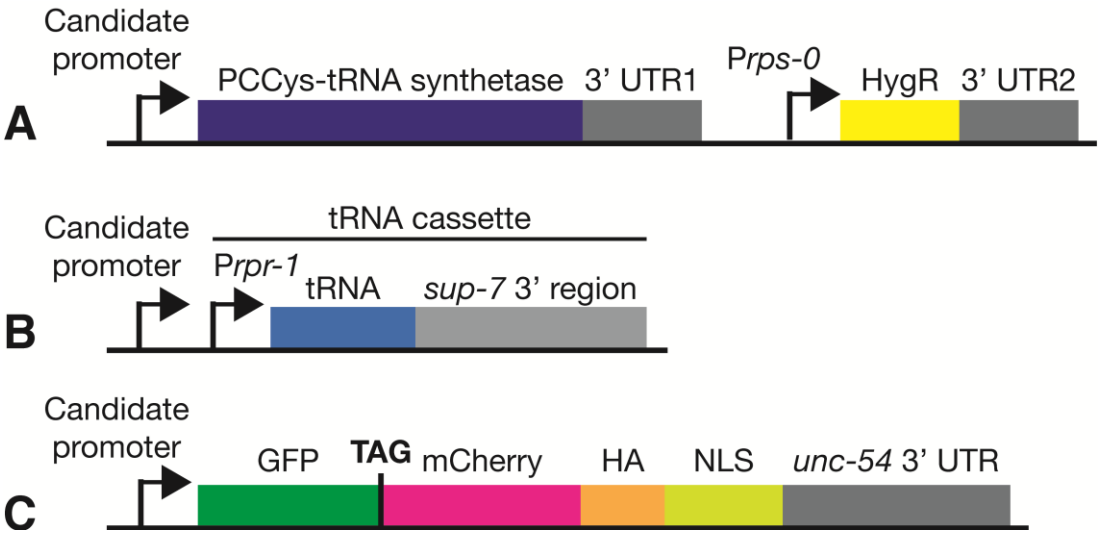

Figure 2.5 Genetic components for testing PCCys incorporation in neurons.

Adapted from Greiss *et al.*[99]. **(A)** The plasmid of PCCys-tRNA synthetase and hygromycin B resistance. **(B)** A variant of tRNA construct that contains two sequential promoters. The upstream one was added for potential tissue-specific expression[260], while the downstream one (*rpr-1* promoter) for transcribing the tRNA region. Some tRNA plasmids have five copies of the tRNA cassette that consists of an *rpr-1* promoter, a tRNA coding region and a terminator. **(C)** The gene of dual-colour reporter that is only translated as a nucleus-localised fusion protein upon PCCys incorporation. All “Candidate promoter” denotes a same candidate

selected from the list of all chosen promoters (namely, P*gcy-32*, P*dat-1*, P*ncs-1*, P*ttx-3*, P*hen-1*, P*ser-2*, P*flp-18*, and two versions of P*inx-7*[253-259]) for the three plasmids of the orthogonal PCCRS/$tRNA_{CUA}$ pair and fluorescent reporter for each reporting strain.

I cultured transgenic worms from these bombardments with PCCys supplied and then screened them for the proxy mCherry signal for PCCys incorporation. Incorporation was observed in the majority of the animals of strain BZ91.4.1 (Figure 2.6A). The strain was generated from bombardment no. 91 and has the transgenes driven by promoter of *gcy-32*. Gene *gcy-32* gets expressed (by P*gcy-32*) in only four oxygen-sensing neurons (i.e., AQR, PQR, and left and right URXs) for encoding a soluble guanylate cyclase[261]. Notably, the incorporation-reporting fluorescence appears in all and only the four O_2-sensing neurons of these worms. Then I conducted the same incorporation screens on transgenic worms of the other candidate promoters and compared the corresponding results to that of BZ91.4.1. In the reporting strain with the use of P*dat-1* and the reporting strain of P*ncs-1* usage, correspondent target neurons (dopaminergic neurons[255] and *ncs-1*-expressing neurons[254], respectively) also showed nucleus-localised red fluorescence (Figure 2.6B&C, respectively). However, the incorporation only presented in a fraction of the expected cells and only some of the

PCCys-fed worms. Also, hardly any incorporation was observed in the remaining strains. This preliminary data (summarised in Table 2.2) suggests that I should proceed to implement PCCys for protein control in the oxygen-sensing neurons with P*gcy-32*. Overall, it is very infrequent that all promoter-specific neurons get PCCys incorporation and the incorporation levels vary among the neurons of the same animal. And the FPKM data (from WormBase[150]) of a promoter only reflects the average expression level among different neurons of the driven gene. Possibly due to the heterogeneous protein expression levels of the same promoter in different cells and the variable inheritance of the transgenes in extrachromosomal arrays[132, 157], the dual-promoter arrangement is still unable to attain in all the target cells an adequate transcriptional level of the orthogonal $tRNA_{CUA}$ for an observable level (by the reporting fluorescence) of amber suppression upon PCCys supply. For the orthogonal tRNA expression in later experiments, I decided to always use mono-promoter (i.e. only P*rpr-1*) design first before experimenting with other modification means on the tRNA construct.

Table 2.2 PCCys incorporation results of different neuronal promoters.

Of the "x/y" in the column "Incorporating strains", "x" means the number of strains that showed incorporation and "y" means the number of transgenic strains from the bombardment. "+" is for an incorporation ratio well below 1/2 , "++" for approximately

50%, and “+++” for presence of incorporation in the vast majority. YA, the young adult stage of worms.

BZ no.	Promoter	Median FPKM in YA	Incorporating strains	Incorporation (of animals)	Incorporation (of cells)	Reference of promoter used
91	*gcy-32*	0.3	3/4	++	+++	[253]
92	*dat-1*	0.5	1/4	+	++	[255]
132, 133, 139	*ncs-1*	2.2	1/6	+	++	[254]
140, 141	*ttx-3*	0.1	2/7	+	+	[258]
166	*hen-1*	0.9	0/5	-	-	[258]
180	*ser-2*	0.7	0/3	-	-	[257]
182	*flp-18*	6.9	0/5	-	-	[256]
184	*inx-7 (short)*	9.5	0/4	-	-	[259]
186	*inx-7 (long)*		0/6	-	-	

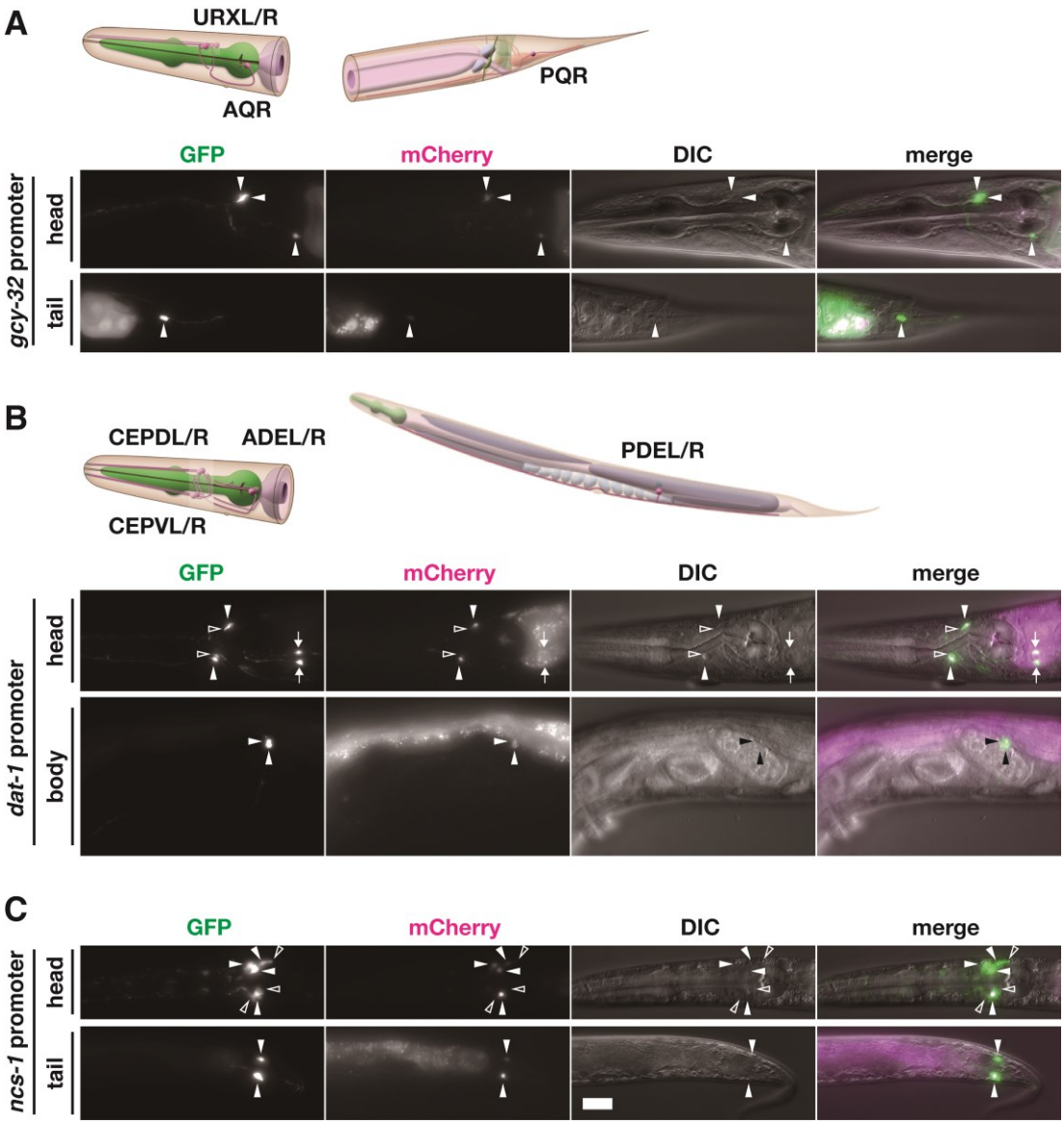

Figure 2.6 The efficiency of PCCys incorporation varies for different neuronal types.

Worms fed on PCCys-NGM plates were imaged through the microscope channels for GFP and mCherry. **(A)** The promoter of gene *gcy-32* expresses the construct in four oxygen-sensing neurons (URX pair and AQR in the head, and PQR in the tail) in worms of Strain BZ91.4.1. All of them show mCherry signal in nuclei that reflects PCCys incorporation. **(B)** The promoter of *dat-1* (encoding a dopamine:sodium

symporter[255]) drives the expression in eight dopaminergic neurons (the dorsal and ventral pairs of CEP and ADE pair in the head, and PDE pair in the mid section of the body) in worms of Strain BZ92.3.1. Half of the CEP neurons and PDE cells show mCherry fluorescence, but not the rest CEPs. Also gut auto-fluorescence[262, 263] masks possible red fluorescence of PDE. **(C)** The promoter of *ncs-1* (neuronal calcium sensor 1[254]) enables the transgenes to be expressed in multiple sensory neurons and interneurons in the nerve ring and the tail in worms of Strain BZ132.1.1. Among the cells expressing GFP, approximately half have no observable mCherry. Solid arrowheads indicate cells detected with red fluorescence, hollow arrowheads mark the neurons bearing no mCherry signal and arrows point the cells with uncertain state of incorporation. For all images, scale bar = 20 μm. Schematics of worm neurons are adapted from illustrations of WormAtlas[264].

2.2.2 Optimization of PCCys delivery

Various means of NCAA delivery to *C. elegans* have been developed and analysed by previous studies, which is displayed in Chapter 2.1.2. Notably, feeding worms PCLys in liquid culture without bacterial food as well as the use of detergent Triton X-100 (0.1%) were shown to improve PCLys uptake[167]. Meanwhile, growing worms on solid agar plates supplemented

with Ala-(Dansyl)Ala resulted in more robust incorporation of the chemical as compared to supplying it in liquid culture[165]. The reports altogether suggest that the optimal method varies for delivering different NCAA kinds. When I use water to dissolve PCCys, I found that it is not as hydrophilic as PCLys, which was also indicated by the fact that rough PCCys products from chemical synthesis were washed and desalted by H_2O and hydrophilic solvents (isopropanol and ethanol)[241]. Therefore, the best way of delivering PCLys should not be just copied to PCCys. In my optimisation experiments I focused on feeding worms on agar plates supplemented with PCCys of varying concentrations, either in the presence or absence of Triton X-100.

NGM plates containing concentrations of 0, 1.25, 2.5, 5 and 10mM were made. At the concentration of 10mM, PCCys precipitates began appearing after only a few days of storage of the newly made plates, possibly due to gradual dehydration in storage. For this reason, 10mM was the maximum concentration tested. Approximately 600 starved L1 larvae from a synchronised population of strain BZ2.4.2, the transgenic strain with the strongest fluorescent signal of ubiquitous PCCys incorporation (in Chapter 2.2.1.1), were transferred onto a plate of each concentration. A 50-µl drop of OP50 suspension from a fresh liquid culture was added after 40 min. Then the worms were left to grow at 20℃ for 3 days before being observed under

fluorescent microscope and then lysed for western blot analysis (results in Figure 2.7).

For all PCCys concentrations, the incorporation-indicating nuclear mCherry signal appeared, mainly in the worm intestine. Conversely, no mCherry signal could be observed in animals grown on NGM plates in the absence of PCCys. Incorporation levels were determined by western blot against the C-terminal HA tag of the full-length reporter protein (schematic in Figure 2.5C). The signal intensity of each protein band was quantified from the scans of exposed and developed films (Figure 2.7B). In brief, PCCys was successfully incorporated in worms. The experiment shows that in the absence of detergent, 1.25mM and 10mM of PCCys seem to have the highest incorporation (by HA signal normalised against GFP) whilst with Triton added 2.5mM and 10mM appear to be the best concentrations for encoding PCCys. Also, triton supplement is shown to approximately double the incorporation efficiency. Below is the detailed analysis of the western blot results.

All lanes on the anti-GFP blotted membrane display a band of approximately 27 kD. In the parallel protein blotting by anti-HA antibody, a band of the full-length reporter presents in all groups with PCCys supplied. The GFP signal was detected via Pierce ECL (enhanced chemiluminescent) reagent,

while the relatively low level of full-length protein requires HA detection with the use of the much more sensitive SuperSignal West Femto reagent, so the reporter bands did not manifest with the anti-GFP antibody.

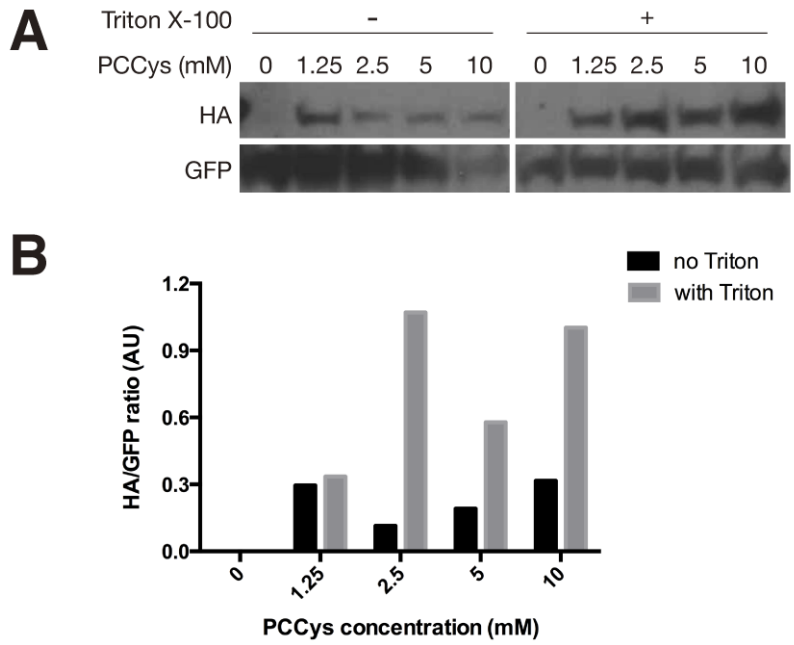

Figure 2.7 The efficiency of photo-caged cysteine incorporation varies under different conditions.

The strain BZ2.4.2 expresses the dual-colour incorporation reporter ubiquitously, especially strongly in intestinal cells. Approximately 600 worms of BZ2.4.2 were fed PCCys of various concentrations (0, 1.25, 2.5, 5 and 10mM) with or without the supplement of Triton X-100 (0.1%) as a permeabiliser. AU, arbitrary unit. **(A)** Images of western blot of GFP and HA (attached to mCherry) contained in the worm samples. The upper band is the GFP::PCCys::mCherry::HA fusion protein and the lower band is GFP. The ratios of the respective integrated grey values of the protein bands are

plotted in **B**. **(B)** The bar graph represents HA to GFP (intensity) ratios and these fractions are not to scale of the relative expression level, as the chemiluminescence of the two antibody sets were detected differently. All bar graphs in **B** are not quantitative, as no replication of the western blot experiment has been performed.

Detergent addition generally increases the amount of PCCys incorporated (shown as HA bands) and enhances the incorporation efficiency of PCCys (reflected by the HA/GFP ratio), except for the 1.25mM group. The concentration groups of 2.5mM, 5mM and 10mM exhibit similar intensity of HA bands in either condition of Triton addition. Without the presence of Triton X-100, the 1.25mM group shows higher total incorporation than the other three groups. The reason why a concentration of ≥ 2.5mM even renders less absolute incorporation is unclear to me. Inversely, with the detergent added, more PCCys molecules are incorporated in the three groups of higher concentrations. This improvement is possibly due to an increased uptake of PCCys via cuticle permeation with Triton X-100 present.

However, from 0mM group I observed decreased GFP level with Triton X-100 present. For concentrations of 1.25mM and 2.5mM, similar differences also exist between groups with/without Triton, which might be a sign of the effects of detergent. On the contrary, the samples of concentrations higher than

2.5mM have more GFP with the detergent. Since I only performed one such experiment, the inconsistent signals of GFP among different groups cannot provide me with sufficient information to make a conclusion and repeat experiments are required for further investigation.

The HA/GFP bar of each experimental group is proportional to the ratio of the PCCys incorporation frequency via amber suppression to the translation termination frequency at the amber, of the same reporter transcript. It reflects how efficient the orthogonal $tRNA_{CUA}$ gets charged by PCCRS with PCCys and suppresses the amber codon for incorporation, in competition with eRF1-mediated release of the nascent peptide chain upon amber stop codon[265, 266], to which I gives the name “relative incorporation”. There is a general trend that the relative incorporation increases as the PCCys concentration gets raised. Meanwhile, the highest relative incorporation of PCCys is achieved by 1.25mM without Triton X-100 and by 2.5mM when Triton was added. In comparison, more than 2-fold improvement from Triton X-100 addition can be observed for most concentrations (except 1.25mM). As the positive impact increases the most from 1.25mM to 2.5mM, the PCCys incorporations of the no-Triton 2.5mM, 5mM and 10mM groups are possibly limited by the PCCys uptake through feeding. I thus infer that the

incorporation capacity of the orthogonal system can be utilised with concentrated PCCys best in the presence of a cuticle permeabilsing reagent.

The 2.5mM PCCys-NGM plates supplemented with 0.1% Triton is the optimal condition of a general incorporation of the NCAA. In the meanwhile, the best PCCys concentration with no Triton falls at approximately 1.25mM. For the later experiments I decided to choose the concentrations of 2mM (in between 1.25mM and 2.5mM) and 5mM (the one used in the literature for mammalian cultured cells[241], as a reference).

In my retrospect, further experiments are required with improvements as below. How different conditions of PCCys delivery affect the protein expression that is irrelevant to amber suppression cannot be concluded from merely immunoblotting GFP, as the gene of GFP is carried in the form of extrachromosomal array whose inheritance from parent cells/worms to the daughter cells/progeny are highly mosaic[132, 157]. And the resultant GFP bands are unable to show if the same amounts of samples were actually loaded in protein electrophoresis, although I have taken extra effort in controlling this variable by synchronising all sample worms to the same developmental stage and controlling worm numbers in the protein samples of each group of Western blot experiments. Therefore, future experiments

should include the blotting of a stably expressed endogenous protein like ß-actin[267] as an additional reference for the loading step. Also, by comparing the bands of GFP and this loading control protein, I will be able to assess whether and how the expression of transgenes are influenced by different conditions of PCCys supply.

As the signals of HA and GFP have to be detected from two set of reagents of different sensitivities, HA/GFP is only proportional to but not exact number of the incorporation efficiency. A better assessment of the frequency of PCCys incorporation per transcript, requires a reference worm strain that expresses a GFP::mCherry fusion protein regardless of amber suppression. Then the incorporation frequency can be quantified by comparing (the grey value of) the band of GFP::PCCys::mCherry upon incorporation to the band of GFP::mCherry fusion generated by endogenous aaRS/tRNA pairs in protein synthesis.

2.3 Discussion

From the initial experiments I attained the incorporation of PCCys into the GFP-mCherry reporter throughout the nematodes. On this basis, by using a

neuronal promoter to drive the expression of both orthogonal aaRS and fluorescent reporter, PCCys incorporation can be confined to the correspondent group of neurons. And this has been demonstrated in three independent cases (of P*gcy-32*, P*dat-1* or P*ncs-1*). Afterwards, I varied some parameters of PCCys delivery for higher yields of PCCys-incorporated proteins. Besides NCAA concentration, the presence/absence of surfactant in the NGM plates is an evident variable to the incorporation effects. However, prolonged exposure to a non-ionic detergent such as Triton X-100 (octylphenol ethoxylate) may disturb the development and homeostasis of the animal. Indirect evidence include the toxicity of Triton to the planarian *Dugesia japonica* of a high environmental sensitivity[268], as well as the significant inhibition of NP-9/Tergitol (nonylphenol ethoxylate, a structurally analogous nonionic detergent) and metabolite NP (nonylphenol) on body length, locomotion and life span of *C. elegans*[269]. Considering the potential hazard, I would not continue to add Triton X-100 into the mixture of NGM agar and PCCys for casting plates.

When comparing the imaging results, I noticed mosaicism in both transgene expression and PCCys incorporation. The green fluorescence seems less intense in some cells than others among all targeted neurons, with the use of either ubiquitous or neuronal promoter. Secondly, the mCherry signal from

incorporation is generally weak in neuronal cells and is often observable in only some of target neurons of one worm. Heterogeneous effects of a certain promoter in different neuronal cells may also contribute to the phenomena, in addition to the aforementioned expression variability of extrachromosomal arrays. In order to successfully and efficiently ablate target cells in subsequent light-uncaging experiments, I still had to improve the current incorporation yields. One approach to better incorporating the NCAA while restricting the proteomic burden of neurons by not needing to overexpress the genetic components, is to raise the efficiency of the orthogonal pair in both tRNA acylation and mRNA translation.

The yield of the incorporated protein via amber suppression *in vivo* is inevitably limited by (1) nonsense-mediated decay, the endogenous mechanism of degrading mRNAs that contain a premature amber codon (NMD), and (2) by the amber-binding competition between orthogonal $tRNA_{CUA}$ for polypeptide elongation factors and the release factors for terminating translation[265, 270, 271]. Accordingly, there are two possibly compatible technical routes to improvement - (a) to reduce and even to abolish the activity of the internal competitor(s), and (b) to enhance the orthogonal aaRS/tRNA pair.

The strategy of the aforementioned PCCys incorporation attempts was to break Limit(1) via Route(a) in building transgenic strains from the genetic background *smg-2(e2008)*. This null allele of *smg-2* gene inactivates nonsense-mediated mRNA decay pathway, since SMG-2 protein as an ATPase-helicase plays an essential role in NMD[161]. It is also the most conserved participator in this mRNA surveillance pathway[272]. In contrast to its yeast orthologue UPF-1, worms with completely dysfunctional SMG-2 are still viable and only exhibit mild phenotypes like a swollen yet functional vulva (hermaphrodite) or bursa (male)[161]. Nevertheless, nonsense-mediated decay exerts regulation extensively on the transcriptomic homeostasis of the animal[273]. From a screen of the *C. elegans* transcriptome, Muir *et al.* have identified 680 high-confidence substrates of NMD that are transcribed mainly from pseudogenes but also from coding genes of nuclear hormone receptors and F-box A proteins[273]. Corresponding to role of NMD in mRNA quality control, most of these genes are for nucleic acid binding, mRNA splicing and mRNA metabolic processes[273]. Besides, less direct mRNA substrates of the surveillance pathway encode participators in stress response, immune and defense systems whose effects could be too subtle to manifest in the generic examinations for viability or locomotion[273].

Study by Sakaki *et al.* revealed the association between endoplasmic reticulum (ER) stress (especially, unfolded protein response/UPR) and nonsense mediated mRNA decay in worms[274]. More specifically, SMG-6 protein is localised on ER and knocking down each of *smg-1*, *smg-4* and *smg-6* genes leads to ER stress as well as higher sensitivity to the stress[274]. Although the study did not conduct *smg-2* inhibition due to a lack of the correspondent RNAi and mutant strains, the central role of SMG-2 in NMD pathway and relationship between NMD components SMG-1/4/6 and ER stress suggests me to extrapolate the known results to SMG-2[274]. It is also biologically reasonable that a pathway for degrading aberrant mRNAs eases the burden of ER in disposing unfolded or misfolded proteins, and eliminating NMD will be detrimental to worms, especially in circumstances when the secretory protein production is challenged[275, 276].

Another NMD-related physiological situation that has been mentioned in the literature is ageing. One of the characteristics of ageing is a decline in the quality control of biomolecules; consequently accumulation of misfolded proteins encoded from genes with a premature stop codon leads to proteotoxicity[277, 278]. Observations by Son *et al.* found that NMD activity generally declines as adult worms become older[278]. DAF-2 is a receptor of insulin/insulin-like growth factor 1 (IGF-1) and *daf-2* knockout worms have a

longer lifespan than the wild type[279]. NMD activity is shown enhanced in these mutant worms as well, with indispensable participation of SMG-2[278]. Overexpression of *smg-2* does not extend the lifespan of wild-type worms, but *smg-2* knockdown significantly diminishes the longevity effect of *daf-2* dysfunction[278]. Furthermore, such influence has tissue specificity, i.e., *daf-2* mutant worms with only neuronal *smg-2* inhibition of activity presents lifespan decrease[278]. And the whole-worm effect from *smg-2; daf-2* double knockdown can be rescued by neuronal expression of *smg-2*[278]. Tissue dependent regulation of nonsense decay is reflected by the phenomenon in worms that the senescent decline in NMD activity of neurons is the least among that in all types of cells during worm ageing[278], which is also confirmed in mammalian studies[280]. The experimental evidence relates with my project plan in the following two aspects: 1) higher NMD activity in neurons may at least partially contribute to the lower level of neuronal PCCys incorporation observed in the experimental strains expressing the dual-colour reporter; 2) since the genetic encoding of PCCys in my objective is meant to be a versatile tool for different research purposes (longevity study included), *smg-2(e2008)* mutant strains are not suitable for probing possible applications of the technique.

In the meanwhile, transgenic strains made from the wild type strain N2, for instance BZ132.1.1, still showed an observable level of PCCys incorporation in some neurons. This also suggests a re-consideration of using the *smg-2* knockout background for future transformation. In view of the potential side effects of knocking out *smg-2*, I decided to generate N2-based transformants in subsequent incorporating and uncaging experiments, which requires me to increase the efficiency of the orthogonal pair, i.e. taking the aforementioned Route(b).

When pursuing this enhancement goal, it is natural to focus on the overall intracellular concentration of the exogenous components for expressing the orthogonal aaRS/tRNA. In the meantime, the sub-cellular distribution of orthogonal components has been neglected. Only recently was a nuclear localisation signal in the sequence of pyrrolysyl-tRNA synthetase *Mm* PylRS identified[281]. When expressed in human cells, the protein appears mainly in the nucleus, away from where tRNAs and the translation machinery reside[281]. The low cytoplasmic concentration of the tRNA synthetase leads to insufficient acylation of tRNA, no matter how abundant orthogonal tRNAs and the NCAA supply are. By fusing a short nuclear export signal (NES) of the form alpha of human protein kinase inhibitor gene (PKIα)[282, 283] to the N-terminus of PylRS, Nikić *et al.* relocated the protein (named as NES-PylRS)

to the cytosol and increase the incorporation efficiency of Boc-Lys by 15 fold[281]. The approach can be used for modifying the orthogonal pair I used previously.

Increasing the amount of tRNA is another direction of improvement, as common arrangements such as using multiple copies of the tRNA gene expression cassette have proved effective in yield enhancement in other systems(reported in bacteria[284], yeast[285] and mammalian cells[286]). In the practice I also transformed four-copy tRNA plasmids for most of my strains and it has already contributed to the success of PCCys incorporation.

In addition, tRNA structure can also be modified for higher *in vivo* stability and higher affinity with the cognate synthetase and the host translation machinery[271, 287, 288]. Guo *et al*. generated a library of *Mj* $tRNA^{Tyr}_{CUA}$ variants by evolving their binding interface with EF-Tu and many of the variants have shown a higher rate of NCAA incorporation in *E. coli* [289]. As the tRNA-binding site of eukaryotic elongation factor eEF1 is structurally similar to EF-Tu, this tRNA optimisation strategy is likely to also enhance the efficiency of NCAA incorporation in eukaryotic cells[270, 271].

Recently, Serfling *et al.* rationally optimised the sequence of *Mm* $tRNA^{Pyl}_{CUA}$ and tested resultant candidates for incorporating *Nε*-(*tert*-butyloxycarbonyl)-L-lysine (Boc-Lys) and *Nε*-(benzyloxycarbonyl)-L-lysine (Z-Lys) in mammalian cells[290]. Two sets of candidates have been designed: Series M (M for Mutant) tRNAs retain the scaffold of *Mm* $tRNA^{Pyl}_{CUA}$ with one single base or base pair replaced by the corresponding conserved nucleotide(s) of human tRNAs[291]; Series C (C for Chimera) tRNAs have their structures bearing the similarities with both *M. mazei* $tRNA^{Pyl}_{CUA}$ and bovine mitochondrial $tRNA^{Ser}_{UGA}$[292]. The tRNAs that are orthogonal to the host endogenous aaRSs and compatible with the host cell context are selected for a second-round adaptation for encoding a specific NCAA[290]. Three directions have been taken for generating new candidates of Series M: 1) re-combining the beneficial mutations observed in the previous round, 2) arranging an A-box motif in D arm[82], and 3) reconstructing bases connecting the T arm to the acceptor stem[290]. The second way was also used in second-round improvement of Series C tRNAs. In the ncAA-adaptation round, the candidates were tested for incorporating Boc-Lys and Z-Lys. Parallel improvements with both NCAAs were observed for Series M tRNAs, while C-tRNAs all showed the highest incorporation with Boc-Lys. Particularly, M15 ($tRNA^{M15}_{CUA}$) is the best of Series M tRNAs and C15 ($tRNA^{C15}_{CUA}$) is the best of Series C tRNAs[290].

Subsequently, the intracellular concentration, amino acylation rate, and post-transcriptional modification of the tRNA variants were examined and compared with the original tRNA (labeled as PylT). The intracellular concentration of C15 is the highest, followed by M15 and finally PylT, meanwhile the acylation rate stably stays at 50% of all tRNAs prepared[290]. The running speeds of both C15 and M15 in Northern blot were higher than that of PylT, suggesting these two modified tRNAs have higher structural stability[290]. It can be due to a more favorable tertiary interaction between the D loop and the T loop, along with the stabilising role of GC pairs in the arms of the tRNA[290]. Post-transcriptional modifications e.g. switching from uridine to pseudouridine at nucleotide position 39 and 55 also help stabilise the T loop and the anticodon stem of M15 tRNA[290, 293, 294].

Afterwards, Serfling *et al.* tested C15, M15 and PylT tRNAs for their respective efficiency of NCAA incorporation into CRF1R, a G protein-coupled receptor[290]. The yield of incorporated CRF1R was the highest with M15. Then they examined whether or not further incorporation enhancement can be attained by a combined use of these tRNAs and the aforementioned improved version of synthetase, NES-PylRS[281]. NES-PylRS/$tRNA^{M15}_{CUA}$ and NES-PylRS/$tRNA^{C15}_{CUA}$ do no bring no additional improvement to the incorporation efficiencies of their respective controls (using the original

PylRS), possibly because that the efficiency boost brought by M15 or C15 has already saturated the NCAA incorporation in the protein synthesis, even with the imperfect PylRS[290]. Benefits from adding NES could be discernable when the concentration of the NCAA or the orthogonal tRNA got limited, so the combined usage helps guarantee the best yield[290].

The combination of NES-aaRS and $tRNA^{M15}_{CUA}$/$tRNA^{C15}_{CUA}$ were lately used for improving the orthogonal incorporation system of *C. elegans*[169]. This study not only reassured compatibility and orthogonality of the tRNA variants in worm cell context, but also demonstrated by proxy fluorescence and immunoblot assays that the NES-aaRS/$tRNA^{C15}_{CUA}$ pair is 50-fold more efficient in incorporating PCLys than the old system[169]. In next chapter, I am going to display and discuss the change brought to my experiments by the modified tRNAs and NES attachment to the orthogonal tRNA synthetase.

Chapter 3: Light activation of photo-caged caspase can ablate single cells

3.1 Introduction

Cysteinyl aspartate specific proteinases or caspases, as core mediators of apoptosis, can be engineered for *in-vivo* cell ablation. I will first introduce the pathway of apoptosis in *C. elegans* as well as in humans, then introduce the foundational work on engineering this protease, and finally describe how to create a caspase-based ablation tool with improved controllability and specificity by using genetic code expansion. The experimental realisation of the precise removal of several types of worm cells via light-controllable caspase tool, in living worms, will be presented in Results section.

3.1.1 Apoptotic pathway and caspase of *C. elegans*

“To be, or not to be, that is the question”[295] not only for a free-will person facing a critical dilemma, but also for cells, unconscious but fundamental units of the physiological activities of any multicellular organism. The choice “not to be”, as a genetically programmed cell fate, is of the same importance as other

cellular activities to the entire organism, as it ensures correct development, tissue homeostasis, and protection from disease[296, 297].

Vertebrate cells “committing suicide” in a controlled fashion were first described in a monograph in 1951[298]. Later on, evidence of programmed cell death in invertebrates was also found: in the pupae of the American silkmoth, the abdominal intersegmental muscles broke down and disappear at the end of metamorphosis[299]. Since then, many studies have made subcellular observation of this physiological activity and refined its morphological definition[300], so that it could be better distinguished from necrosis, an mode of cell death marked by organelle/cell swelling, cytoplasmic membrane rupture and intracellular content release[301], and thus the word “apoptosis” emerged[300]. Coined from “falling off” in Greek, “apoptosis” was firstly used to conceptualise the morphological hallmarks on the course of the controlled cell deaths i.e. the formation of cytoplasmic remnants as small round bodies and the engulfment and degradation of the cell debris by neighbouring cells, in human tissue samples of hematoxylin-eosin (HE) staining or under the electron microscope[300]. And the term “programmed cell death (PCD)” had been interchangeable with “apoptosis”, before the categorisation and characterisation of other non-passive, physiological death modes than apoptosis in the last

decade[302, 303]. Now we know that apoptosis is one essential type of programed cell death, with its non-inflammatory characteristics so far in contrast to other known death modes of a cell[301-303].

Apoptosis has been discovered in various kinds of tissues of model organisms ranging from invertebrates, nematode[304] and fruitfly[305], to mammals like mouse[306]. In 1977, Sulston and Horvitz found cells undergoing apoptosis and determined their identities in the *C. elegans* developmental lineage [304]. Afterwards, it was in *C. elegans* that pro-apoptotic proteins, CED-4 and a caspase CED-3, were first identified[307]. The mechanism of apoptosis is highly conserved through the evolution of metazoans[302] and the findings achieved in nematodes were insightful to the mechanisms of apoptosis in other metazoan species: subsequent studies of fruitfly, mouse and human found that their respective homologues to the key components of the worm apoptotic machinery[308].

Using differential interference contrast (DIC) microscopy, worm cells undergoing apoptosis can be identified by the typical morphological alterations they exhibit[113, 304]. At the beginning, blobs form at the perinuclear area of a cell. Later the nucleoplasm becomes less granular and more refractile under illumination before the nucleus gradually shrinks to

disappearance. Afterwards, the entire cell has its refractility increased to become button-like, and finally the boundary of the dying cell disappears as it gets engulfed by neighbouring cells. The microscopically observable morphological changes of apoptotic cells makes it possible to identify participating genes in apoptosis through genetic screens.

Genetic studies have showed that the aforementioned morphological features mark only one of the four major stages of apoptosis[309]. The initiatory stage for a cell to dismantle itself is fate specification, during which either external or internal stimuli are transduced into apoptogenic signals to lift the brake on this death-promoting pathway[309]. Subsequently, the cell-autonomous execution of the death program starts. During this stage, destructive proteolysis is activated throughout the cell and results in the following successive subcellular changes: chromatin condensation, DNA laddering, loss of mitochondrial membrane potential, loss of plasma-membrane phospholipid asymmetry, and detachment from the matrix[300, 310, 311]. In next stage, the dismantled cell gets phagocytosed by other cells in its neighbourhood[113, 312]. The fourth stage is the degradation of the debris of the dead cell by phagolysosome, which takes place in the cytoplasm of the engulfing cell[313].

Three rounds of apoptosis autonomously occur during the life cycle of a hermaphrodite worm[309]. The first round takes place during embryogenesis at 250-450 minutes post fertilisation and results in the removal of 113 cells. The second round happens in the L2 stage and removes a further 18 neurons. The final round is in adulthood and removes approximately half of all germ cells from the germline[309]. In addition to the developmental and physiological cell deaths, in the germline of adult worms there are also DNA damage-induced apoptosis and infection-induced apoptosis as the defence against harmful mutations and pathological bacteria[309, 314, 315]. These apoptosis events, especially their execution by caspase, are under tight and sophisticated control.

Gene *ced-3* was first identified from screening the worm mutants whose cell death had been eliminated[307]. Subsequent cloning and sequencing work found that *ced-3* encodes a cysteine protease with a significant residue similarity with human caspases[316, 317]. The translation product is an inactive precursor or zymogen (written as proCED-3) that consists of a Prodomain, a p17 Long Subunit (LS) and a p13 Short Subunit (SS)[318]. The long subunit harbours a highly conserved pentapeptide QACRG sequence which contains a catalytic cysteine residue (C358)[316, 318]. In the three-dimensional structure of proCED-3 (Figure 3.1A), this critical cysteine

resides deep in the loop L2 and is inaccessible to potential proteolytic substrates of the caspase, which underlies its inactivity[319]. Maturation of proCED-3 into active CED-3 requires cleavage at three sites, NFVDAPT, DSVDGVP and DNRDGPL, with the breakage of the respective C-terminal peptide bonds of aspartates D220, D374 and D388[318]. The cleavage is indeed reciprocal proteolysis of a proCED-3 pair (autocatalysis)[320, 321]. *In-vitro* experiment found that when two zymogenic monomers homodimerise, the L2 loop of the second molecule (denoted as L2' below) forms a loop bundle with adjacent L2 and L4 loops of the first monomer to stablise the open conformation (of the first monomer), thereby exposing the catalytic site and consequently activating the zymogen[320, 321]. However, the *in-vitro* presence of CED-3 dimers in these studies was probably due to the artificially high protein concentration for crystallisation[319, 322]. *In vivo*, these zymogens exist as monomers, and both their self-dimerisation and activation are under tight restraints, for example by CSP-2/3 [323-325] and NPP-14[326].

CSP-2 and CSP-3 have amino acid sequences highly similar to the subunit(s) of CED-3[323]. Consequent competitive binding of CSP-2 and CSP-3 with CED-3 subunit(s) reduces the chance of auto-dimerisation of CED-3 zymogens in germline cells[325] and somatic cells[324], respectively. NPP-14

is a nuclear-pore protein on the cytoplasmic site of outer nuclear membrane. By binding the Prodomain of CED-3, NPP-14 not only directly hinders the dimerisation and subsequent autocatalysis of proCED-3 monomers, but also tethers the caspase to cell nucleus[326]. These regulators are instances of an extra level of control for restricting caspase activity in the cells programmed to survive, in addition to the essential pathway of apoptosis that I will introduce as below.

The core pathway of apoptosis activation in *C. elegans* comprises EGL-1/BH3, CED-9/Bcl-2, CED-4/Apaf-1, as well as CED-3[327-331] (Figure 3.1B). It is with the facility of multimerised CED-4 molecules (apoptosome) that the autocatalysis between two CED-3 zymogens can happen[332]. However, two CED-4 molecules as in an asymmetric dimer together bind to only one molecule of CED-9 that is localised on the mitochondrial membrane. Hence, CED-4 is unable to assemble into an apoptosome and also spatially sequestered from CED-3 zymogens that are tethered to the outer nuclear membrane[322, 326, 331]. An apoptotic stimulus will promote the expression of EGL-1 (a BH3-only protein) that is an antagonist of CED-9. EGL-1 directly binds a hydrophobic surface cleft of CED-9 through its extended BH3 (Bcl-2 homology region 3) α-helix[333, 334]. The binding site for EGL-1 does not overlap with the CED-4-binding area of CED-9[334]. Not only the affinity

between CED-9 and EGL-1 is higher[335], but also this interaction induces allosteric rearrangement of CED-4 binding surface on CED-9[333]. Thus the interaction of EGL-1 and CED-9 releases CED-4 from the latter[332, 336]. Subsequently, these free CED-4 dimers translocate to the nuclear envelope with the facility of its receptor SUN-1/Matefin on the inner nuclear membrane[336, 337]. The direct binding of CED-4 and SUN-1 brings CED-4 close to CED-3 zymogens at the nuclear periphery, promoting direct interactions between CED-4 and CED-3[337]. The presence of the CED-4 apoptosome negates the inhibition on CED-3 from structural analogues CSP-2/3 and nuclear protein NPP-14, and facilitates the reciprocal activation of two proCED-3s. Therefore, the release of CED-4 from the CED-9/CED-4 complex upon EGL-1 interaction and the subsequent activation of CED-3 by CED-4 apoptosome mark the no-return point for apoptosis activation[327-331].

Four CED-4 dimers further oligomerise to make a funnel-shaped apoptosome (as an octamer in respect of the CED-4 number)[322, 331, 332]. Through the direct interaction of CED-4 (via a shallow surface pocket and CARD/caspase recruitment domain) with both L2’ loop and CARD in the Prodomain of proCED-3, the funnel-like apoptosome attracts two CED-3 zymogens (proCED-3) into the close proximity of its "hutch"[319, 330]. Particularly, the

proCED-3-CED-4 binding through the flexible L2' loops induce steric activation of catalytic C358 in the L2 of at least one proCED-3, which subsequently facilitates the two monomers to perform reciprocal auto-proteolysis of each other and to mature into their active form[319, 322, 332, 338]. This caspase specifically recognises consensus amino acid sequence DXXD (X for any amino acid) for cleaving its numerous substrates, including the proCED-3s as yet inactive[339, 340]. The initial activation of the executive caspase CED-3 from CED-4 evokes a positive feedback of cascade activation, in which more and more proCED-3s mature and cleave substrates; hence the autonomous process of cell death becomes irreversible.

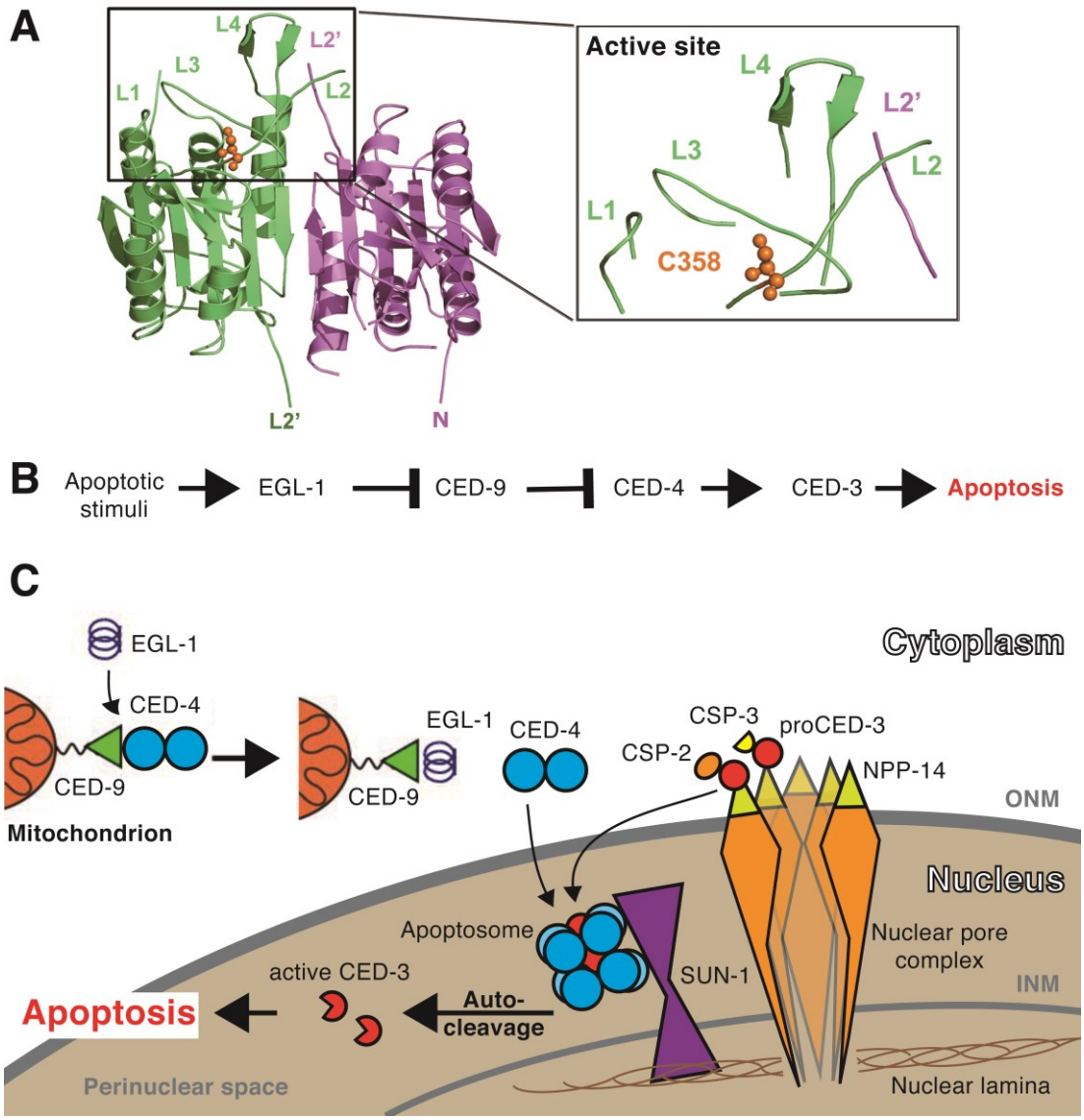

Figure 3.1 Schematic of *C. elegans* caspase CED-3 and the canonical apoptosis pathway.

(A) Structure diagram of two dimerised CED-3 zymogens in green and magenta, respectively (PDB code: 4M9R). Adapted from Huang *et al.*[319]. The four loops (L1-L4) surfacing the substrate-binding groove of one CED-3 molecule are in a catalytic conformation: loop L2' of the adjacent zymogen interacts with L2 and L4 as a stabiliser, so that L2 with the catalytic cysteine (C358, shown as spheres) is

suspended at one end of the groove and ready for catalysis. **(B)** The core pathway of apoptosis in *C. elegans*. An apoptotic stimulus promotes EGL-1 to relieve the inhibition of CED-9 on CED-4, so that CED-4 can facilitate the activation of CED-3 zymogens and then active CED-3 proteins execute proteolytic steps of apoptosis[327-331]. Positive and negative interactions are displayed as "--->" and "---|", respectively. The protein-protein interaction details and potential additional regulators are depicted in (**C**). **(C)** Adapted from a figure made by Sebastian Greiss, personal communication. The anti-apoptotic protein CED-9 is localised on the mitochondrial outer membrane and it sequesters a dimer of pro-apoptotic CED-4 proteins[322, 326, 331]. The perinuclear localisation of the CED-3 zymogen (proCED-3) is due to its interaction with NPP-14 of the nuclear pore complex[326]. In addition, proCED-3 is bound and inhibited by either one of its paralogues (CSP-2 and CSP-3)[323-325]. In cells destined to die, EGL-1 inhibits CED-9 from binding CED-4[332, 333, 335, 336]. In the current model of the apoptotic pathway in the literature, free CED-4 dimers translocate to the perinuclear region where they interact with SUN-1/Matefin and each four CED-4 dimers assemble into an apoptosome[336, 337]. SUN(Sad1 and UNC-84 domain)-1 sits across the inner nuclear membrane with its inside part binding the nuclear lamina[337, 341]. CED-4 apoptosome promotes the maturation of CED-3 through the reciprocal cleavage of two proCED-3 molecules into a functional caspase[319, 322, 332, 338]. Active CED-3 then irreversibly executes the later steps of apoptosis[316, 318, 342]. There

is so far no evidence that apoptosome-facilitated CED-3 auto-cleavage take place in the perinuclear space, despite that the illustration hypothetically depicts the process in the shaded range of “Nucleus”. ONM: outer nuclear membrane. INM: inner nuclear membrane.

The above description of protein-protein interactions (illustrated as Figure 3.1C) does not encompass the full regulatory process from EGL-1 expression to CED-3 activation. For instance, alternative splicing of *ced-4* generates two proteins, CED-4S and CED-4L[343]. Whereas CED-4S promotes cell death, CED-4L inhibits apoptosis[328, 343]. Both spliced proteins are antagonised by CED-9, but it remains unclear which of them interact(s) with SUN-1[343]. The presence of SUN-1 is also required for nuclear maintenance of CED-4 and the activation of apoptosis[337]. This suggests a possibility that CED-4 dimers further multimerise in the perinuclear region to form apoptosome for facilitating proCED-3 auto-activation. However, so far there is no evidence of the exact locations of apoptosome assembly and facilitated CED-3 autocatalysis.

3.1.2 A comparison of the intrinsic apoptotic pathway in human and in *C. elegans*

Many core apoptotic components of *C. elegans* correspond to human proteins that participate in the intrinsic apoptotic pathway[297, 302, 308, 344]. EGL-1 belongs to the BH3 (Bcl-2-homology domain 3)-only superfamily[345]. CED-9 is a homologue to human protein Bcl-2 with high similarities of both amino acid sequence and protein structure[327]. CED-4 is a homologue of human Apaf-1 (apoptotic protease activating factor-1)[346]. The apoptosis executor CED-3 shares sequence and functional similarities with fruitfly caspase drICE, mouse caspase NEDD-2, and human Caspase-1 (or Interleukin-ß-Converting Enzyme or ICE), Caspase-2, Caspase-3 (also named CPP32) and Caspase-9[309, 316-318, 347]. Not only are these proteins highly conserved, but the activation process of metazoan intrinsic apoptosis is also similar[302].

There are several differences in the apoptotic pathways of humans and *C. elegans*[302](Figure 3.2). For example, mammalian Bcl-2 does not directly bind Apaf-1 to exert inhibition. Instead, it acts indirectly through inhibiting pro-apoptotic BAX (Bcl-2-associated X protein) and BAK (Bcl-2 antagonist/killer). BAX and BAK belong to the BAX family and one of their activities is to increase mitochondrial membrane permeability, which leads to the release of cytochrome c from the inner mitochondrial membrane into the

cytosol[348-352] to activate death-promoting Apaf-1[346, 353, 354]. Another function of BAX/BAK is to activate IAP (inhibitor of apoptosis) antagonists such as ARTS (apoptosis-related protein in the TGF-β signalling pathway)[355, 356] so that the latter become able to antagonise the constitutive block from X-linked inhibitor of apoptosis (XIAP) on the auto-activation of both initiator caspases and executor caspases[357, 358]. So far, no *C. elegans* analogue of this branch of the mammalian apoptotic pathway has been identified[297]. Caspase proteins are also sources of the difference: in human cells, initiator caspases like Caspase-9 are first activated and then they proteolytically activate executors Caspase-3 and 7 to advance the death cascade[359]. On the contrary, there is no differentiation between the initiator and the executor among the caspase proteins of *C. elegans* and the only caspase, CED-3, bears the similar protein sequence and the functions to both Caspase-9 and Caspase-3[316-318, 360-362].

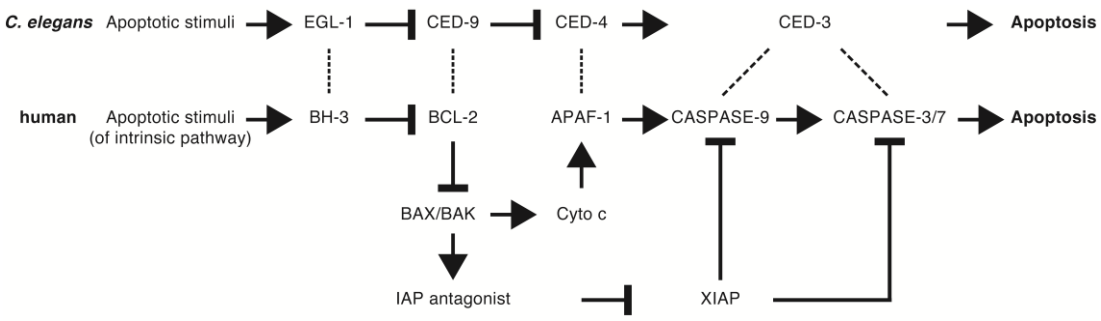

Figure 3.2 A comparison of the apoptotic pathway in *C. elegans* and in human.

Compared with the canonical pathway of *C. elegans* apoptosis, intrinsic apoptosis of human cells has additional regulatory steps and components. Without its upstream antagonist BH-3, BCL-2 (CED-9 homologue) inhibits the activity of Apaf-1 (CED-4 homologue) through two intermediate proteins, BAX/BAK and cytochrome c ("cyto c") of mitochondria[302]. BCL-2 also antagonises the function of caspases in apoptosis through XIAP, which has no known equivalent in nematodes[302]. Instead of the execution role of its homologous CED-3[361, 362], Caspase-9 initiates the signal transduction cascade of apoptotic proteolysis by cleaving other executive caspases like Caspase-3 and Caspase-7, and the latter two proteases cut a range of proteins to destroy the cell[353, 359]. IAP, inhibitor of apoptosis. XIAP, X-linked IAP. The proteins connected by dash line are homologues to each other and CED-3 is homologous to both Caspase-9 and Caspase-3[316-318, 360-362].

3.1.3 Caspase engineering for ablating cells

I contemplated engineering both human Caspase-3 and worm CED-3 for the envisaged optical initiation of apoptosis. To achieve the goal, the exogenously expressed caspase proteins should be immune to the negative regulation of *C. elegans* endogenous apoptotic pathway. Activated CED-3 is able to cleave proCED-3[318], therefore it is possible to induce the auto-activation of both

exogenous (engineered) and endogenous CED-3 zymogens for an apoptotic cascade by the manual initial activation of “seed” caspases. It requires an adequate amount of prospective optically activatable CED-3 molecules to present in the target cells and me to manually initiate (via light) CED-3 cleavage in experiments. It is similarly feasible to use human Caspase-3 as the “seed” caspase with below reasons. First, Caspase-3 and CED-3 are structurally similar and both have a QACRG catalytic sequence within their long subunits[316, 318-320]. Second, the two caspases recognise the same amino acid sequences DXXD for substrate cleavage (X denotes any amino acid of which DEXD is likely the optimal) and have many substrates in common[339, 340, 363]. For example, they both cut PARP, a nuclear enzyme substrate to be cleaved at the early onset of apoptosis, at the sequence DEVD[342, 360]. Besides, activated Caspase-3 is able to cleave a CED-3 zymogen at the desired sites (D220, D374 and D388), especially D374 in the site DSVD[318]. Caspase-3 may not directly cleave a small fraction of CED-3 substrates, due to its slightly different recognition preference from that of CED-3 (DEVD for Caspase-3 while DETD for CED-3)[363, 364]. Nevertheless, Caspase-3 is highly like to trigger CED-3-specific substrate cleavage by activating CED-3.

Caspase-3 has been widely applied for cell ablation in several model organisms. Thus far, three major technical routes have been explored for controlling its activation[365]: i) controlling the expression of an endogenous caspase, through for instance specific promoters or DNA recombinase; ii) directly activating endogenous caspases, via for example small molecules[366]; iii) engineering a caspase protein for the higher controllability[367-372]. There are many ways of the third route: 1) constructing and specifically expressing constitutively active caspase variants, e.g. with subunits reversely interlinked[373] or key residue substituted[368]; 2) fusing a self-dimerising (recruiter) domain to caspase N-terminus[367] or to each of its subunits (LS and SS) that are separately expressed[180], so monomers or split subunits can be brought together for autocatalysis to assemble a functional caspase; 3) installing a controllable element into either the subunit linker (of Way1) or the recruiter (of Way2) so to add an extra level of activity induction, for instance using a protease cleavage site[370], a light-sensitive LOV domain[371] or a chemically inducible estrogen receptor domain[369]; 4) substituting a key residue of the endogenous caspase with a photo-caged amino acid to make the protease light-controllable[67, 372]. I will go further into the technical details of the third route below, as it has demonstrated how I can engineer Caspase-3 as well as CED-3 for prospective optical control for inducing apoptosis.

It has been shown early in the history of caspase engineering that subunit rearrangement endows the enzyme with constitutive activity[373]. Srinivasula *et al.* reconstituted the DNA sequence of Caspase-3 and Caspase-6 and the resultant products (rev-Caspase-3 and rev-Caspase-6) from their N- to C-terminus are composed of Short Subunit (p12, SS), Prodomain (with an additional caspase cleavage site) and Long Subunit (p17, LS)[373]. The proteolytic efficiency of rev-Caspase-3/6 on their idiosyncratic substrates in apoptosis, PARP and Lamin A, is comparable to that of their respective native counterparts[373]. Srinivasula *et al.* also reported that rev-Caspase-3/6 are constitutively active for mediate apoptosis in two mammalian cell lines with nearly 90% killing efficiency[373]. Later, structural analysis on human caspases indicated that by reversing the order of Caspase-3 subunits, the L2 loop of Short Subunit gets extra degrees of freedom for its spatial movement and hence the subunit can freely bind and catalyse a substrate[320]. Similar result has been reported upon inverting subunits of the fruitfly homologue caspase drICE[374]. The data altogether suggests that exposing the L2 loop of caspase Short Subunit is a key to keeping a caspase active.

Further evidence came from a Caspase-3 V266E variant which was reported to have constitutive activity of proteolysis[368]. In the native Caspase-3 zymogen, residue V266 presents in an inter-subunit linker that sits at the

interface of dimerisation[364, 375]. With a V266E (valine-to-glutamate) mutation, the inter-subunit linker gets expelled from the interface, allowing stable interactions to form between the loop bundles, especially L2 and L2' loops[376, 377]. However, this caspase version is not an ideal basis of an ablation tool. When was expressed in human cells, the V266E variant displayed much lower catalytic efficiency of cleaving canonical substrates such as PARP than the wild-type Caspase-3[364]. More importantly, the mutant caspase barely cleaves endogenous caspase zymogens[364]. It means that the Caspase-3 (V266E)-mediated apoptosis relied solely on the expressed V266 variant and if I chose to photo-cage this caspase variant for cell ablation, I might have to overexpress it to obtain satisfactory killing effects.

In-vivo induction of apoptosis from engineered caspases has been demonstrated by Chelur *et al.* in *C. elegans*[180], which is an important reference to my work on Caspase-3 and CED-3. In this approach named recCaspase, the two subunits (LS and SS) of Caspase-3 are separately expressed and then brought together for caspase assembly through their linkage with a self-dimerising pair of anti-parallel leucine-zipper domains (a N-terminal NZ and a C-terminal CZ)[180, 378]. Specifically, CZ is attached to the N-terminus of LS and the NZ domain is linked to the C-terminus of SS,

which means that the SS-NZ fusion product obtains a free L2 loop at its N-terminus for constitutive proteolytic activity. When being extrachromosomally expressed in worm mechanosensory neurons, recCaspase-3 killed approximately 80% of the cells[180]. Worm caspase CED-3 that had been engineered in the same way (i.e. recCED-3) showed an ablation efficiency of above 60%[180]. The rearranged constitutively active recCaspase-3 or recCED-3 is able to induce apoptosis in the absence of CED-4[180], indicating that the exogenously expressed engineered caspase can act independently of the endogenous activation of the apoptotic pathway.

Differently from recCaspase-3 and recCED-3, hitherto photo-caged variants of caspase have only been assayed *in vitro*[67, 372]. Nevertheless, these studies have demonstrated that a photo-caged caspase can be generated via genetic code expansion and this variant can be activated with the level of spatiotemporal control so far unachievable by other caspase activation means. To endow Caspase-3 with such controllability, Wu *et al.* substituted the catalytic cysteine C163 by an *ortho*-nitrobenzyl cysteine (NB-Cys) during the protein translation of the host yeast cells[67]. Then they lysed the yeasts, UV illuminated the lysate to uncaged NB-Cys-incorporated Caspase-3 within and compared *in-vitro* proteolytic activities of the caspases before and after UV illumination[67]. As expected, the caging moiety, nitrobenzyl, prevented the

caspase from functioning, whilst uncaging of the NB-Cys residue reactivated approximately 40% of all caged enzymes[67]. In a more recent study, Wang *et al.* used the same caging group to block the access of proteolytic substrates to Caspase-3 catalytic site, so to make the caspase optically controllable[372]. They chose the residue M61 in the vicinity of the active site for incorporating an *ortho*-nitrobenzyl tyrosine (NB-Tyr)[372]. Subsequent assays on mammalian cells in which the caged-caspases expressed showed efficient apoptosis upon uncaging light[372]. However, the caged state of this Caspase-3 variant also presented significant background activity, approximately 10-20% of the cleaving capacity of the activated state[372].

With the above studies for reference, the final step of my theoretical preparation is to decide where to "install" the "light switch" of the prospective activatable caspase and which photo-caged amino acid to use. I will go into that in next section.

3.1.4 Structural and mechanistic considerations for caspase light-activation via PCCys

Previous work has demonstrated that the cysteine (C358) of the highly conserved pentapeptide sequence (QACRG) in CED-3 active site is required for CED-3 to cleave substrates[316, 379-383]. Figure 3.3A illustrates the reaction mechanism[379-383]. First, C358 first gets deprotonated by H315 and then it attacks the carbonyl carbon of the N-terminal peptide bond of the asparate residue of the caspase recognition sequence, generating an oxyanion intermediate in a thioester linkage to C358 of CED-3. Next, the oxyanion initiates a unimolecular elimination that leads to breakage of the substrate peptide bond. A water molecule that gets deprotonated by H315 attacks the carbon of the thioester bond (nucleophilic addition), from which an oxyanion forms and again undergoes elimination and the thioester bond is therefore broken. In this step, the sulfur of C358 accepts a proton from H315, so that CED-3 returns to the state before cleaving its substrate. Unsurprisingly, substitution of C358 by serine abolishes all of the cleavage activity[318]. Therefore, among the several cysteine residues of CED-3, C358 is the most suitable site for installing an “ON/OFF switch”.

For Caspase-3, the conserved catalytic cysteine C163 has the equivalent role[317, 382, 383]. Replacing C163 of Caspase-3 by glycine deactivates the

human caspase in the transfected cells[384]. Moreover, the sulfur of C163 thiol is the main target of irreversible caspase inhibitors[379, 385-387]. A major category among the inhibitors is peptide analogues made up of two parts: a tetrapeptide for caspase recognition with substrate specificity (e.g. DEVD for Caspase-3) and an electrophilic "warhead" for forming a covalent bond with the nucleophilic thiol of C163[379, 385-387] (example mechanism shown in Figure 3.3B). And the inhibition mechanism applies to CED-3 as well.

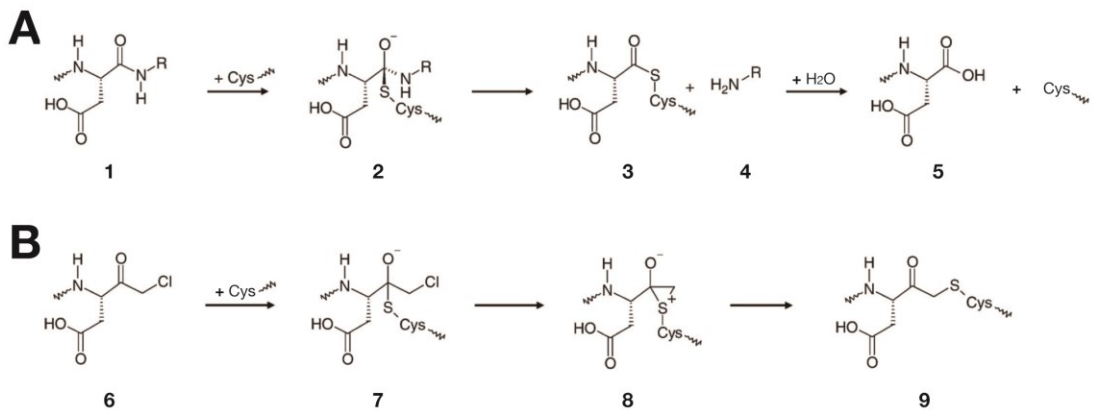

Figure 3.3 Mechanisms of proteolysis by caspase and of caspase inhibition by chloromethyl ketone as an example.

Adapted from Ganesan *et al.* [379]. **(A)** Peptide bond cleavage reaction catalysed by caspase[379-383]. In the first step, catalytic cysteine (shown as Cys) attacks the carbon of the scissile peptide bond between asparate and the other residues required for caspase recognition in the substrate protein (**1**). The resultant tetrahedral intermediate (**2**) has the substrate and enzyme linked, and its oxyanion

undergoes elimination that breaks the peptide bond, producing a thioester adduct (**3**) with the caspase and the C-terminal fragment (**4**) of the original polypeptide. The thioester bond then gets attacked by deprotonated water, which generates an oxyanion that initiates elimination to release the N-terminal fragment (**5**) of the original peptide as well as the caspase. **(B)** Postulated mechanism of the reaction between caspase and chloromethyl ketone, an inhibitor with a halogen methylketone "warhead". The thiolate anion of caspase attacks the inhibitor keto group with an intermediate (**7**) produced, and then the halide is released upon the formation of a sulfonium intermediate (**8**). The oxyanion of **8** initiates an elimination that breaks the three-membered circle and the resultant product (**9**) still has the caspase thiol alkylated.

After the prospective substitution of C163 (of Caspase-3) or C358 (of CED-3) by a caged derivative of cysteine, the caging group is highly likely to disrupt the nucleophilic attack by the sulfur atom of the catalytic cysteine, thereby inhibiting the caged caspase. This has been confirmed by Wu *et al.* that incorporation of *ortho*-nitrobenzyl cysteine at Caspase-3 Residue163 was shown to inactivate the protease *in vitro*[67]. The caging group I chose for caging caspase proteins, 1-{4',5'-(methylenedioxy)-2'-nitrophenyl}ethyl, have the following advantages over the nitrobenzyl group (also discussed in Chapter 2.1.1). First, its absorption spectrum is red-shifted compared with

that of nitrobenzyl[226, 234, 235, 241]. As light of a longer wavelength is less toxic to biomolecules and gets less attenuated by biological tissues on its path, *in-vivo* uncaging this group requires relatively weaker illumination and the process is benign to living cells[236-238, 388]. Furthermore, the group *per se* is bulkier than NB-Cys and hence more reliable for a complete deactivation of caspase catalytic site. A possible counterexample is the aforementioned (in Chapter 3.1.3) NB-Tyr caged caspases that showed residual cleaving activities[372].

In Chapter 2 I have demonstrated the first time the incorporation of *S*-[(*R*,*S*)-1-{4',5'-(methylenedioxy)-2'-nitrophenyl}ethyl]-L-cysteine (PCCys) in living animals. With this initial success, in below sections I will show how PCCys was applied for optically controlling caspase activity, for the sake of precise cell ablation.

3.2 Results

3.2.1 Design of site-specific caging of caspase

To create a photo-caged caspase for *in-vivo* use, two requirements need be met: 1) before being uncaged, the caged caspase stays inactive so the host

cells stay alive and functional; 2) upon UV illumination, Caspases with their caging group removed will mature into functional enzymes and initiate apoptosis of the illuminated cells.

Caging the critical cysteine is highly likely to deactivate caspase, which meets Condition 1 and has been backed up by previous research[67]. However, the second condition requires much engineering of the caspase. Considering that natural translation products of caspase genes are zymogens, the caged cysteine needs to be inserted into a caspase that is constitutively active, so that the enzyme becomes functional upon uncaging, without the need for further activation steps that would depend on upstream components of the apoptotic pathway. Such a constitutively active version has been created by linking subunits of Caspase-3 in a switched order (rev-Caspase-3)[373] (Figure 3.4A). The caspase variant was chosen as the foundation for introducing *in-vivo* control through photo-caged cysteine incorporation and subsequent UV uncaging.

In my design of the plasmid for expressing a light-activatable Caspase-3, a linker connects the sequences of two subunits in a reverse order (i.e., the sequence of Short Subunit is in the upstream) (Figure 3.4B). The linker consists of the proCaspase-3 Prodomain and two flanking cutting sites of the

caspase: the 5' one, translated as Asp-Glu-Val-Asp-Gly, is artificially inserted, while the 3' one naturally exists between Prodomain and Long Subunit. In the meantime, the 135th codon of Long Subunit for the cysteine critical to the caspase function (i.e., C163) has been replaced by an amber stop codon (TAG). The coding nucleotides of the caspase are optimised for the nematode species[243] and artificial introns (designed by Fire *et al.*[389], sequences in Chapter 4.1) are used instead of the endogenous introns for expression enhancement[390, 391].

In the absence of PCCys, the amber codon (TAG) stops translation of the construct, resulting in a Caspase-3 truncate that lacks the critical residues (C163 included) for its function. In contrast, successful incorporation in the presence of PCCys will allow the translation of the sequence downstream of the amber codon and the production of a full-length Caspase-3, but the caging group of PCCys still presents in the caspase catalytic pocket and blocks the proteolytic activity. As illustrated in Figure 3.4B, downstream to Caspase-3 sequence are the sequences of a F2A and a mCherry. F2A originates from foot-and-mouth disease virus (FMDV) and is able to trigger a skip of ribosome (termed as "StopGo") that releases the upstream polypeptide while resumes translating the remaining nucleotides of the same mRNA by the same translation complex[392, 393]. I used F2A in my construct

for expressing mCherry::HA::NLS as a separate peptide from the engineered caspase, thereby avoiding a possible functional interference of the otherwise covalently attachment (through a linker or not) of the fluorescent protein to the caspase. Usage of NLS (Nuclear localisation sequence) here gives two benefits: it concentrates the mCherry in the nucleus so the red signal can be better observed, and the nuclear localisation *per se* is a second readout of the mCherry fluorescence apart from a possible background[252]. Additionally, the C-terminal fusion of HA with mCherry facilitates Western blots of the fluorescent protein, reflecting consequently the presence of PCCys-incorporated Caspase-3 (or caged Caspase-3). A CFP is designed for screening transgenic strains as well as a read-out of the expression level of the entire construct. That is because CFP will be polycistronically expressed via the *trans*-splicing effect of Splice Leader 2 (SL2), with or without PCCys incorporation[394, 395].

Since *C. elegans* caspase CED-3 exhibits sequence and structural similarity to its functional equivalent, human Caspase-3, I designed the construct for CED-3 in the same way as for Caspase-3, except for three adjustments. First, it is the codon of C358 that gets substituted with the incorporation-directing amber codon; second, the original introns of CED-3 gene are replaced by artificial introns[389] and finally, the inserted cleavage site at N-terminus of

the Linker is rather DEVD (Asn-Phe-Val-Asp), specific to *C. elegans* CED-3[318].

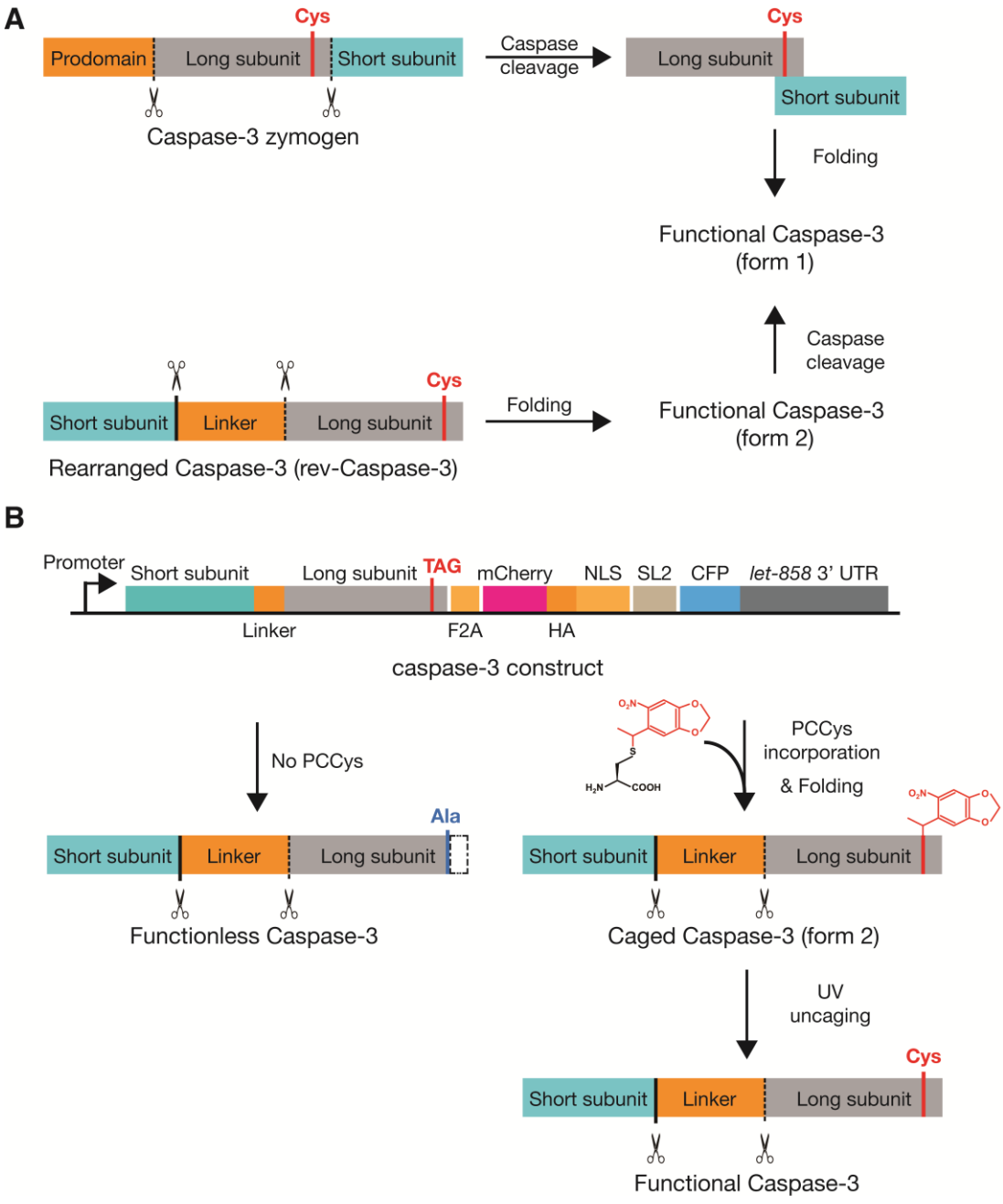

Figure 3.4 Schematic of the wild type and engineered derivatives of human Caspase-3.

All Linkers (labelled orange boxes) consist of the Prodomain region of Caspase-3 zymogen and a caspase cleavable DEVD (Asp-Glu-Val-Asp) sequence (vertical solid black lines). Canonical amino acids are presented as 3-letter codes. Important cleavage sites of caspase are marked as black lines (dotted or solid line) with "scissors" marks beside. The critical cysteine at position 163 is highlighted in red and in bold. **(A)** The polypeptide translated from human Caspase-3 exists as a zymogen. Caspase cleavage at the cutting sites removes the Prodomain and separates each subunit, enabling the product to fold to a functional structure (form 1). An engineered version of Caspase-3 has the subunits connected by a removable Linker consisting of a cleavage site and the Prodomain, in reversed order (i.e. the originally downstream Short Subunit now gets translated first). Subsequent to translation, the polypeptide folds into active caspase (form 2). The resulting caspase remains functional after being further cleaved. **(B)** Top: Schematic of the plasmid for expressing Caspase-3 in *C. elegans*. The critical C163 of the Long Subunit is replaced by amber stop codon (TAG) for PCCys incorporation. F2A denotes the 2A peptide from FMD (foot-and-mouth disease) virus that triggers a "ribosomal-skip" when being translated[396, 397], releasing the upstream polypeptide (the engineered Caspase-3) meanwhile initiating the translation of the mCherry::HA::NLS in the downstream. NLS, the nuclear localisation sequence of EGL-13[252]. SL2 or Spliced Leader 2 (from *gpd-2/gpd-3* operon) ensures simultaneous expression of both upstream and downstream genes[394, 395], so CFP is expressed in all

transgenic worms. The N-terminal part, a non-functional truncated Caspase-3 (ending with Ala162/A162 of the Long Subunit) is produced in the absence of PCCys incorporation. PCCys incorporation allows the generation of full-length Caspase-3 as well as the downstream mCherry that reports successful incorporation events by fluorescence. The caged protein stays inactive until the critical cysteine becomes unblocked upon UV illumination.

3.2.2 PCCys-mediated optical control of Caspase-3 for apoptotic ablation *in vivo*

3.2.2.1 Light activation of PCCys-caged caspase ablates neurons

As shown in Chapter 2.2.1.2, neuronal incorporation of PCCys into the GFP-mCherry reporter at the internal amber codon can be achieved for various classes of neurons. I attempted to express photo-caged Caspase-3 in oxygen-sensing neurons (AQR, PQR and URX) and touch sensory neurons (ALM, AVM, PLM and PVM). The success of cell-specific PCCys incorporation and UV uncaging in these two neuronal classes indicate the robustness of this designed tool across the nerves system of *C. elegans*.

3.2.2.1.1 Oxygen-sensing neuron

I used the promoter of *gcy-32* to specifically express the constructs of the engineered caspase and PCCRS in O_2-sensing neurons[253]. A ubiquitous promoter (either P*sur-5* or P*wars-1*) was used to express hygromycin B resistance gene throughout the nematode for the antibiotic screen of transgenic worms. Transgenic lines were constructed by biolistic bombardment of the genetic constructs into worms of Strain ZG610 that expresses GFP (from promoter of *gcy-37*) in the oxygen-sensing neurons[398]. This allows the apoptotic induction to be assayed by disappearance of the GFP-expressing neurons. As P*gcy-37* works in only these four neurons, it is also easy to count the number of the ablated neurons. Transgenic worms surviving the hygromycin B selection were further screened for CFP fluorescence under a dissecting microscope. Unlike *smg-2(e2008)* animals, ZG610 has a functional nonsense mRNA mediated decay (NMD) machinery. NMD is expected to decrease the amount of Caspase-3 mRNA due to its internal (premature) amber codon, whereas the expression of CFP, the transgenic reporter in the polycistronic mRNA, will not be affected. This is because that the SL2 splicing process happens prior to the translation-dependent recognition of premature termination codons in the NMD pathway[161-163].

To begin with, I decided to test whether the previously published subunit-reversing strategy of rev-Caspase-3[373] can be reproduced in worm caspase CED-3 to render the rev-CED-3 constitutively active. The construct designed for the engineered CED-3 has the exact arrangement as the rev-Caspase-3 of Figure 3.4A, except that the catalytic cysteine C358 is unchanged. Bombardment (no. 84) of the construct into ZG610 worms generated transgenic worms, whose adults were then checked for neuronal apoptosis through fluorescent microscope.

In the adult worms of Strain BZ84.1.2, AQR presents itself as an apoptotic body (without any dendrite) and one URX has become shapeless debris, while PQR and the other URX can be no longer observed under the microscope (Figure 3.5B). The result suggests that the rev-CED-3 of my design is capable of executing apoptosis of target cells. In comparison, ZG610 worms, the negative controls, have all their O2-sensing neurons in place with intact processes.

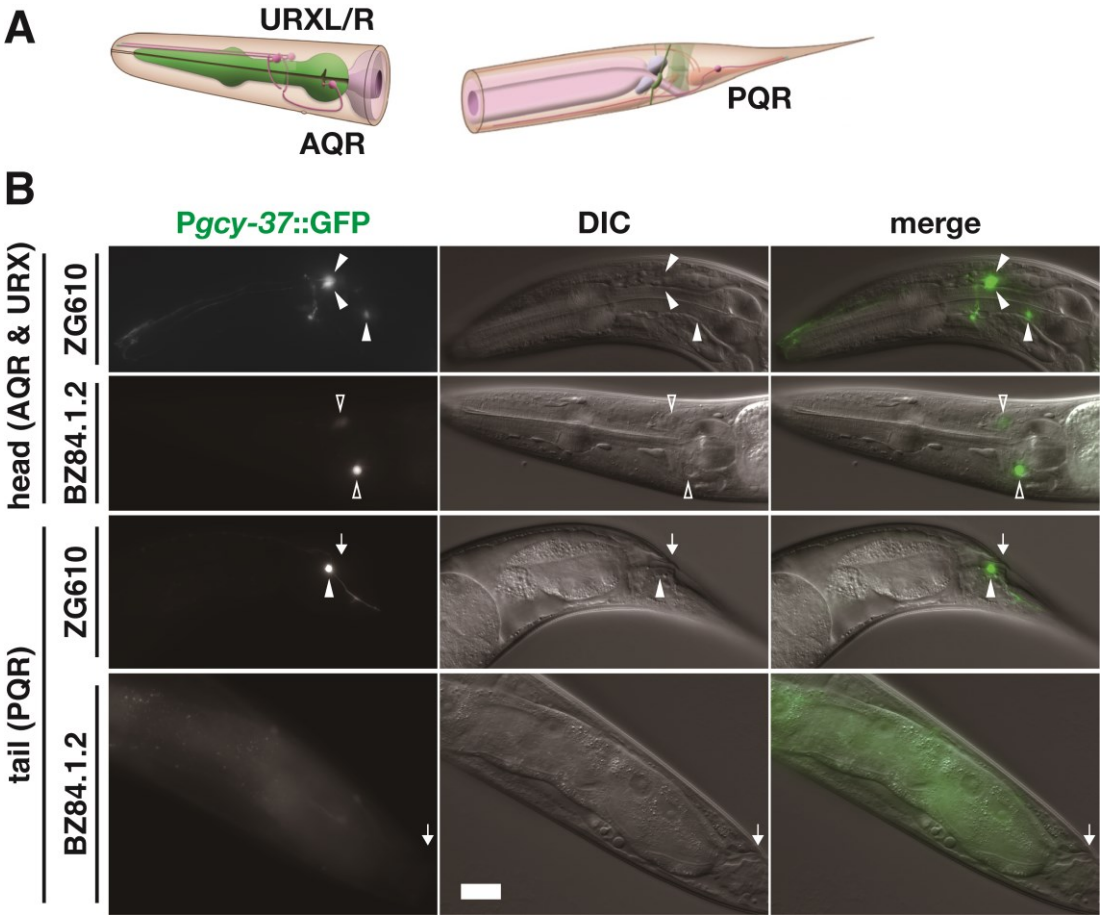

Figure 3.5 Rearranged CED-3 (rev-CED-3) can autonomously ablate oxygen-sensing neurons *in vivo*.

(A) A schematic of O_2-sensing neurons in *C. elegans*[264]. AQR and a structurally symmetric pair of URX are in the head. A characteristic AQR sits slightly posterior and ventral to the centre of the terminal bulb, while typical URX cells reside dorsal to the isthmus and at midline between the metacorpus and the terminal bulb. However, their locations in individual worms may vary (as those shown in Figure 3.5B). Among the four neurons, PQR is the only that is located in the tail, slightly posterior to the rectum (marked by arrow; the arrowheads show neurons). **(B)** Negative control worms (Strain ZG610) expressing GFP only in the four O_2-sensing neurons, as well

as experimental worms (BZ84.1.2) transgenic of rev-ced-3 construct were imaged in their early adulthood. The AQR, URXs (top line) and PQR (Line 3) of ZG610 worm all have healthy soma and processes with GFP fluorescence. Contrarily, in the caspase-expressing worms, mere the dying AQR and one URX cell with no dendrite could be detected (Line 2) and no PQR was left (bottom line). Solid arrowheads point healthy neurons with an intact cellular structure and hollow arrowheads mark the debris or corpse of neurons with ongoing apoptosis. All images follow the same 20-μm scale bar.

Then I proceeded to generate transgenic worms of PCCys-incorporable caspase (the version with amber stop codon, Figure 3.4B). CFP-positive strains were selected and then transferred onto PCCys-NGM plates to test for incorporation and uncaging-induced apoptosis. PCCys plates of 2mM and 5mM were used for assaying each strain. Among candidate strains whose worms showed cyan fluorescence in all O_2-sensing neurons, Strain BZ52.2.8 had the highest penetrance as well as a high transmission rate of progeny, and chose this strain to perform the follow-up PCCys incorporation and uncaging experiments. I cultured synchronised nematodes of L1 stage on NGM plates without/with PCCys supplements (of 5mM) for 3 days. The obtained adults, either received whole-worm UV illumination for 5 minutes or not, were all imaged on the fourth day of the serial experiments.

The tail-located PQR neuron was first checked under microscope, as from the lateral it is surrounded by less tissue than the head neurons are. To ensure that the target cell was indeed gone, and the disappearance of a fluorescent cell is not due to out-of-focus microscopy, I performed serial imaging along the z-axis of the mounted worms and projected all these layers in image analysis to obtain the final figures. As displayed in Figure 3.6B, PQR expresses CFP under all conditions while only displays mCherry signal when the worm was provided with PCCys. The PQR of the PCCys-incorporated worm disappeared after UV illumination, indicating that uncaging PCCys-caged caspase proteins leads to ablation of the neuron. I saw no disappearance of PQR in animals only supplied with PCCys but not illuminated and in animals illuminated but with no PCCys supplied. Therefore, mere PCCys incorporation or UV illumination (of the experimental intensity) is unable to kill the neuron.

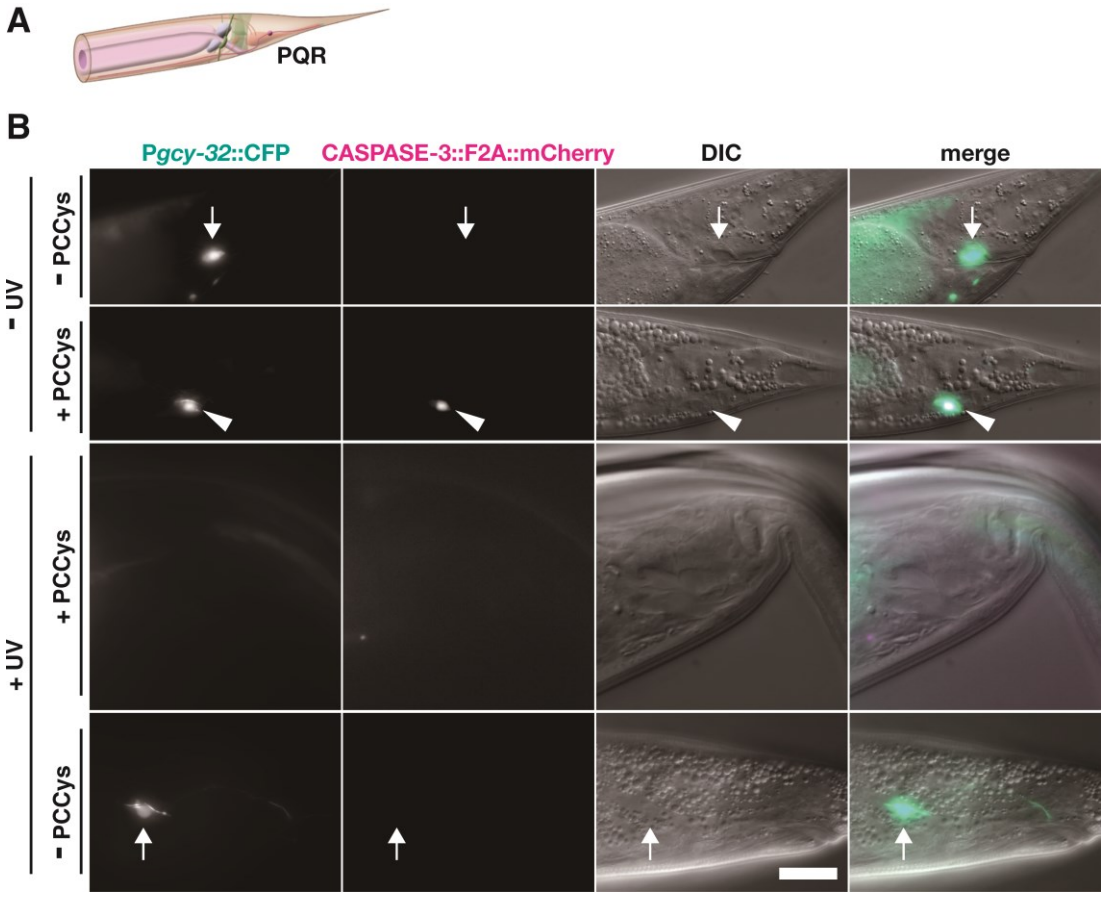

Figure 3.6 Neuron PQR killed by activated PCCys-Caspase-3 upon whole-worm UV illumination.

(A) A schematic of the O_2 sensing neuron PQR in *C. elegans*[264]. **(B)** Transgenic worms grown on NGM plates without UV illumination (top line) have only CFP expressed in PQR but show no mCherry signal, while worms feeding on PCCys in the dark (second line) express CFP in the cytoplasm and mCherry in the nucleus of this neuron. PCCys-fed worms (third line) but not NGM-grown worms (bottom line) lose PQR after UV uncaging. Arrowheads mark the intact PCCys-incorporated PQRs and arrows point the neuron without incorporation. All images are in the same scale, scale bar = 20 μm.

The experiment results of AQR (Figure 3.7) are comparable to those of PQR. In the presence of photo-caged cysteine, the activity of the caged Caspase-3 in AQR can be restored by whole-worm UV illumination. In contrast, PCCys incorporation alone without UV treatment or UV treatment alone without prior PCCys feeding was not able to trigger the death of targets.

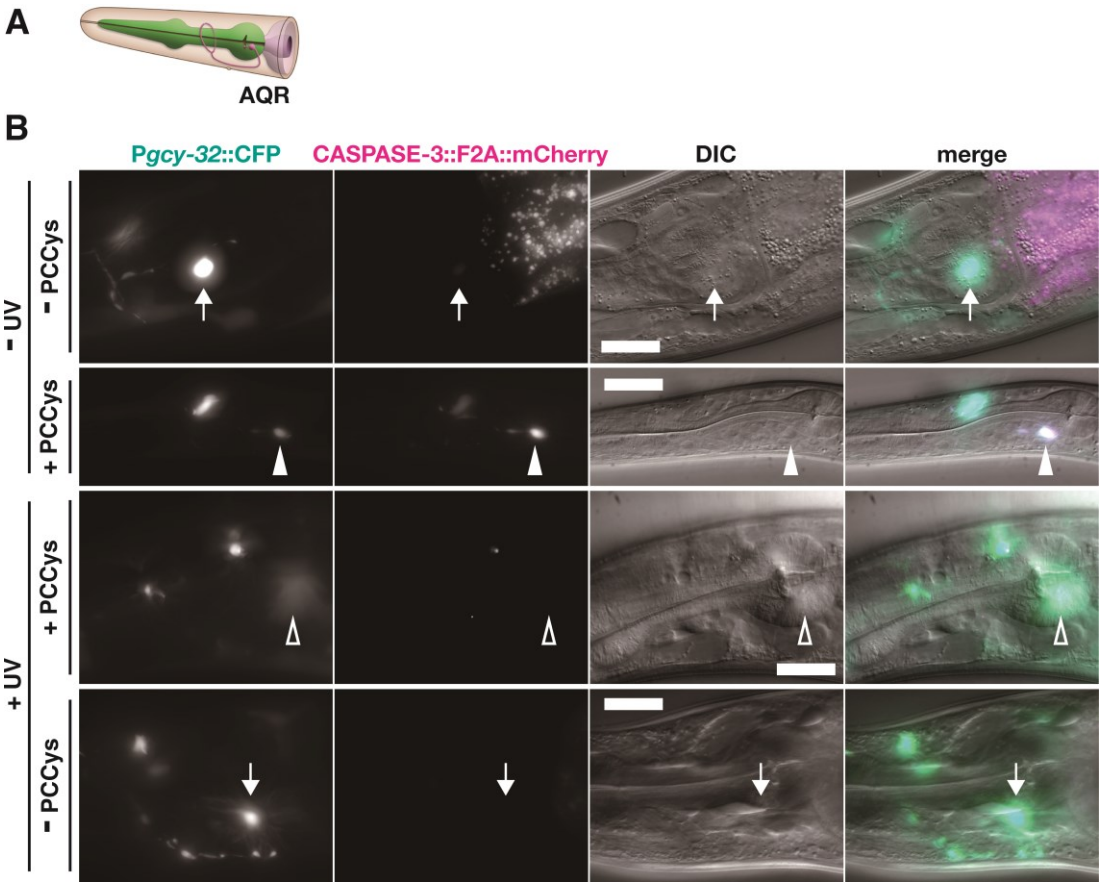

Figure 3.7 Neuron AQR killed by activated PCCys-Caspase upon whole-worm UV illumination.

(A) A schematic of the O_2 sensing neuron AQR in *C. elegans*[264]. **(B)** Worms from NGM group with no UV (top panel) have CFP-fluorescent and intact AQR soma and

processes. In comparison, the nematodes supplied with PCCys but no UV exposure (a L2 larva, second line) also show nuclear fluorescence of mCherry. AQR cells in worms grown on PCCys (third panel) underwent apoptosis and have been removed after UV uncaging. In contrast, UV illumination alone does not induce neuronal death (bottom images). Arrows point and healthy AQRs with no red signal, solid arrowheads mark healthy AQRs with mCherry expression and hollow arrowheads locate the degraded AQR. Scale bar = 20 μm.

The neuron pair URX are also located close to the terminal bulb of the worm pharynx. Despite the slightly different incorporation efficiencies between the left and right cells, both PCCys-incorporated neurons undergo apoptosis after UV uncaging, as only deformed debris can be detected at their original positions (Figure 3.8).

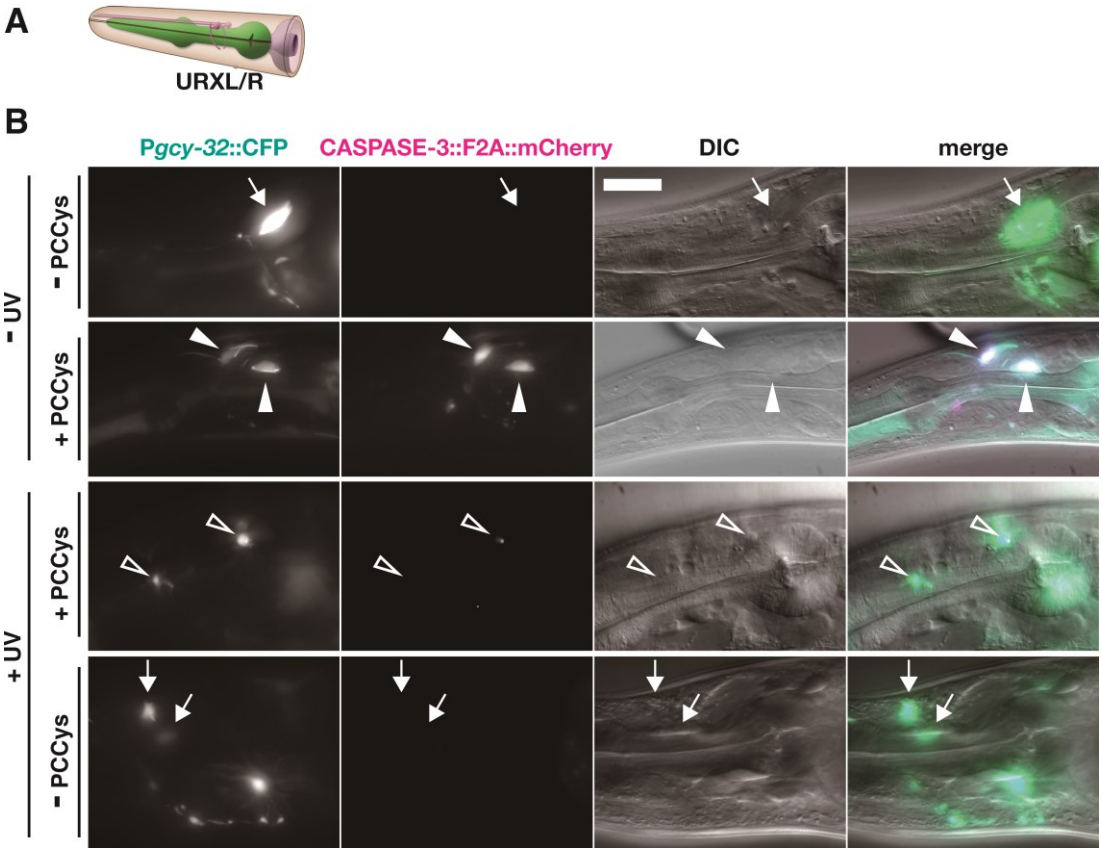

Figure 3.8 Neuron URX killed by activated PCCys-Caspase upon whole-worm UV illumination.

(A) A schematic of the O_2-sensing neuron pair of URX in *C. elegans*[264]. **(B)** URX neurons of all transgenic worms expresses CFP, but only those of worms grown on PCCys show nuclear mCherry fluorescence, either without (second row) or with UV (third row). Meanwhile, UV illumination along with cell-specific PCCys incorporation is able to ablate the neurons (the third row). Worms on NGM do not lose any URX cell subsequent to the UV exposure (bottom row). Solid arrowheads point intact URXs. Scale bar = 20 μm.

Taken together, the imaging results demonstrate that the caspase version I developed is capable of inducing apoptosis of all cells of a specific neuronal class. Apoptosis only occurs upon two steps in tandem: addition of photo-caged cysteine and exposure to 365-nm light. The data also suggests that silencing the NMD pathway is not a necessity for producing a sufficient amount of PCCys-incorporated caspase. I therefore continued to perform all subsequent experiments in the worms of N2 background.

Variability presents in several experimental aspects among these O_2-sensing neurons. First, the intensity of CFP fluorescence varied from cell to cell within the same worm (the second panel of Figure 3.7B and the bottom panel of Figure 3.8B), so the P*gcy-32*-driven expression level of the caspase construct is not equal among different cells of the neuronal class. Besides, the absolute level of PCCys incorporation, positively correlated with mCherry signal strength, also differs from cell to cell (the second line of both Figure 3.7B and Figure 3.8B). Moreover, the incorporation-dependent apoptosis of each cell type manifested differently: at the time point when PQR of a worm was entirely removed (i.e., no trace of the cell was left) after an uncaging experiment, in the original location of AQR an amorphous fluorescence remained and a dying URX can still exist as a disk-like apoptotic body. The seemly neuron-specific responsiveness to optical uncaging, however, is

deduced roughly from a rather small collection of worms I checked or imaged. For better evaluating the cell-wise and the overall ablation efficiencies of this light-activatable caspase, I would assay post-UV apoptosis of the neurons on an adequate number of the BZ52.2.8 worms (all experimental details in Methods).

To prepare the assay, firstly (on Day 0) a synchronised L1 population of Strain BZ52.2.8 was transferred onto either NGM or 5mM PCCys plates and cultured for 3 days. Then the worms of each culture condition were picked to an unseeded NGM plate and illuminated for 5-minutes with 365-nm light (Day 3). On Day 4 the worms of all treatments were checked individually and in a randomised order under a fluorescent microscope. According to the following criteria, each O_2-sensing neuron (either present or already eliminated) was given a semi-quantitative score: “1” for normal neuronal morphology, “0.5” for an abnormal shape or size, and “0” for complete disappearance. Two scores respectively of left and right URX cells were added up. I performed five replication experiments, with Ailish Tynan assisted me in giving scores to the neurons for minimising confirmation bias, and then I statistically analysed the scoring data as below.

To quantify the efficiency of cell ablation after optical uncaging of the

photo-caged Caspase-3, I decided to conduct analysis separately for each neuronal type. In this way, I would be able to discover possible differences in apoptotic responsiveness of PQR, AQR and URX. For each neuronal type, I averaged the scores of each treatment group (i.e., either PCCys-absent or PCCys-present) for each assay replicate. The averages of all replicates were gathered and their distribution was tested by a Shapiro-Wilk normality test[399]. All these tests show p-value results above the threshold 0.05, thus the null hypothesis of Shapiro-Wilk test that assumes the normal distribution of data should be accepted, meaning that I should choose a parametric method to compare two groups. I thus analysed all replicate data of the "without PCCys" group and the "with PCCys" group, using Student's t test with Welch correction. As is shown in Figure 3.9, a significant difference between caged and uncaged states could be statistically validated for each type of neuron. Therefore, the whole-worm optical activation of caspase is capable of apoptotically ablating specific class of neurons in living worms.

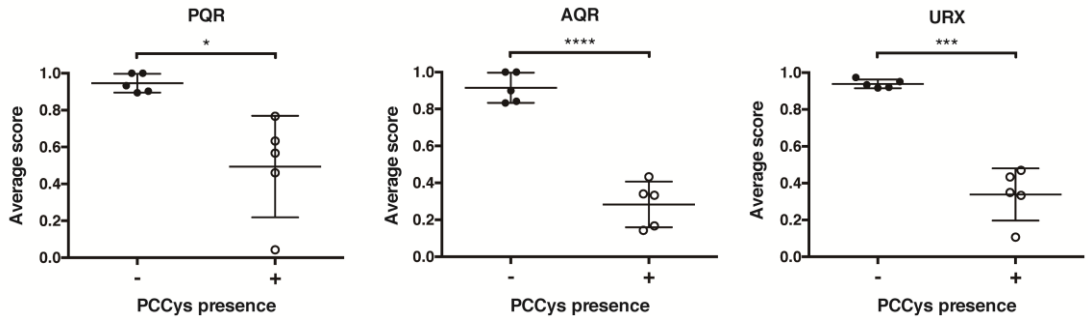

Figure 3.9 UV-activated death of each PCCys-incorporated O_2-sensing neuron is statistically significant in the cell ablation assay.

For each of the total five replicates of the ablation assay, the state of oxygen-sensing neurons (i.e., PQR, AQR, URX) was evaluated under fluorescent microscope in between 25-30 worms that had been UV illuminated after being grown either with or without PCCys supply. The normality of the data set of each neuronal type was tested by Shapiro-Wilk normality test[399] and subsequent significance tests were performed using Student's t test with Welch correction. Error bars show standard deviation of all replicates. * is for a p-value < 0.05, *** for < 0.001 and **** for < 0.0001.

Furthermore, because from every individual worm I have obtained a complete array of scores for all oxygen-sensing neurons, with no data missing, for either treatment group ("-PCCys" vs. "+PCCys") I am able to compare all neuronal types in a single statistical test. The normality assumption of each group has already been confirmed, so I tested the data as well as their derivatives after logarithmic and square root transformations for homogeneity

of variances by Bartlett test. Although the p-values of the original data are all below 0.05, the p-values of the base-2 logarithms are indeed higher than the significance threshold, indicating they have the same variance level (i.e., the null hypothesis of Barrlett test). Then I chose one-way ANOVA (pair-matched, with the Geisser-Greenhouse correction) to test the logarithmically transformed data (plotted in Figure 3.10) and obtained the result of p-values over 0.05. Therefore at the time point of microscopic scoring (24 hours after UV illumination), there is no statistically significant difference between any two types of neurons, in regard of their physiological reactions to both the worm sample preparation and the serial experiments required for light-activatable apoptosis.

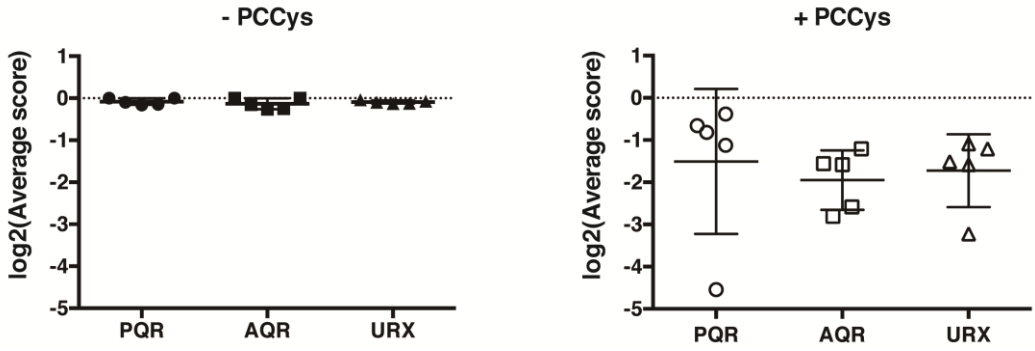

Figure 3.10 The responses to UV illumination do not significantly vary among the different O_2-sensing neurons under the same condition of PCCys supply.

The base-2 logarithmic derivatives of the same data sets as in Figure 3.9 are plotted separately for the cases without or with PCCys supply. All three types of neurons were altogether tested for their differences by one-way ANOVA and the variance observed turns out to be statistically insignificant. Error bars show standard deviation of the replicates.

As for the observed variability of early apoptotic manifestations of different O2-sensing neurons, one possible contributor is the different surroundings of individual neurons: for a sample worm mounted sideway on a slide, PQR is the nearest from the top side, so ultraviolet light illumination of the whole-worm mode has the shortest transmission path through the tissue above the target, i.e., the least attenuation to reach this target to activate more caspase proteins. As for the head neurons, AQR resides left to the median axis, the effect depends on which side the worm lies. Nevertheless, the left and right cells of URX present accordingly varied speed of apoptosis: as I can see in the third row of Figure 3.8B, the left URX cell has already become shapeless debris while the right counterpart appears as a disk-like apoptotic body. Nevertheless, the inconsistent attenuation of UV light in uncaging different cells can be technically solved in single-cell targeting with the help of microscope-mounted UV laser.

So far, I have demonstrated in oxygen-sensing neurons that by using light to activate PCCys-caged Caspase-3, I am able to efficiently ablate the specific group of cells. I then proceeded to optimise the orthogonal system for encoding PCCys and to test the anticipated efficiency enhancement by light-activating ablation of the same class of neurons.

Two means of optimising the orthogonal aaRS/tRNA pair have been detailedly discussed in Chapter 2.3. By attaching a nuclear export signal to the tRNA synthetase (NES-aaRS), Nikić *et al.* raised the cytosolic concentration of the synthetase, thereby improving the incorporation efficiency of Boc-Lys[281]. Serfling *et al.* showed that adopting rationally designed variants of $tRNA_{CUA}$ (i.e., $tRNA^{M15}_{CUA}$ and $tRNA^{C15}_{CUA}$) alone or in combinatory use with NES-aaRS can increase the incorporation yield of both Boc-Lys and Z-Lys[290]. However, both studies tested their optimisation means by the incorporation of lysine-based NCAAs in only mammalian cells, so the uncertainty of their respective *in-vivo* effects in *C. elegans* should not be neglected.

Concurrently with the optimisation work on the PCLys-incorporating orthogonal PCKRS/$tRNA_{CUA}$ pair in worms by the above two means, by Davis *et al.* in the lab[169], I decided to test the potential enhancements of using

NES-PCCRS and $tRNA^{M15/C15}_{CUA}$ on PCCys incorporation first in O_2-sensing neurons, and later in mechanosensory neurons (shown in next section). I therefore designed the serial bombardments (no. 194-202) to N2 worms: each bombardment trial implements one combination of the 3 versions of synthetase (the original PCCRS and the variants with either PKIα NES[282, 283] or Smad4 NES[400]) and 3 versions of tRNA (the original $tRNA^{Pyl}_{CUA}$, $tRNA^{M15}_{CUA}$ and $tRNA^{C15}_{CUA}$) (Table 3.1). Correspondent plasmids for this experimental plan have been constructed and transformed for generating transgenic strains (Kieran Baxter assisted me in the bombardments and worm screens). Worms of each strain were grown on plates in the absence or presence of PCCys for 3 days and then observed under fluorescent microscope.

Table 3.1 Bombardments of each combination of the orthogonal aaRS/tRNA variants in oxygen-sensing neurons.

Plasmid details of the bombardment no. 194-202 are in Appendix. Nuclear export signals (NES) from human genes (PKIα[282, 283] and Smad4[401]) had been inserted in two of the PCCRS plasmids. In the column of tRNA, PylT stands for the original *Mm* $tRNA^{Pyl}_{CUA}$, M15 stands for $tRNA^{M15}_{CUA}$ and C15 is for $tRNA^{C15}_{CUA}$. “PCCys- mCherry+ Strains” column contains the numbers of strains that have red mCherry fluorescence without PCCys supplement. “PCCys+ mCherry+ Strains”

column shows the numbers of strains with mCherry fluorescence after being cultured in presence of PCCys. N/A stands for “not applicable”, for I had too few worms of the transgenic strain.

Bombardment No.	NES of PCCRS	tRNA	Transgenic Strains	PCCys- mCherry+ Strains	PCCys+ mCherry+ Strains
194	/	PylT	4	1	1
195	/	M15	6	3	3
196	/	C15	7	1	1
197	Hu PKIα	PylT	7	4	3
198	Hu PKIα	M15	6	4	4
199	Hu PKIα	C15	7	2	3
200	Hu Smad4	PylT	6	3	3
201	Hu Smad4	M15	3	3	3
202	Hu Smad4	C15	5	N/A	N/A

In the no-PCCys worms of many strains from bombardment no. 194-201, I found mCherry signal in the nuclei of all oxygen-sensing neurons. This fluorescence did not result from the encoding of PCCys at the inserted amber site, but by appearance is indistinguishable from the mCherry shown upon PCCys incorporation. Figure 3.11 displays typical images taken from such worms of two independent strains, BZ194.2.1 and BZ197.6.1. Compared with control animals of Strain ZG610 that only express GFP in the neurons, worms of these transgenic strains have O_2-sensing neurons with similarly bright GFP, and CFP in the cytosol and mainly nuclear mCherry expression. In the

meanwhile, all these cells retain normal shapes with healthy-looking neutrites, so they do not contain active caspases.

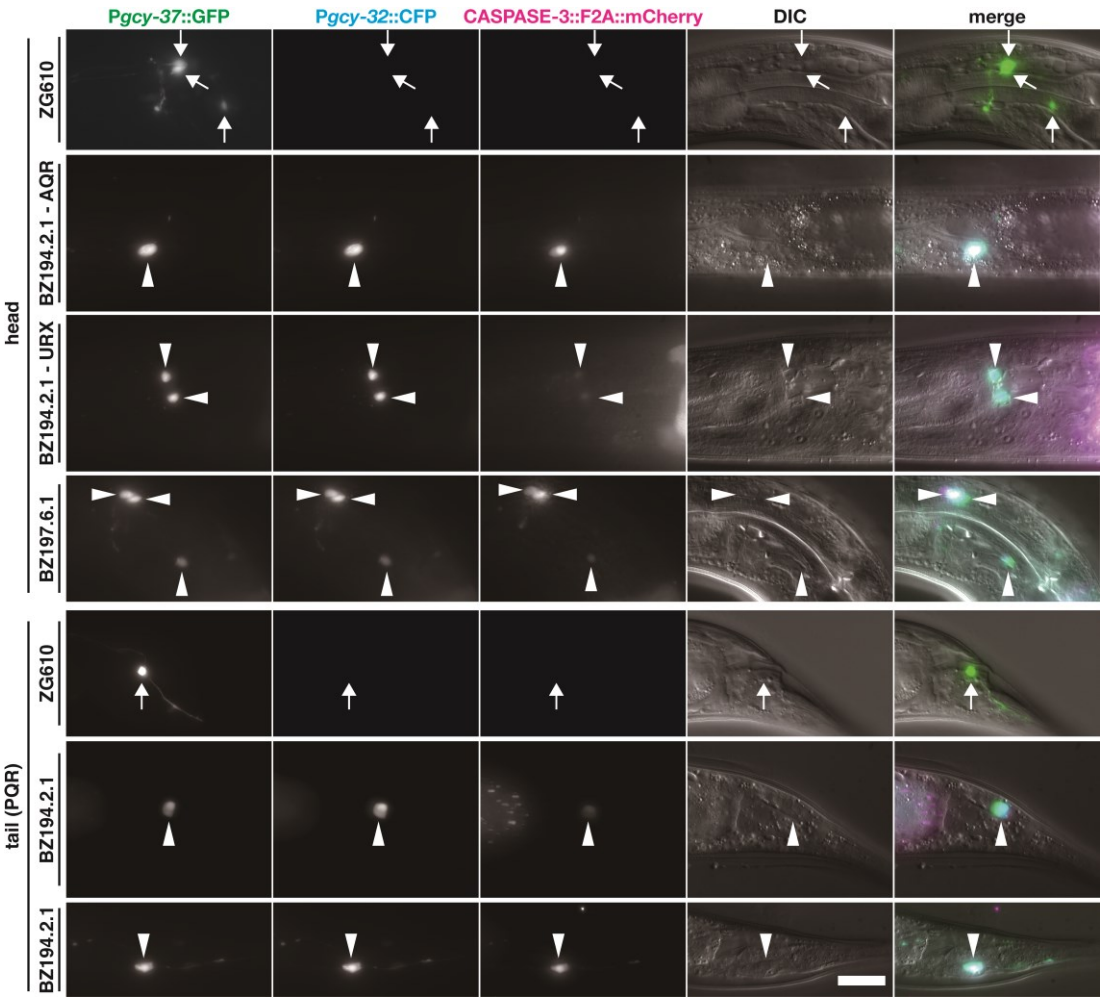

Figure 3.11 Transgenic strains express nuclear mCherry in the O_2-sensing neurons with no PCCys incorporation.

Worms of strains ZG610, BZ194.2.1 and BZ197.6.1 were all kept on NGM plates before imaging. Compared to the negative control (ZG610, the same images of Figure 3.5B) with only GFP fluorescence in the AQR (head), URX (head) and PQR (tail), mCherry signals of varied brightness are observed in the nuclei of the cells in

BZ194 and BZ197 worms, co-localising with the background GFP signals and the transgenic-reflecting CFP fluorescence. All images follow the same 20-μm scale bar.

Nonspecific amber read-through in the absence of NCAA has been previously reported by Parrish *et al.* [165], possibly due to the production of gene product fragments or antisense RNAs in the DNA rearrangements of extrachromosomal array formation[132, 157]. There is also possibility that a mutation might have occurred in the caspase-3 plasmid (ZX122), especially at the amber stop codon, which allows downstream translation and subsequent production of dysfunctional Caspase-3 proteins and the downstream mCherry proteins. In either case, all cells of the antibiotic-selected transgenic worms should contain arrays with the mutation, which is not difficult to detect. Therefore, I prepared lysates from worms of the representative strains that I imaged (BZ194.2.1 and BZ197.6.1), and amplified and sequenced the caspase-3::F2A::mCherry::HA::NLS::SL2 region of ZX122 plasmids within the lysate.

Sequencing results (Figure 3.12) showed no mutation in any of my samples. I then stepped back to inspect the design of the genetic component used to seek possible explanations. Suspected contributors to the peculiar fluorescent signals will be discussed in Chapter 3.3.1.

Figure 3.12 No sequence change detected in the caspase-expressing plasmid (ZX122) for the serial bombardments or the worm lysates.

DNA sequencing peak charts of the caspase-3 coding region in the plasmid (ZX122) in its dilution for the serial bombardments (no. 194, 197 included) as well as in the worm lysates of strains BZ194.2.1 and BZ197.6.1. Sequences of the incorporation-directing amber stop codon and its vicinity all match the original design.

During apoptosis, cells display morphological alterations like perinuclear blobs, nuclear shrinkage and take on a button-like appearance under DIC optics. However, the final state of a complete cell ablation, the disappearance of the target, is invisible *per se*. In the above experiments, I have to express at least one fluorescent protein (GFP from strain background and/or CFP from a transgenic plasmid) as a visual marker. The disappearance of the

fluorescence reflects the death of the cell targeted and hence the caspase activity.

The rationale is however not completely rigorous, as a cell could die non-apoptotically of factors other than caspase activation. The contribution of experimentation can be excluded via setting appropriate controls, but there are more direct ways to indicate apoptosis, for instance, using terminal deoxynucleotidyl transferase to label the characteristic nuclear DNA fragmentation of apoptotic cells[402, 403], or to directly measure or visualise caspase activity[404, 405]. Thus I tried several caspase-reporting variants of fluorescent proteins (later termed as caspase reporter), namely ZipGFP, GC3AI and VC3AI, whose fluorophore maturation depends on the caspase cleavage of a linker sequence[404, 405] (Figure 3.13A&B). I built the coding sequence of each of these reporters into an operon downstream to the P*rps-0*-driven HygR gene of the tRNA synthetase construct (Figure 3.13C). Then I bombarded the respective new construct of each caspase reporter along with the plasmids of the orthogonal tRNA and the caspase into worms. The obtained mutant strains (BZ136-138) have the same genetic composition with the strain BZ52.2.8, except for an additional caspase reporter that gets ubiquitously expressed (by P*rps-0*). To test whether or not each caspase

reporter works, I cultured transgenic worms on PCCys-NGM plates and then UV uncaged them.

Three strains in total showed the incorporation of PCCys in at least one oxygen-sensing neuron. At this stage, I did not see any sign of fluorescence from the caspase reporter (upper line of Figure 3.13D). Around 1 day after I UV-illuminated the worms, I observed faint signal of the caspase reporter VC3AI at the anticipated site of URX. The emission spectrum of Venus-based VC3AI overlaps with that of the transgenic marker CFP, so I had to observe and image VC3AI through the channel of YFP, for distinguish VC3AI signal from CFP fluorescence. I observed the co-localisation of the yellow light (of VC3AI) with both CFP and mCherry (lower line of Figure 3.13D), so the caspase reporter VC3AI is able to indicate the cleavage activity of activated caspases. In fact, my practice of avoiding fluorescent bleed-through by using the YFP filter set greatly reduces the signal intensity as well as the dynamic range of VC3AI, so this trial failed to attain real-time demonstration of caspase cleavage activity. Nevertheless, this result is another piece of evidence of my optical re-activation of caspase in oxygen-sensing neurons.

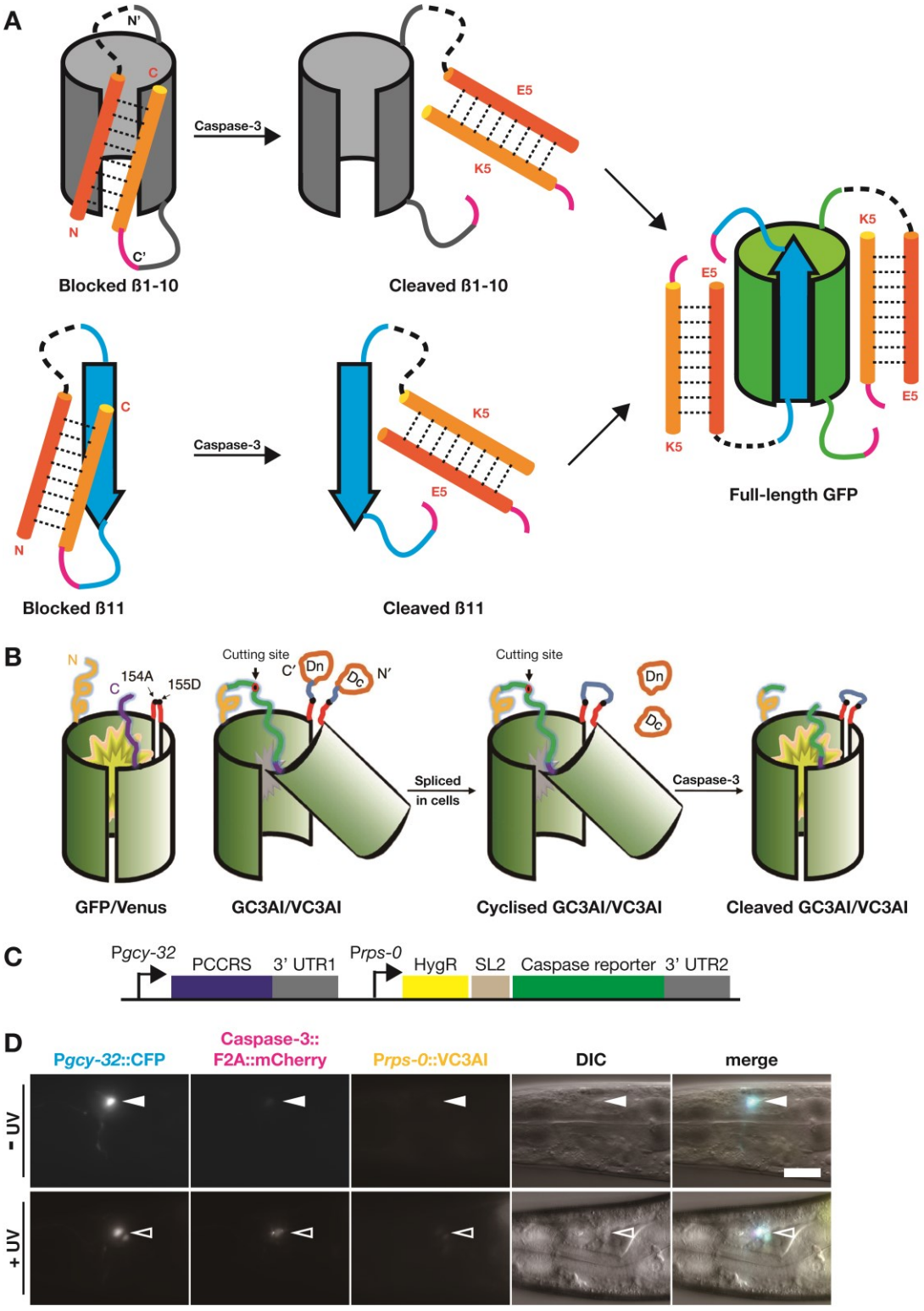

Figure 3.13 The indication of the ongoing apoptosis by caspase reporter.

(A) Schematic of the caspase reporting mechanism of ZipGFP. Adapted from To *et al.*[407]. ß1-10 (barrel in grey) of the 11 ß strands of GFP is flanked by E5 (in red) at N terminus (denoted as N') and by K5 (in orange) at C terminus (denoted as C'). E5 and K5 are coiled coils that can heterodimerise with each other[406]. Linkers are shown as black dashed lines and magenta solid lines. The magenta lines contain cutting sites (residues DEVD) of Caspase-3. ß11 (in blue) is flanked with K5 and E5 in the same manner. K5 and E5 block the assembly of ß1-10 and ß11, which can be removed by the Caspase-3 cleavage at residues DEVD within their linkers. Then the free ß1-10 and the free ß11 can assemble a full-length GFP that emits fluorescence (indicated in green colour). **(B)** Schematic of the caspase reporting mechanism of GFP/Venus-based caspase-3-like protease activity indicator (GC3AI/VC3AI). Adapted from Zhang *et al.*[404]. Yellow lines and purple lines represent respectively the N- and C-terminal sequences outside of the ß-barrel structure and they are linked with an insertion of Caspase-3 cutting site (residues DEVD) in GC3AI/VC3AI. Also in the design of both GC3AI and VC3AI, residues A154 and D155 are split and each linked through additional residues GSAEY and CFNEF (in red) with the N- and C fragments of *Npu DnaE* intein (Dn and Dc, respectively). Dn and Dc can associate spontaneously to self-splice from the C3AI sequence and *trans*-ligate the residues GSAEY and CFNEF[408, 409], so that GC3AI/VC3AI is cyclised and not fluorescent. Cleavage of Caspase-3 at the cutting site will restore the green fluorescence of the chromophore. **(C)** The genetic construct for expressing the above caspase reporters.

The reporter sequence is placed with a SL2 in the downstream of the P*rps-0*-driven HygR gene, in the same plasmid of PCCRS. **(D)** PCCys-fed worms were imaged before and (24 hours) after UV uncaging. A YFP filter set was chosen for separating the signal of caspase reporter VC3AI from that of CFP. No fluorescence was seen through the YFP channel in living neuron (marked by solid arrowhead in top panel), but faint yellow signal co-localised with both CFP and mCherry in the apoptotic cell (marked by hollow arrowhead). Scale bar, 20 μm.

3.2.2.1.2 Touch receptor neuron

The results of activating engineered Caspase-3 in oxygen-sensing neurons have reached the objective of ablating specific group of cells and then I decided to reproduce the optical control of this caspase version in another class of neurons and later in other tissues (Chapter 3.2.2.2). A range of cell ablation methods introduced in Chapter 3.1.3, such as optogenetic ablation tools [188, 203] and recCaspase-3[180], have the killing effects tested in touch receptor neurons. To compare my engineered caspase with them, I decided to specifically express the caged caspase in these mechanosensory neurons by using a 471-nucleotide promoter of *mec-7*[410]. Gene *mec-7* is a ß-tubulin gene and its protein product together with MEC-12 (an α-tubulin) assembled 15-protofilament microtubules that are unique to the six touch

receptor neurons (ALMs, AVM, PVM and PLMs) of the *C. elegans* soft touch circuit[411-414]. The mRNA level of *mec-7* is high during late development and adulthood (median Aggregate expression estimate is around 13.4 in L4 larva and 10.9 in adults, from WormBase[150]).

In addition, I intended to compare the incorporation efficiency between the two ways of tRNA expression: one is the sequential arrangement of a Pol II promoter (here is P*mec-7*) and a Pol III promoter (P*rpr-1*) (dual-promoter design, Figure 3.14A), and the other uses one P*rpr-1* in a tRNA cassette (mono-promoter design, Figure 3.14B). The dual-promoter design was adopted and proven only in neuronal incorporation of PCCys in the fluorescent reporter (Chapter 2.2.1.2), whereas the aforementioned success of PCCys incorporation in caspase used the tRNA plasmids of the second design. The tRNA construct of either dual- or mono-promoter design was co-transformed into N2 worms with the plasmids expressing the engineered caspase and the orthogonal synthetase PCCRS (Figure 3.14C), in either bombardment no. 144 or no. 203, respectively. One transgenic strain was obtained from bombardment no. 144 and four strains from bombardment no. 203. These worms show cyan fluorescence from the expression of the caspase plasmid, in all the six touch receptor neurons, suggesting that the promoter is able to initiate cell-specific transcription of the coding sequences

of caspase as well as of CFP. However, none of the transformants shows the proxy red fluorescence of PCCys incorporation when being supplied with the amino acid, even in cells with high levels of CFP. I speculated that the final incorporation efficiency of the current orthogonal system turns out below the detection threshold; hence I did not proceed to try uncaging the PCCys-grown worms to see if target neurons would still be ablated. It was in later experiments (of Figure 3.15-3.18 and those in Chapter 3.2.2.2) that I came to realise that very low level of PCCys-incorporated caspase once uncaged might be sufficient to kill the target cells, even this level of incorporation does not concurrently generate observable amount of mCherry. Nevertheless, I sought incorporation improvements via the aaRS/tRNA optimisation methods I tested in O_2-sensing neurons: the attachment of NES to the tRNA synthetase[281] and the sequence optimised $tRNA_{CUA}$ variants ($tRNA^{M15}_{CUA}$ and $tRNA^{C15}_{CUA}$)[290].

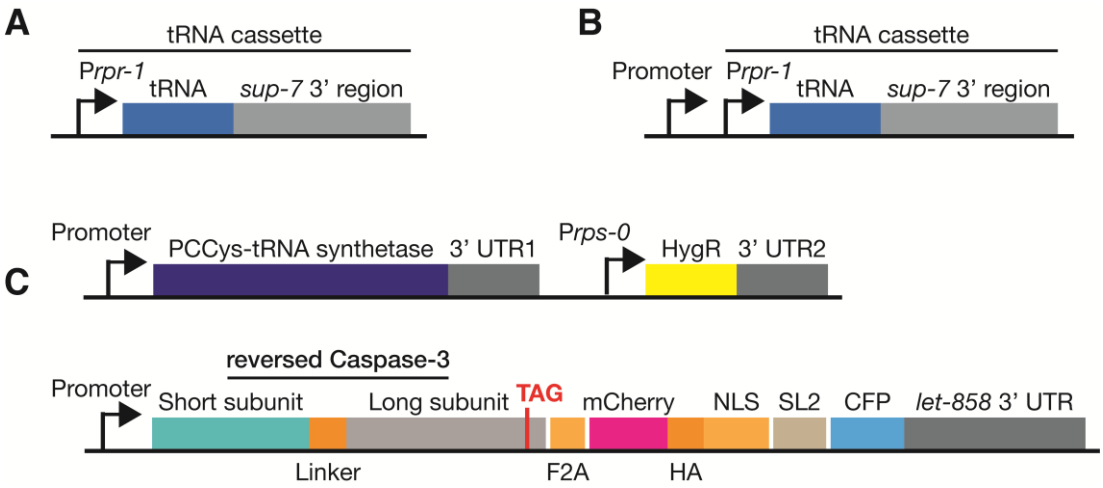

Figure 3.14 Schematic of the genetic components used initially for caspase activation in mechanosensory neurons.

(A) The orthogonal tRNA plasmid of bombardment no. 144 has two promoters, P*mec-7* and P*prp-1*, arranged in tandem. **(B)** The P*prp-1*-driven $tRNA_{CUA}$ construct used in bombardment no. 203. **(C)** Maps of the common plasmids of bombardments no. 144 and 203 for expressing the PCCys-tRNA synthetase and the engineered caspase.

I built constructs of 3 versions of synthetase (the original PCCRS, the PKIα NES-PCCRS variant and Smad4 NES-PCCSRS, all driven by P*mec-7*) and 3 versions of tRNA (the original $tRNA^{Pyl}_{CUA}$, $tRNA^{M15}_{CUA}$ and $tRNA^{C15}_{CUA}$, all driven by P*prp-1*). All possible combinations of the orthogonal PCCRS/tRNA pair except the original *Mm* PCCRS/$tRNA^{Pyl}_{CUA}$ were tried with eight experiments (bombardment no. 204-211, listed in Table 3.2).

Transgenic strains generated from each bombardment were each assayed for PCCys incorporation, in which L1 larvae were transferred to PCCys-NGM plates, cultured for 3 days, and then microscopically checked for the red mCherry signal. Worms that showed PCCys incorporation were subsequently divided into two groups ("-UV" and "+UV"), transferred onto two correspondent new plates, and only "+UV" group were illuminated for 5 minutes in a UV crosslinker. Afterwards, both groups were cultured for another day and then imaged under fluorescent microscope. Table 3.2 summarises the experimental results of worms obtained from each bombardment group upon PCCys supplementation and UV uncaging experiments.

Control groups "-PCCys, -UV" and "-PCCys, +UV" were not designed here at the time, as I regarded the above experiments preliminary and would like to progress faster to the next step of my plan: the proof-of-principle PCCys incorporation and ablation experiments on cells of other tissue types. Then I would come back to conduct a "formal" experiment that contains all control groups. However, I was only able to perform these "preliminary trials" before a breakout of worldwide pandemic and the consequent travel disruption and quarantines. Nevertheless, not having a full set of comparison groups is a flaw in some of my experimental designs (other cases pointed out in

correspondent sections), which reminds me to attach equal weight to various experiments at different stages of my project.

Table 3.2 Summarised PCCys incorporation and cell ablation results of the bombardments of each combination of the orthogonal tRNA synthetase/tRNA variants in touch receptor neurons.

Results of bombardment no. 203 (no optimisation on the original pair of PCCRS/tRNA) were included for comparison. Plasmid details of the bombardment no. 203-211 are in Appendix. For the nuclear export signals inserted into plasmids of PCCRS, “Hu PKIα” stands for the NES of human protein kinase inhibitor gene (form alpha)[282, 283], and “Hu Smad4” means the NES of the human protein Smad4[401]. In the column of tRNA, PylT stands for the original *Mm* $tRNA^{Pyl}_{CUA}$, M15 represents $tRNA^{M15}_{CUA}$ and C15 is for $tRNA^{C15}_{CUA}$. The experiments were conducted on N2 worms. “Incorporating Strains” column contains the numbers of strains that have succeeded in incorporating PCCys. “Ablating Strains” column has the numbers of strains that have shown cell apoptosis after UV caging.

Bombardment No.	NES of PCCRS	tRNA	Transgenic Strains	Incorporating Strains	Ablating Strains	Touch receptor neurons ablated
203	/	PylT	4	0	0	/
204	/	M15	2	1	0	/
205	/	C15	7	5	2	All cells for both strains

206	Hu PKIα	PylT	3	1	1	All cells except PVM
207	Hu PKIα	M15	3	1	1	ALM and PLM
208	Hu PKIα	C15	4	4	2	All cells except PVM for both strains
209	Hu Smad4	PylT	2	1	0	/
210	Hu Smad4	M15	8	8	4	All cells
211	Hu Smad4	C15	4	2	1	All cells

It seems to me that the incorporation efficiency can be increased by either using NES-tRNA synthetase alone (inferred from comparing bombardment no. 203 with no. 206 and no. 209) or using the optimised $tRNA_{CUA}$ alone (inferred from comparing bombardment no. 203 with no. 204 and no. 205). However, the transgenic strains from each bombardment group were too few for me to determine the isolated improvement by individual optimising means. Combinational use of both NES-PCCRS and the optimised tRNA variant promotes PCCys incorporation in worms, which is most evident from comparing bombardment no. 208 and no. 210 with bombardment no. 203: all transgenic strains from the former two groups exhibit mCherry fluorescence of PCCys incorporation. Subsequently, a larger amount of PCCys-incorporated caspases are available for mediating apoptosis of the target neurons upon UV

uncaging. Representative images of worms from bombardment group no. 210 are displayed in Figure 3.15 to Figure 3.18.

The pair of ALM neurons is located approximately at the first trisection point of the body axis on either lateral side of the nematode[415] (Figure 3.15A). The neuron sends a straight dendrite within the dorsal sublateral cord and underneath the hypodermis to the head, ending at where the pharynx locates. In the PCCys-supplemented, unexposed worm (first row of Figure 3.15B), ALM looks healthy and expresses CFP in both its soma and the extending processes. Despite the faint yet visible mCherry signal inside the cell area, I was sceptical about the PCCys incorporation suggested, till I found the complete removal of ALMs (no CFP or mCherry fluorescence detected) in worms by activated caspases from UV uncaging (bottom row of Figure 3.15B).

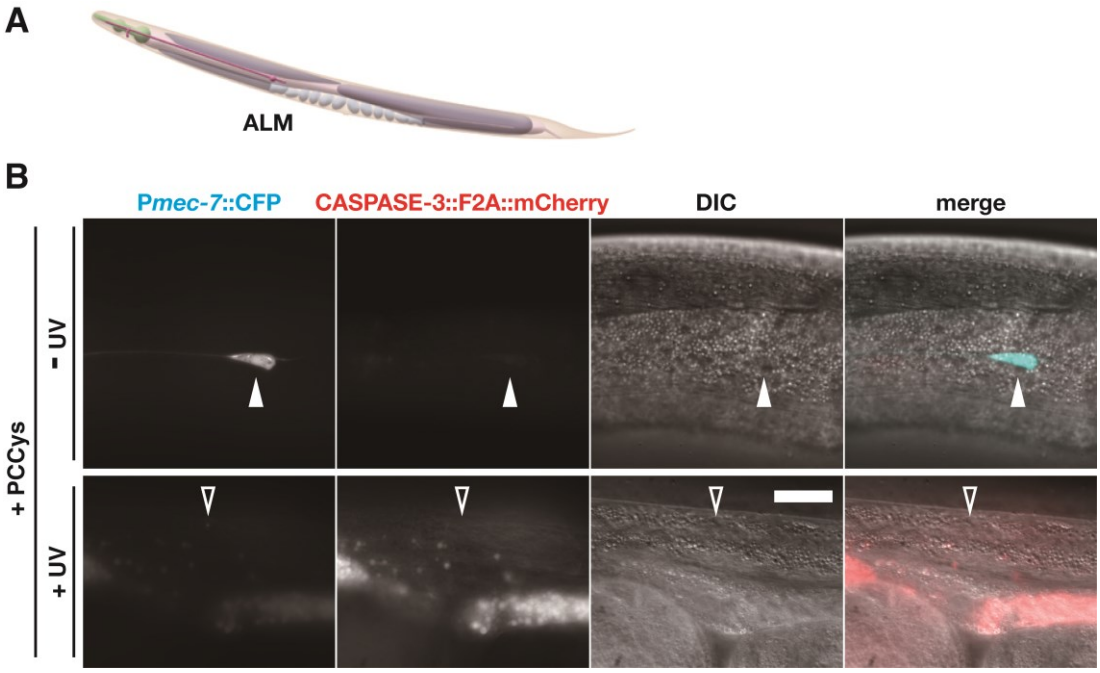

Figure 3.15 Neuronal pair ALM killed by activated PCCys-Caspase upon whole-worm UV illumination.

(A) A schematic of the left cell (ALML) in the symmetric pair of touch receptor neurons ALM in *C. elegans*[264]. **(B)** The ALML of the worm that got no UV illumination (upper) has healthy soma and processes showing cyan fluorescence and low-intensity signal of nuclear mCherry. In comparison, the worm exposed to UV (lower) loses the neuron completely. Solid arrowhead marks the healthy ALM and hollow arrowhead locates the ALM ablated. All images are in the same scale and scale bar = 20 μm.

With a main process projects firstly ventrally and then extends within the ventral nerve cord (VNC) towards the metacorpus in the head, neuron AVM is responsible for sensing mechanical stimulus in the anterior part of the

nematode[415]. In the PCCys-fed worms kept away from UV, CFP signals the intact AVM and its expression of the caspase construct. Compared with the ALM shown above, brighter mCherry expression is observed in the cell nucleus. After being illuminated by 365-nm light, the neuron underwent apoptosis and left fluorescent dot-like debris the next day.

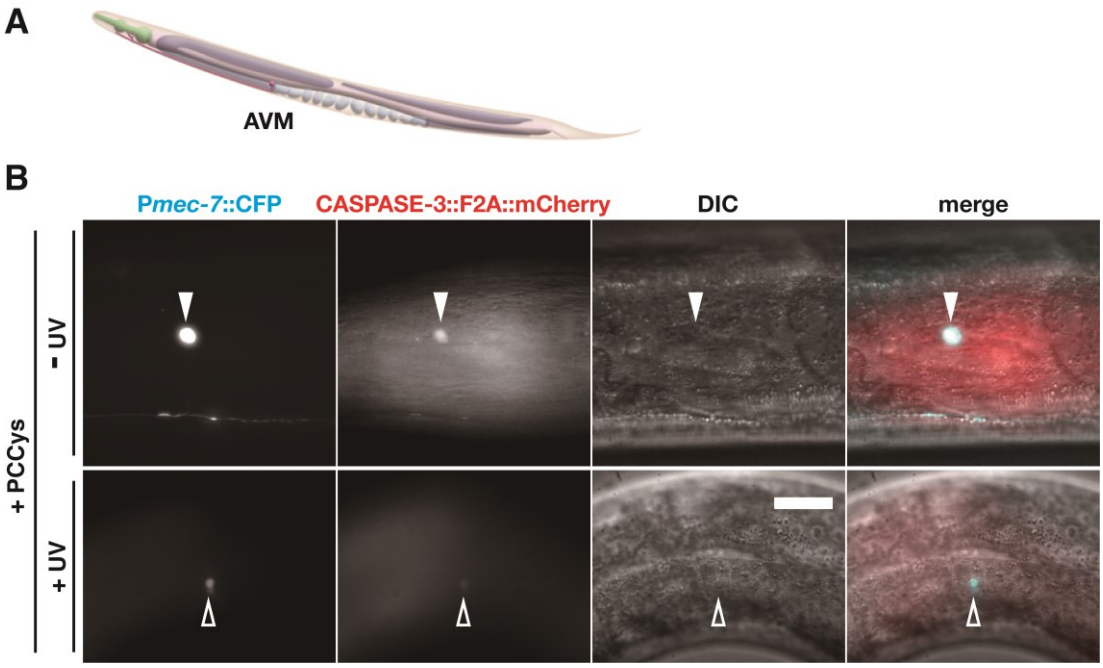

Figure 3.16 Neuron AVM killed by activated PCCys-Caspase upon whole-worm UV illumination.

(A) A schematic of the touch receptor neuron AVM in *C. elegans*[264]. **(B)** Worms grown on PCCys-NGM plates, both with (top panel) and without (bottom panel) UV uncaging, were imaged. The worm without UV exposure has its AVM (marked by solid arrowhead) expressing CFP in the soma (strongly) and the dendrite (faintly), and the incorporation-reporting mCherry weakly in the nucleus. In the illuminated

worm, only AVM debris (located by hollow arrowhead) has been detected slightly ventral to the cuticular alae via cyan and red fluorescence. All images are in the same scale and scale bar = 20 μm.

The symmetrical pair of PLM neurons is located posterior to the rectum in *C. elegans*. The long major processes from these cells join the ventral sublateral cord to reach the vicinity of the vulva[415]. PCCys-incorporating PLM neurons are in normal shape with bright cytoplasmic CFP fluorescence and nuclear mCherry signal before UV illumination. 24h after uncaging, some of them begin to have signs of apoptosis: an amorphous debris, fluorescent in both CFP and mCherry channels, shows up at the anticipated location (the bottom row of Figure 3.17B). Close to this remnant, there is a roundish structure with much weaker CFP and no mCherry. The shape without protrusion can be a slightly damaged PLM cell early in its dying course, as the level of PCCys incorporation appears low yet effective in activating caspase for apoptosis.

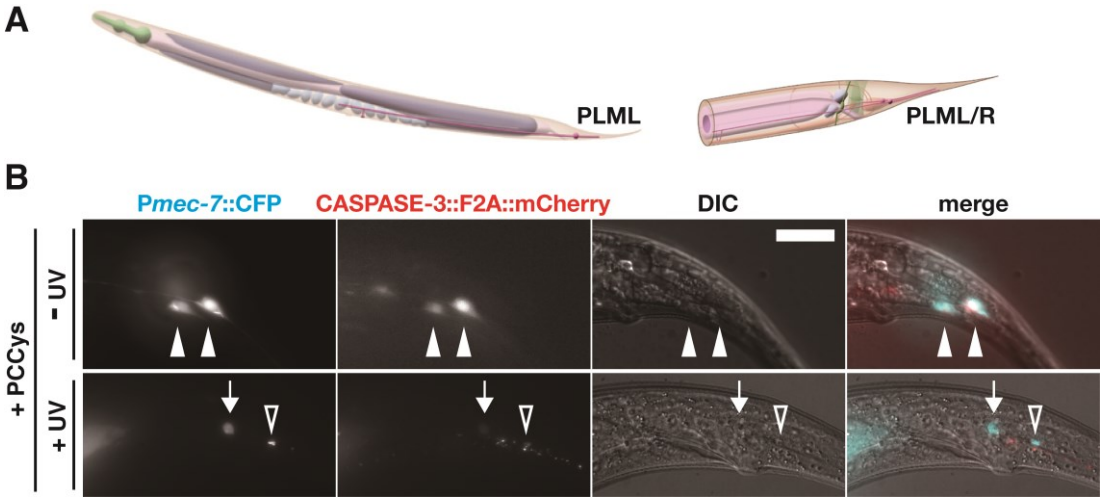

Figure 3.17 Neuronal pair PLM killed by activated PCCys-Caspase upon whole-worm UV illumination.

(A) A schematic of the touch receptor neuron PLM pair in *C. elegans*[264]. **(B)** Both top and bottom line of images show worms supplied with PCCys. PLML and PLMR of worms in the no-UV group (top) are observed with soma CFP and nuclear mCherry. On the contrary, worms treated with UV (bottom) only show apoptotic remnants of PLM cells. Normal PLMs are marked by solid arrowhead and a surely apoptotic PLM is marked by hollow arrowhead. Arrows point the neuron without detectable mCherry signal. All images are in the same scale and scale bar = 20 μm.

The PVM neuron is located approximately at the posterior point of trisection and slightly left to the median axis of the worm body. The process of the unipolar neuron grows ventrally to join the ventral nerve cord and then projects towards anterior half of the animal[415]. The dendritic segment in VNC lies parallel to that of AVM and in the side view, the anterior one third of

the PVM process overlaps with the posterior half of the AVM process. In PCCys-fed worms of the UV negative group, PVM looks normal in both DIC and CFP channels while the cell lacks mCherry signal. On the contrary, UV-exposed worms that have incorporated with PCCys contain no healthy PVM, and the typical images of an apoptotic PVM show deformed remains in the anticipated locus (Figure 3.18B).

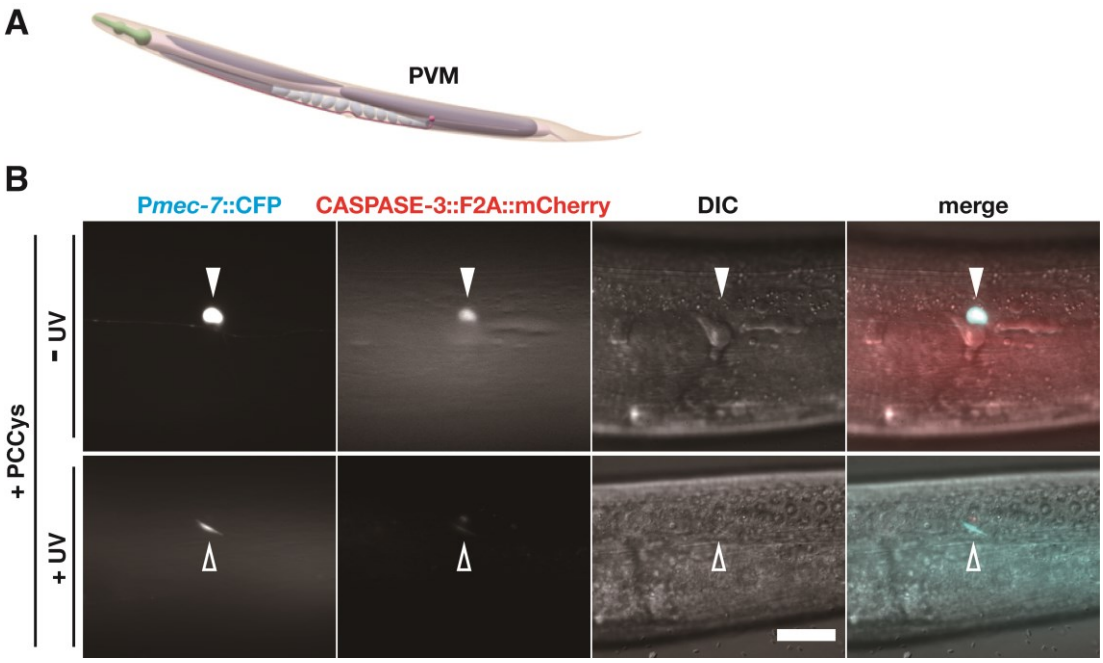

Figure 3.18 Neuron PVM killed by activated PCCys-Caspase upon whole-worm UV illumination.

(A) A schematic of the touch receptor neuron PVM in *C. elegans*[264]. **(B)** Microscopic images of worms on PCCys-NGM plates either with (top row) or (bottom row) without UV exposure. The PVM cell of the worm kept in the dark (marked by solid arrowheads) are observable via CFP and mCherry channels of the microscope.

In comparison, only fluorescent debris of the UV-uncaged neuron (marked by hollow arrowheads) exists with mere cyan fluorescence. All images are in the same scale and scale bar = 20 μm.

The above results demonstrate that the P*mec-7*-driven caspase construct combined with the optimised aaRS/tRNA pair can incorporate PCCys and, upon uncaging, can initiate apoptosis. However, as in the uncaging experiments the control groups “-PCCys, -UV” and “-PCCys, +UV” had not been completely set, so far I cannot exclude the possibility that UV illumination alone triggers death of the targets.

Additionally, from all obtained strains I observed high variability of CFP expression in the progeny of a single worm and in different neurons of the same worm. The variable expression level of the transgenes likely results from the unstable extrachromosomal presence of all the bombarded genetic components during cell mitosis[157]. For instance, the heterogeneous inheritance of the extrachromosomal arrays between the left and right cells of the PLM pair possibly causes the different apoptotic response of PLML and PLMR to uncaging (Figure 3.17B).

To stably express the constructs of the orthogonal aaRS/tRNA$_{CUA}$ pair and the engineered caspase, I decided to genomically integrate the extrachromosomal transgenic array of the transgenes, using either UVC or gamma ray irradiation[133, 416], on candidate strains from bombardment no. 205, 208, 210 and 211. Strain BZ211.2.1.7.1 showed stable cross-generational transmission of the transgenic components that got integrated into the genome (*Is[Pmec-7::smad-4 NES::PCCRS + Prps-0::hygR; Prpr-1::tRNA(C15); Pmec-7::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*). The integrated worms have stable expression and transmission of the genetic components, but in their neurons the CFP fluorescence is so dim that the signal can only be observed in the worms mounted on slides through a 40x objective lens. The weakening possibly came from a reduction of transgene copy number during the integration. Consequently, the incorporation-indicating mCherry signal weakens and becomes largely unobservable even through 63x or 100x objective lens of the microscope, so I was unable to estimate the rate of PCCys incorporation. Also, the low visibility of the marker CFP in the strain disables the assessment of presence/absence of a target cell via fluorescence microscopy. Hence, I would like to cross the integrated transgene into a strain that stably expresses a fluorescent marker in all the touch receptor neurons, so in the offspring the presence or absence

of a neuron as well its morphological changes during apoptosis will clearly revealed.

In the literature, the gene *mec-4* that has been shown involved in touch perception and its promoter has been used for driving protein expressions in only touch receptor neurons[417]. Worms of strain CZ10175 have the reporter construct P*mec-4*::GFP genomically integrated and their touch receptor neurons are stably labelled by green fluorescence. I crossed them with the worms of BZ211.2.1.7.1 to obtain the progeny (denoted as BZ211;CZ10175) whose all touch receptor neurons stably express both GFP and the integrated array that contains caspase with CFP and the orthogonal aaRS/tRNA pair.

In the subsequent cell-ablation experiments, we imaged two groups of L4 and adult worms of Strain BZ211;CZ10175: the NGM grown and UV-unexposed group as the negative control, and the PCCys-supplied and UV-uncaged group. Worms in the latter group were cultured on 5mM PCCys plates for three days and then uncaged in UV crosslinker for 5 min. I left out the intermediate microscopic check of mCherry signal for PCCys incorporation in view of the low expression level of the caspase construct suggested by weak CFP fluorescence in the neurons. One day after UV uncaging, I imaged worms through only the DIC and GFP channels to elevate assay throughput.

Figure 3.19 and Figure 3.20 show representative images of individual types of neurons in worms of both groups.

In the negative control group, BZ211;CZ10175 worms have the neuronal fluorescence resembling that of Strain CZ10175: homogenous GFP signal can be observed in both soma and the anteriorly projecting dendrite of each touch receptor neuron. In contrast, the neurons of worms grown on PCCys-NGM plates started to undergo apoptosis within 24 hours after illumination, with most of the neuronal processes disappeared.

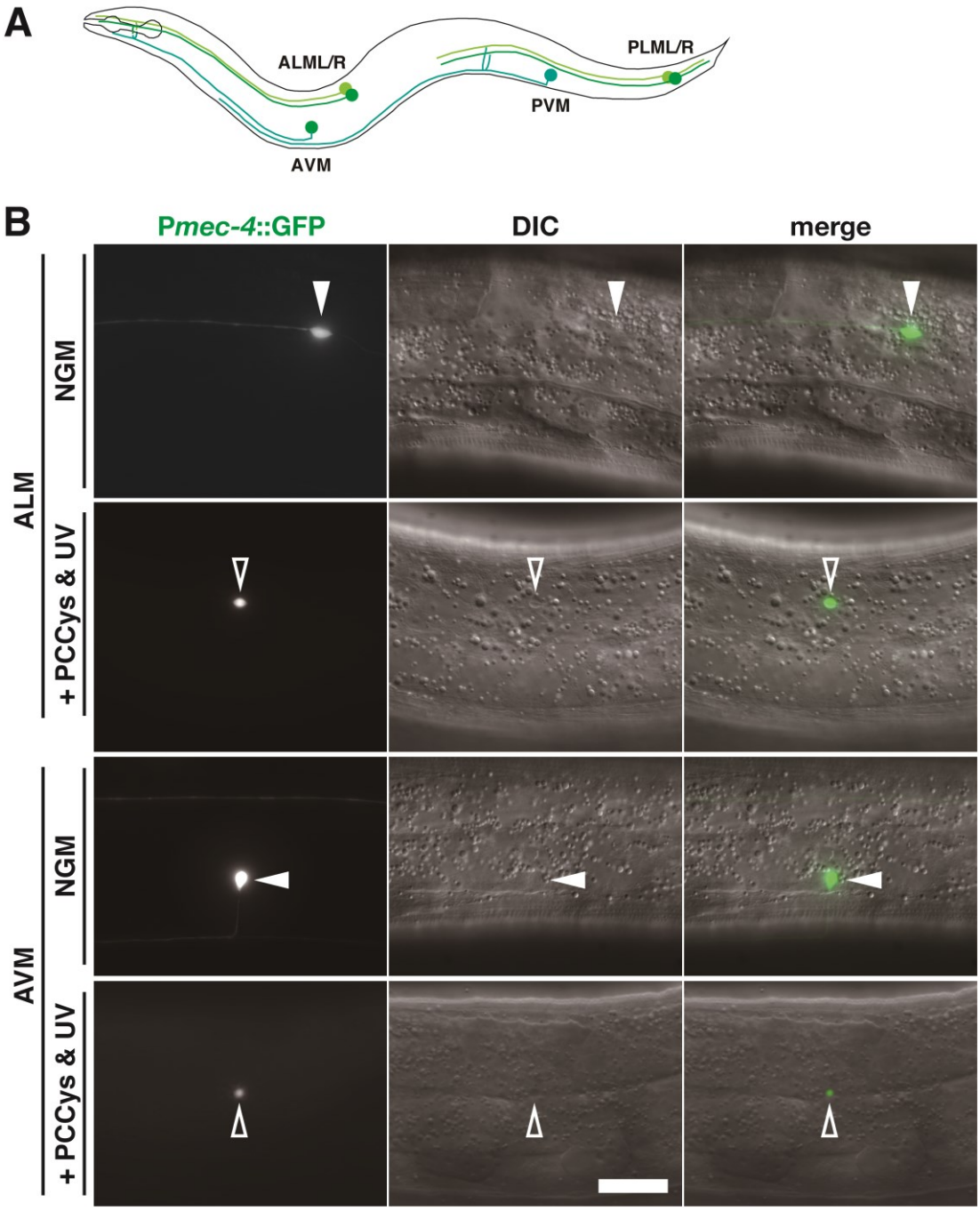

Figure 3.19 Anterior touch receptor neurons ALM and AVM can be killed by activated PCCys-Caspase upon whole-worm UV illumination.

(A) Schematic of all touch receptor neurons of *C. elegans*. Adapted from Chalfie *et al.*[415]. **(B)** Fluorescent images of worms grown on NGM plates without UV

exposure (the negative control, top and bottom rows) and UV-illuminated worms subsequent to feeding on PCCys-NGM plates (median row). ALM and AVM of the control worms express GFP in the healthy-looking somas and dendrites. Meanwhile, PCCys-incorporated and uncaged ALM presents itself as an apoptotic body. Solid arrowheads point the healthy neurons and hollow arrowheads are for the apoptotic cells. All images are in the same scale and scale bar (white) = 20 μm.

(B)

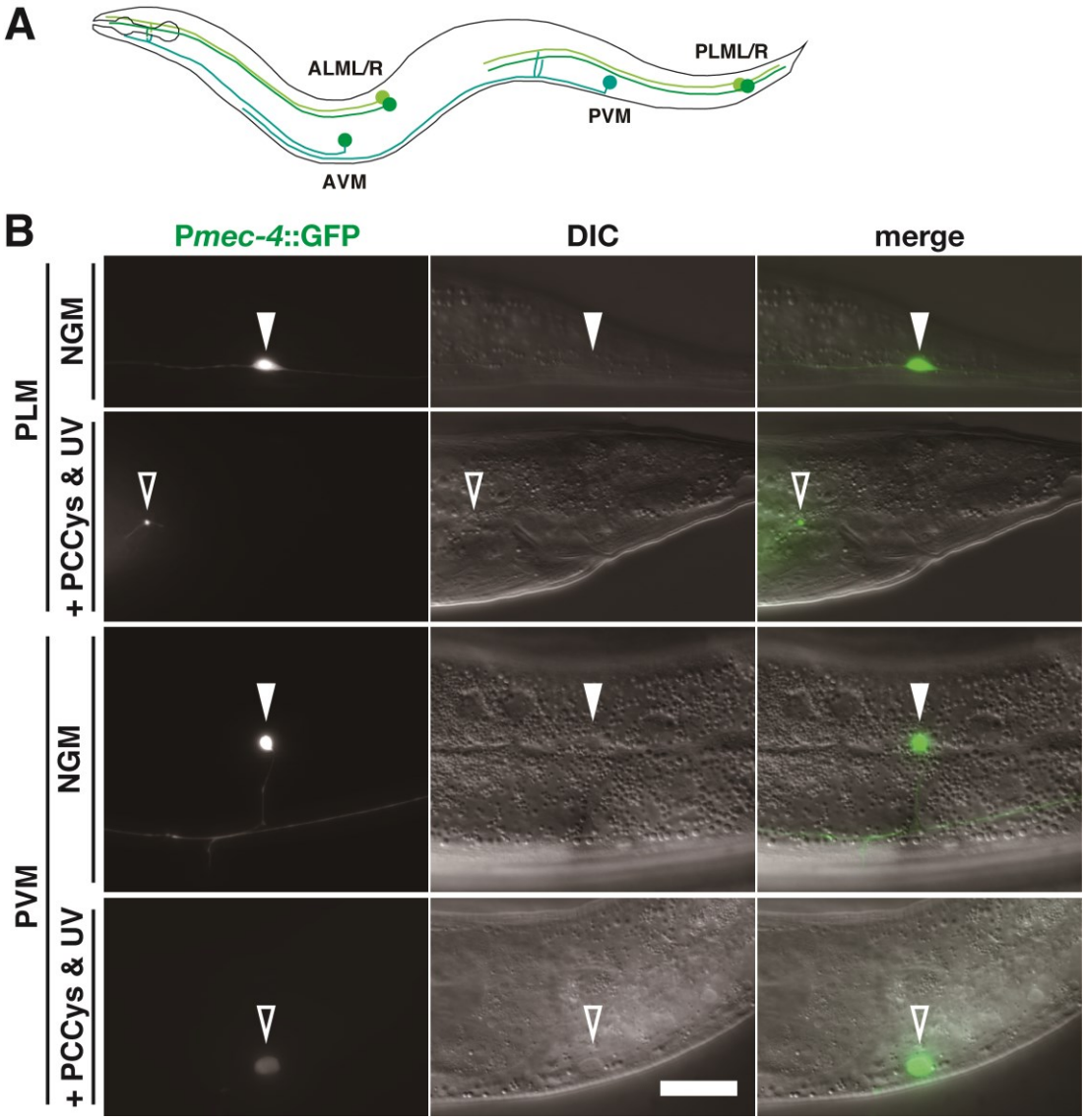

Figure 3.20 Posterior touch receptor neurons PLM and PVM killed by activated PCCys-Caspase upon whole-worm UV illumination.

(A) Schematic of all touch receptor neurons of *C. elegans*. Adapted from Chalfie *et al.*[415]. **(B)** NGM-growing worms of the control group (first and third lines) and the PCCys-supplied and UV-uncaged worms of the experimental group (second and fourth lines) were imaged through GFP channel. Control PLM and PVM cells (shown by solid arrow) exhibit somal and dendritic green fluorescence. Contrarily, the PLM

and PVM of the experimental group undergo apoptosis: the apoptotic PLM presents itself as debris while dying PVM exists as an oval apoptotic body. Hollow arrowheads locate the cells in apoptosis. All images are in the same scale and scale bar (white) = 20 μm.

Still, I observed varied removal speed of different neurons within the same worm after uncaging. Take Figure 3.20 for example, PLMs temporarily presented themselves as fluorescent dots while PVM existed as an apoptotic body at the moment of imaging. Although the sample size was way below the threshold of statistical power and I am unable to conclude if my caspase tool has a higher/lower efficiency in certain neurons, I can still learn the following from the phenomena:

1) Slightly varied efficiencies of multiple contributors (including the promoter, the orthogonal pair of tRNA synthetase and tRNA, the anatomic and postural cell positions combined in UV illumination, and the apoptotic pathway) in combination may lead to differences in the competency to apoptosis between individual neurons. Since similar effects were earlier observed in oxygen-sensing neurons (categorised as chemosensory neurons), and also in the class of mechanosensory neurons here, the variability is not confined to particular classes of sensory neurons, and may be common in most neuronal cells and even cell types other than neurons.

2) Binary assay (i.e., either 100% or 0) of ablation efficiency is not adequate in view of the cell apoptosis variably manifested. Instead, I need a grading system to evaluate the ablating effect of my caspase in the planned assay experiments. The range of neuronal states in the image collections from the above uncaging trials were categorised into Grades 0-4 (criteria details in Methods, Chapter 4.15.2). The numerical distribution of the grades can semi-quantitatively reflect the overall efficiency in individual neurons.

With the grading criteria set, I performed an apoptosis assay simultaneously on three groups of BZ211;CZ10175 nematodes: one group has worms grown on NGM plates and unexposed to UV, one group has worms cultured on NGM and then illuminated by UV, and worms of the third group were grown with 5mM PCCys supplement and uncaged by UV. At the time I did not set a "+PCCys, -UV" group as I was carried to push forward my project: I inferred from Figure 3.15-17 that PCCys incorporation alone does not seem to be pro-apoptotic, meanwhile it was time- and labour-demanding to score the PCCys-incorporated worms in the dark room with only faint red illumination (for avoiding unintended UV exposure). To prepare for the assay, I shifted the culture temperature to 25℃ and prolonged the culture time prior to UV to 72 hours, in order to maximise PCCys incorporation. Additionally, the worms were transferred to floxuridine-supplemented NGM plates just before

uncaging and left there till the microscopic scoring 48 hours later. The chemical prevents worms from laying eggs[418], so eventually only the original generation will be scored. Ailish Tynan assisted me in scoring neurons by fluorescence microscopy, which helps to minimise confirmation bias.

The gathered scores were categorised first by neuronal type then by experimental group and were sorted into the tables of Figure 3.21A. I also plotted the data as an accumulative bar diagram (Figure 3.21B) for presenting the proportions of cells in each grade (i.e., the graded morphological phase of the dying course). For each neuronal type, the bar of “+PCCys +UV” group is very different from the other two, which was then investigated in serial statistical tests as the following.

A

	Grades				
ALM	**0**	**1**	**2**	**3**	**4**
-PCCys -UV	0	0	0	0	24
-PCCys +UV	0	0	1	5	18
+PCCys +UV	13	2	4	4	1
AVM	**0**	**1**	**2**	**3**	**4**
-PCCys -UV	0	0	0	0	12
-PCCys +UV	0	0	0	1	11
+PCCys +UV	7	1	1	2	1
PLM	**0**	**1**	**2**	**3**	**4**
-PCCys -UV	0	0	0	0	24
-PCCys +UV	0	1	1	6	16
+PCCys +UV	19	4	0	1	0
PVM	**0**	**1**	**2**	**3**	**4**
-PCCys -UV	0	0	0	0	12
-PCCys +UV	0	0	0	2	10
+PCCys +UV	8	0	1	3	0

B

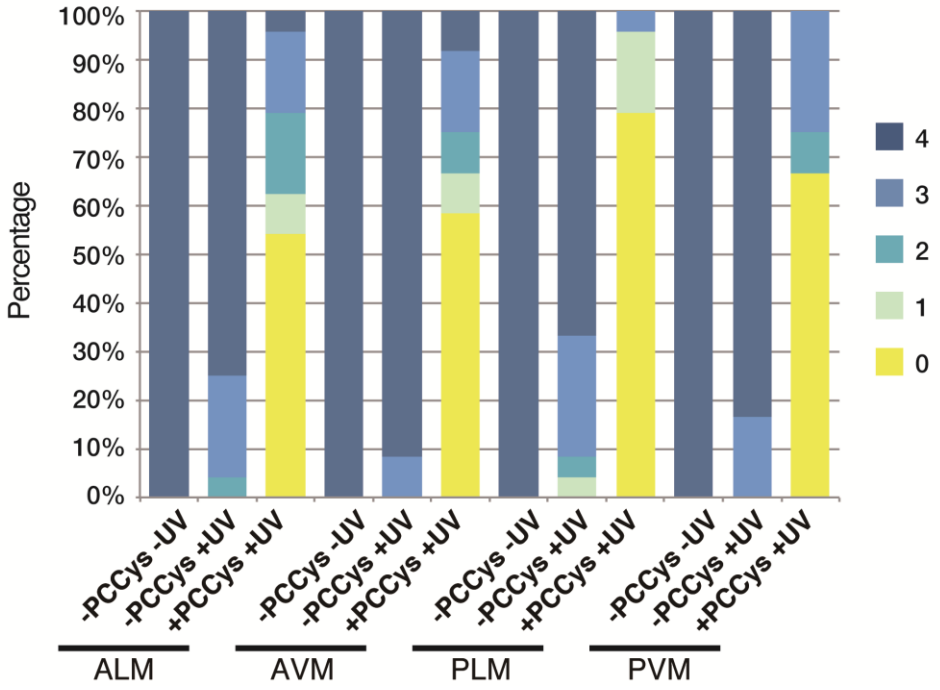

Figure 3.21 Distributional differences in the apoptotic grades of each neuronal type due to PCCys incorporation and light-activateable ablation.

(A) For all types of touch receptor neurons, the numbers of cells in each apoptotic grade (column) and each experimental condition (line) are counted and filled into the table. **(B)** Accumulative bar diagram showing the percentages of cells in each apoptotic grade (colour-coded) of each experimental condition, for each type of touch receptor neurons.

The data of each row of the summarising table (Figure 3.21A) was subjected to Shapiro-Wilk test with the normality assumption, i.e., the data distribution is hypothesised as Gaussian. All the p-values from these calculations are much smaller than the significance threshold 0.05, so the initial assumption (or, null hypothesis) should be rejected, suggesting me select from non-parametric tests. As more than two groups were to be compared simultaneously, Kruskal-Wallis test was chosen and individually conducted for each touch receptor neurons. All p-values obtained are smaller than 0.0001, so I performed post hoc Dunn's multiple comparisons to each two groups and the results are plotted in Figure 3.22. There are greatly significant differences between the PCCys-incorporated, UV-illuminated group and any of the other groups. Therefore, I have validated the efficacious ablation via the optical activation of PCCys-incorporated caspase also in mechanosensory neurons.

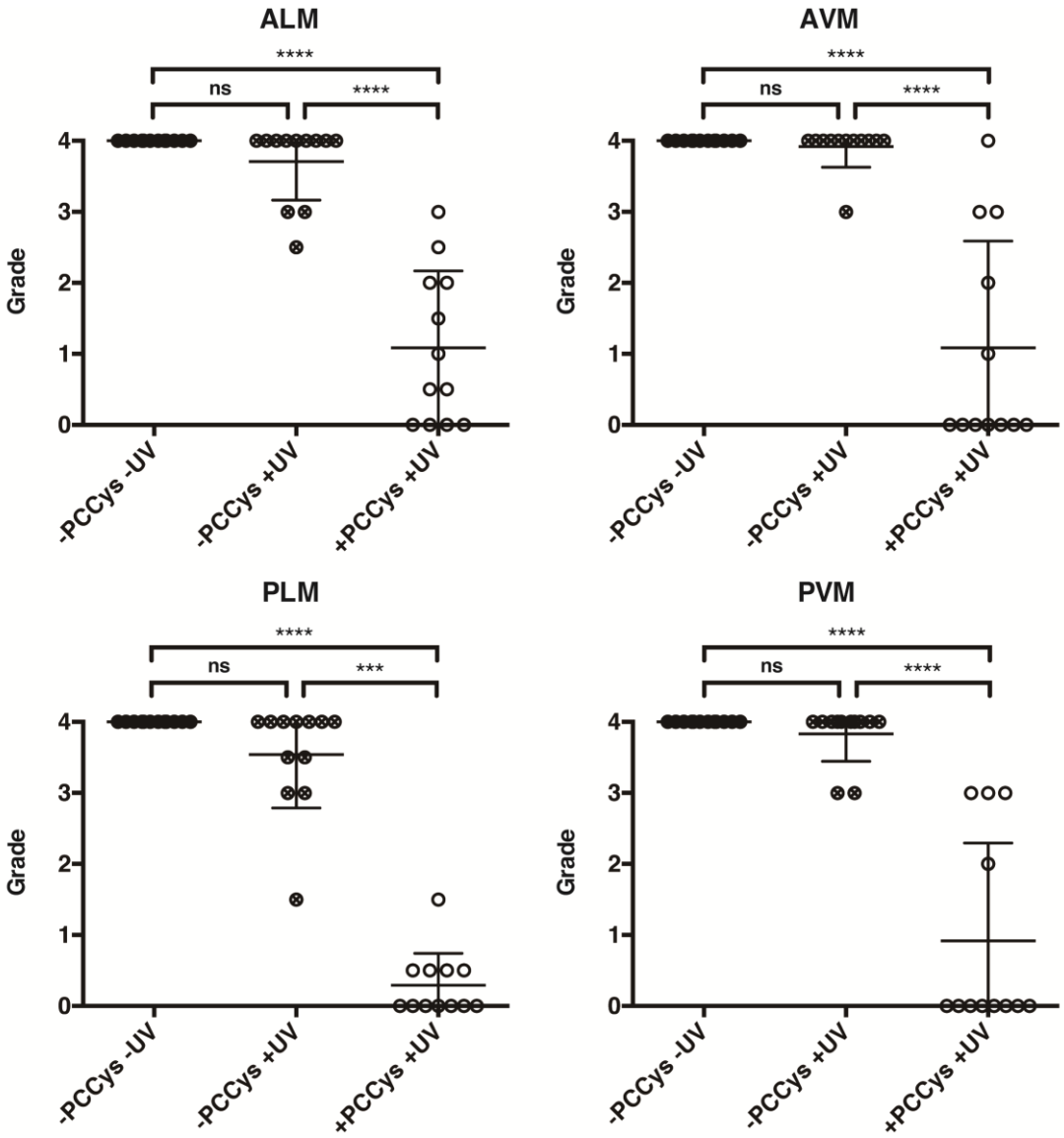

Figure 3.22 The PCCys-incorporated and UV-illuminated neurons are significantly different from the cells of other treatment groups.

For each neuronal type, the distribution of the grading data of each treatment group was analysed by Shapiro-Wilk normality test. Then data of all treatment groups were subject to Kruskal-Wallis test with post hoc Dunn's multiple comparisons. The sample sizes of all groups are 12. Error bars show standard deviation of the data set. “Ns” stands for no significance. *** is for a p-value < 0.001 and **** for < 0.0001.

To find out whether or not the type of touch receptor neurons contributes to the variability of the apoptotic manifestations, I used Friedman rank sum tests to examine the data of each experimental treatment. The results obtained no significance (i.e., p-value < 0.05), so the different types of touch receptor neurons, as long as treated by the same experimental procedure, are not distinct from each other in their morphological manifestations.

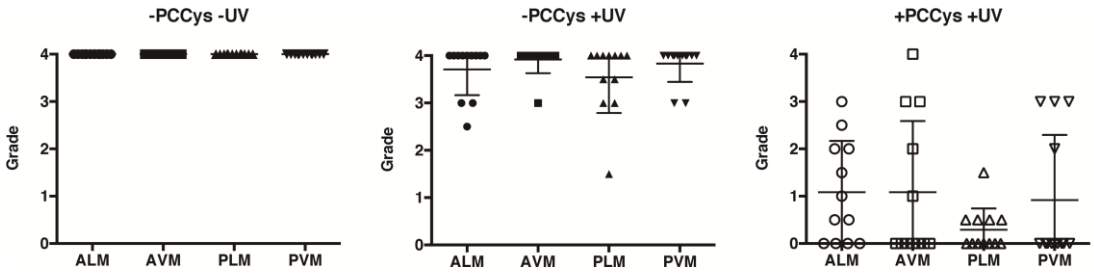

Figure 3.23 No significant difference exists in the apoptosis of different neurons.

For each treatment group, the grading data of all neuronal groups were compared altogether in a Friedman rank sum tests. The sample sizes of all groups are 12. Error bars show standard deviation of the data set.

3.2.2.2 Light activation of PCCys-caged caspase ablates body wall muscles

In the last section, I have demonstrated that with the prerequisite site-specific PCCys incorporation and optical uncaging, the caspase version of my design is capable of ablating two classes of neurons in living worms. To examine the reproducibility of this light-controllable ablation in tissue types other than neurons, I planned to express and test the engineered caspase in *C. elegans* body wall musculature.

Noting the improvement in PCCys incorporation and uncaging by implementing NES-PCCRS in the touch receptor neurons, I decided to try the two NES-attached variants of PCCRS in addition to the original version (as a control) in the first batch of bombardments. The promoter of gene *myo-3* for myosin heavy chain[419] was used for expressing the caspase and the orthogonal (NES-)PCCRS/$tRNA^{Pyl}_{CUA}$ pair in worm body wall muscles.

Table 3.3 Summarised results of PCCys incorporation and cell ablation of using different variants of orthogonal tRNA synthetase in body-wall muscle.

Plasmid details of the bombardment no. 225, 212 and 213 are in Appendix. Both types of the inserted nuclear export signal originate in human (Hu) proteins (PKIα

and Smad4)[282, 283, 401]. PylT stands for *Mm* $tRNA^{Pyl}_{CUA}$. Column “Incorporating Strains” shows numbers of strains that had PCCys successfully incorporated, and Column “Ablating Strains” are numbers of strains that displayed uncaging-dependent apoptosis. “Fraction of incorporating worms” is the proportion of worms that showed PCCys incorporation to the population gathered from randomly picking 15 worms of each strain from a bombardment (e.g., 60 in total for bombardment no. 225), whose numerator is used as the denominator of “Fraction of ablating worms”, which is the proportion of worms that showed cell ablation to the PCCys-containing worms that have been picked of each bombardment group.

No.	NES of PCCRS	tRNA	Transgenic Strains	Incorporating Strains	Ablating Strains	Fraction of incorporating worms	Fraction of ablating worms
225	/	PylT	4	4	4	12/60	6/12
212	Hu PKIα	PylT	10	8	8	71/120	54/71
213	Hu Smad4	PylT	10	9	9	43/135	31/43

From these bombardments I obtained a total of 24 independent strains. When supplied with 5mM PCCys, the majority of these transgenic strains displayed PCCys incorporation. And worms of every incorporating strain had dying muscle cells when checked under microscope approximately 1 day after 5-min whole-worm UV uncaging. Compared with the spindle-like healthy muscles, apoptotic muscle cells usually present themselves as ovals. Also, they contain cytosolic aggregations of fluorescent proteins. Sometimes

cavities would be observed in the original positions of muscle that has been already engulfed. Figure 3.24 includes the representative images under each experimental condition (namely, only PCCys supply, only UV illumination, and the two combined). In the presence of PCCys, many cells express mCherry in their nuclei. After UV uncaging, these muscle cells undergo programmed death and degrade into apoptotic bodies, debris or cavity at the time point of imaging. The ablation effect does not result from a direct optical damage, as the CFP-expressing muscle cells remain healthy after UV treatment.

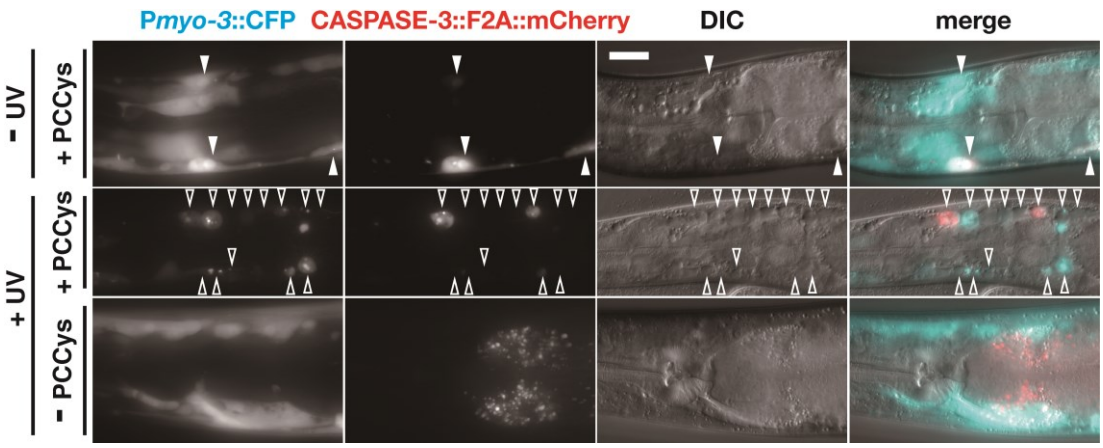

Figure 3.24 Muscle cells can be ablated by activated PCCys-Caspase upon whole-worm UV illumination.

Worms fed on PCCys without UV exposure have normal-shaped body wall muscle that show cytoplasmic CFP and nuclear mCherry (solid arrowheads). Subsequent to UV uncaging, muscle cells with PCCys incorporation (median row) present themselves as either apoptotic bodies or *in situ* cavity (located by hollow arrowhead).

In contrast, PCCys-incorporated cells without UV illumination remain intact (bottom panel). All images are in the same scale and scale bar = 20 μm.

The majority of the transgenic strains of all three bombardments (no. 225, 212 and 213) have shown PCCys incorporation, and all the incorporating strains exhibit cell removals upon “+PCCys, +UV” treatment. Here I am unable to conclude if the NES attachment to PCCRS increased the efficiency of PCCys incorporation in muscle cells. Therefore, I performed a brief count as below. From the PCCys-NGM plates of each bombardment groups, I randomly picked 15 worms of each strain onto the same plate and counted how many of them had the incorporation-reporting mCherry signal (denoted as m). I calculated the ratio of m to the number of all worms picked, to be the “Fraction of incorporating worms” of each group shown in the summary Table 3.3. Next I UV-illuminated each plate with these PCCys-incorporated worms for 5 minutes, and counted how many worms eventually had muscle cells ablated (denoted as n). For each bombardment group, a “Fraction of ablating worms” (in Table 3.3) was obtained by dividing n by m. As shown in the table, worm strains with no-NES PCCRS have the lowest incorporation fraction (20%), while the two groups of NES-PCCRS (bombardment no. 212 and no. 213) expression have their numbers higher than 30% and 50%. In terms of light-activated apoptosis, the no-NES group yielded a 50% rate from the test,

whereas the numbers for the NES-PCCRS groups are over 70%. This difference possibly results from an increased proportion of the orthogonal synthetases in the cytoplasm, poise for charging the orthogonal tRNA with PCCys. Subsequently, more worms of the NES-PCCRS groups had their incorporation levels above the thresholds of both detection and apoptosis initiation, eventually leading to more efficient cell ablation as I observed. In brief, NES attachment to PCCRS improves PCCys incorporation also in worm body wall muscles.

Among all strains from bombardment no. 212 and 213, BZ213.8.1 showed the highest efficiency of cell removal after UV treatment. Previously I always checked the uncaging results the next day of UV illumination experiments, while it might take shorter time in this strain for the optical-activated apoptosis to be observable. To document the apoptotic progression of worm body wall musculature, I performed 5-min UV illumination on the entire body of the PCCys-incorporated worms of BZ213.8.1 and then imaged a batch of typical worms within the population at different time points (Figure 3.25).

Without exposure to the uncaging light, muscle cells of animals grown with PCCys appear normal in their morphology and mCherry proteins are homogeneously expressed in cell nuclei. In comparison, body wall muscles of all post-UV groups show morphological alterations, especially the

redistribution of mCherry from cell nucleus to cytosol as to indicate the dismantlement of nuclear envelope[420]. Ten minutes after UV treatment, a very bright punctum of CFP along with several mCherry puncta were seen in the nucleus of the marked cell. Although such protein aggregations were found at all time points after uncaging, the amount seems to peak from 4 to 6 hours post-UV, for which I do not have an explanation yet. Morphological changes of muscle cells became evident as early as 1-hour post UV, as the cell emphasised in the images of the group (third row of Figure 3.25) has shrunk from its original spindle shape to oval. At the 1-hour point after uncaging, the caspase-initiated breakdown of intracellular structures has further progressed and the nuclear membrane appears to have collapsed[420], allowing the mCherry fluorescence to evenly distribute within the cytoplasm. On a DIC image, the apoptotic cell is visually separated by a cavity from the adjacent cells, which may result from the decreased density of its cytoplasmic contents. For some muscle cells in the PCCys-fed animals, up to 2-hours wait after uncaging is not sufficient for apoptotic changes to become observable, which rarely happen in groups of longer incubation time. Nevertheless, the cells marked in the images of the 2-hour group had high mCherry expression and also appeared very deformed. Four hours after UV illumination, I could not find in apoptotic cells any clear membranous structure for bordering the nucleus, so at this time point cell nucleus has already

disintegrated. Besides, mCherry fluorescence is only observed as puncta that all co-localised with the brightest CFP dots, which possibly signals the degradation of apoptotic cell contents within the phagolysosomes of engulfing cells[313]. Similar effects were seen in more cells in worms imaged 6-hr and 8-hr post UV. In addition to the cells with amorphous cytoplasmic contents, large cavities were observed on many sites of the body wall musculature of UV-uncaged worms after 10-hour incubation. As these cavities show smooth boundaries in the DIC channel and lack signal in the channel of either CFP or mCherry, the apoptotic muscles may be fully degraded within adjacent engulfing cells. This suggests that 10 hours is adequate for the muscle cells to be completely ablated through the apoptotic pathway mediated by my light-activatable caspase variant. More vacuoles appear in replacement of muscle cells in worms imaged 1-day and 2-day post UV. There are still some cells present with normal appearances after 2 days. They all show CFP fluorescence but no mCherry signal. It appeared to me that they have not expressed caged caspase and have therefore not been affected by the uncaging treatment. Such different fate of indivivual muscle cells reflects the high variability of the incorporation efficiency that is possibly due to the extrachromosomal presence of transgenes, and the continuum of cell manifestations to the same optical uncaging ends where some cells do not contain enough incorporated enzymes to initiate apoptosis.

Compared with the tested neuron classes, cells of body wall musculature seem to have particularly heterogeneous responses to the experiments of PCCys incorporation and UV-uncaging. Two factors possibly contribute to this difference: 1) a muscle cell is composed of many sarcomeres and much larger than a neuron, which requires more caspases to be incorporated and uncaged for executing apoptosis; 2) the refraction and absorption along the optical path of ultraviolet light is posture-dependent and the uncaging effects of the top illumination of the UV crosslinker may vary in the body wall muscles on the top side from those on the bottom.

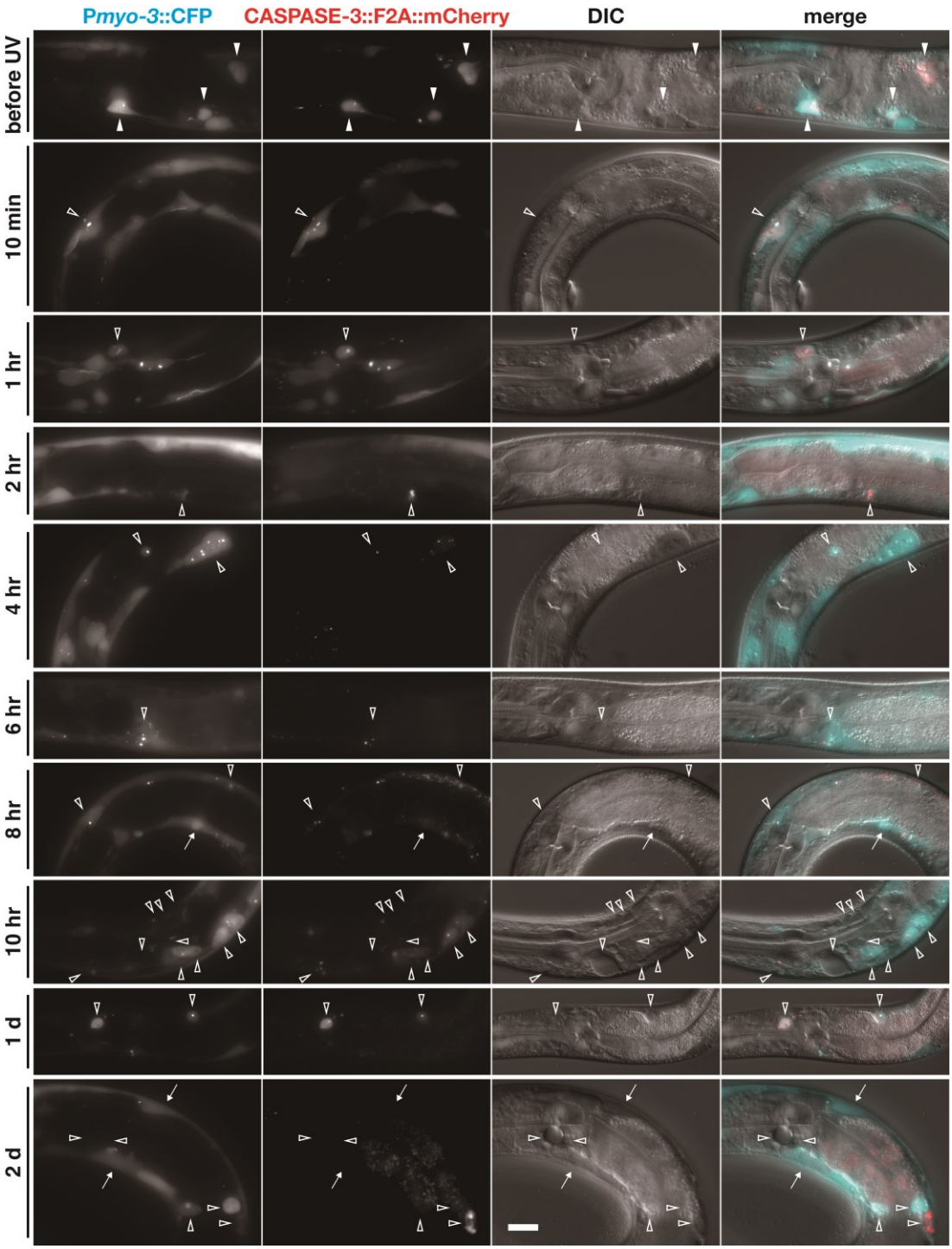

Figure 3.25 PCCys-Caspases activate and mediate apoptosis of muscle cells within hours after UV uncaging.

PCCys-incorporated worms were imaged under fluorescent microscope at different time points before and after whole-worm UV illumination. Except the negative

controls (top row), the body wall muscles of the UV-uncaged animals show apoptotic signs (hollow arrowheads) for all lengths of time post UV (all rows below the top one. hr, hour. d, day). Arrows mark the light-target cells with no PCCys incorporation. All images are in the same scale and scale bar = 20 μm.

In brief, optical uncaging of PCCys-incorporated Caspase-3 is capable of ablating body wall muscles efficiently: the visual sign of apoptosis arose in body wall muscle as fast as approximately 10 minutes after UV uncaging and the most rapid cell removal only took 10 hours.

I then proceeded to optimise the ablation procedure. So far the whole-worm uncaging experiments all set the UV illumination for 5 minutes. However, such a length of illumination is too long for the anticipated single-cell uncaging experiments in which the use of microscope-mounted UV laser requires me to anaesthetise worms. Therefore, here I would like to first examine whether or not shorter illumination times would be sufficient to activate the caged caspase. BZ213.8.1 worms were cultured on 5mM PCCys plates and then UV illuminated for 0, 6, or 30 seconds. After one-day incubation, worms of all these groups were imaged by fluorescent microscope.

As shown in Figure 3.26, unexposed worms have PCCys-incorporated muscle cells in normal shape, while the body wall musculature of all illuminated worms has displayed apparent morphological and fluorescent alterations due to apoptosis. Many cells have been fully ablated in worms of 6-s and 30-s groups, and the few remaining CFP-expressing cells have no detectable mCherry signal of PCCys incorporation. As most of the cells of the UV-illuminated groups have been replaced by the cavities from apoptosis, no significant difference exists in ablation efficacy for the various illumination times. Therefore, UV illumination time down to 6 seconds is sufficient to activate the caged caspase to initiate apoptosis.

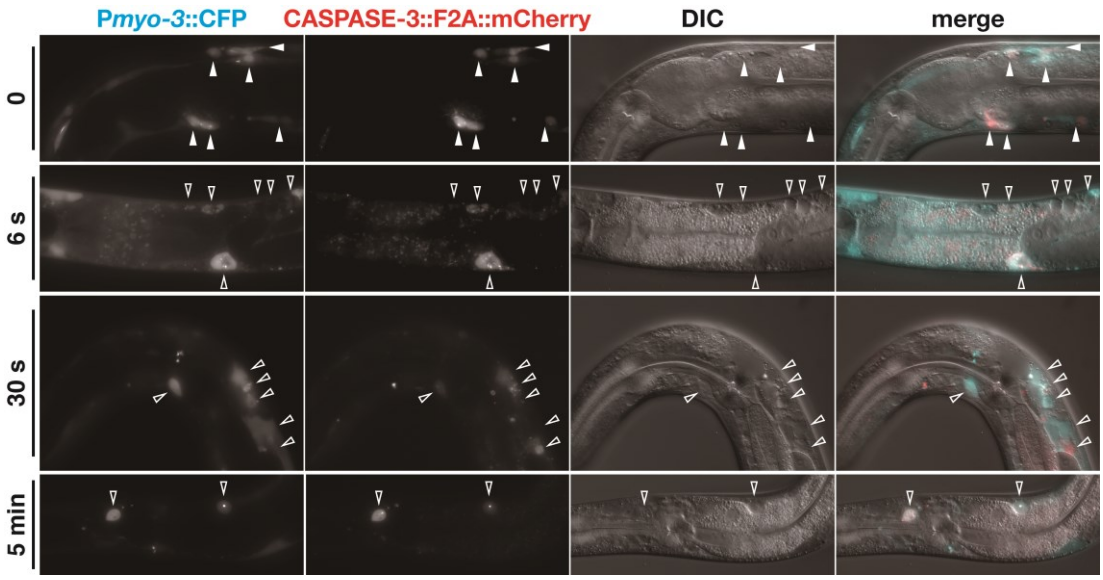

Figure 3.26 Short-time UV illumination on the entire worm is sufficient to activate the PCCys-Caspase for ablating body wall muscles.

PCCys-feeding worms were imaged after being illuminated by UV for 0, 6 or 30 seconds, and the results are compared with images of the worms after 5-minute illumination (the same set of images on the second last row of Figure 3.25). The negative control group (unexposed to UV, top row) worms have intact CFP-expressing muscles, among which some also show red signal of mCherry mainly in their cell nuclei. Instead, worms got illuminated for 6s, 30s and 5 min (the last three rows) display dying and dead muscle cells (hollow arrowheads) from apoptosis, and the severity of cell loss are comparable. All images are in the same scale and scale bar = 20 μm.

3.2.3 Proof-of-principle behavioural experiments with the ablation of touch receptor neurons

The effectiveness of the light-activatable Caspase-3 have been confirmed in neuronal and muscular tissues, so I proceeded to the proof-of-principle experiments for the applicability of my tool in future neurobiological research of *C. elegans* behaviour.

As the name suggests, touch receptor neurons are a class of mechanosensory neurons responsible for sensing soft mechanical forces

exerted externally on the worm cuticle[415]. Their dendrites, being surrounded by mantle (a kind of extracellular matrix) and then engulfed by hypodermal cells, accommodate characteristic 15-protofilament microtubules assembled from tubulins (namely the tubulins MEC-7 and MEC-12)[411-413, 421]. The cuticle area being innervated by its dendrites is the receptive field of one such neuron[415].

I used an established assay for gentle touch[421, 422] to quickly test whether or not the light-activated apoptosis of touch receptor neurons impairs the touch response of these six neurons. These preliminary experiments were performed in a quantitative way, briefly described as the follows (more details in Chapter 4.16). Worms of Strain BZ211;CZ10175 were grown for 3 days either in the presence or absence of 5mM PCCys. They were then divided into two subgroups: worms of one group were uncaged by UV for 5 minutes while worms of the other group were kept in the dark. Two days after illumination, I randomly selected 10 animals for each of the three experimental treatments (namely “-PCCys, -UV”, “-PCCys, +UV” and “+PCCys, +UV”). Additionally, I set a control group of 10 CZ10175 worms that had been cultured in the same condition as the “-PCCys, -UV” group. After acclimatisation, all animals were touched by an eyebrow hair alternatingly at the head stroke site and tail stroke site (as illustrated in Figure 3.27) for a total

of 10 times. The forces exerted were slightly above 10 μN to ensure that only the soft touch receptor neurons would be activated, without any of the harsh touch response stimulated[423] (means of the force measurement is described in Chapter 4.16). A reversal after a head touch or acceleration after a tail touch was counted as one "response". The numbers of responses were recorded separately for head and tail touches as the scores. One complete set of touches was performed on each worm before moving on to the next worm of the treatment group. The counting data was then plotted and analysed statistically (Figure 3.28 and 3.29).

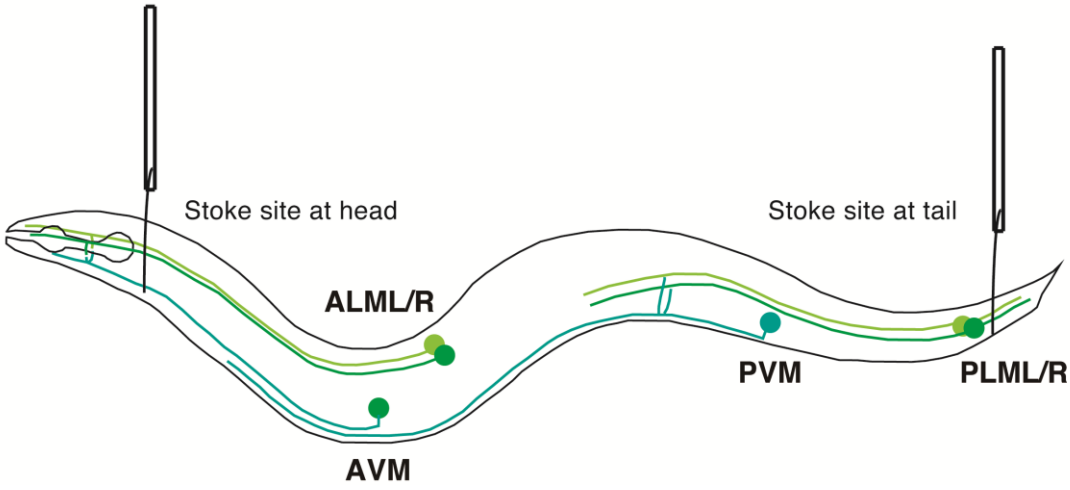

Figure 3.27 Illustration of the stroke points of the soft touch assay.

Each touch requires the hair shaft of the touching tool to sweep perpendicularly to the body axis across the entire side of the worm cuticle[423]. The stroke site for head touches is slightly posterior the terminal pharynx bulb while the site for tail touches is just anterior to the anus[423]. The head touch stimulates both the ALM

pair and AVM, while tail stroke stimulates the posterior touch receptor neurons.

The four sets of head scores of all groups, namely “CZ10175”, “-PCCys, -UV”, “-PCCys, +UV” and “+PCCys, +UV”, were examined for normality by Shapiro-Wilk test. All the p-values are below 0.05, meaning that no set of data follows a normal distribution. Thus I used Kruskal-Wallis test, the non-parametric version of ANOVA, to examine if there is significant variation among the four data sets. The resultant p-value is above 0.05, so the groups are not statistically different from each other in head-touch response (Figure 3.28). The tail-response data of these experimental groups were also put through Shapiro-Wilk normality test and Kruskal-Wallis analysis of variance. The latter test gave a p-value of 0.0038, which indicates significant difference in the tail-touch responses resulted from different experimental conditions of the groups. To figure out which experimental condition underlies the statistical difference, I then conducted multiple Dunn's post-hoc tests comparing every two groups of the four. Among the resultant six p-values, only the one of comparing “-PCCys -UV” and “+PCCys +UV” is adequately small (0.0021), meaning that significantly different responses to tail touch exist between the untreated worms and illuminated, PCCys-supplied worms. The insignificance between “+PCCys +UV” and “CZ10175” or “-PCCys +UV” possibly result from a compromised statistical power of using inadequate numbers of sample

animals. Also, the worms were randomly divided into the group for UV illumination and randomly picked for the touch assays and therefore some might retain functional touch receptor neurons from an incomplete optical activation of caspases. As presented in Figure 3.21, it is only in the “+PCCys +UV” group but not any other group that I have observed animals of a score of 0. Hence, the protocol of this experiment could be further improved by picking neuron-lost animals from the “+PCCys +UV” group under fluorescent microscope for the touch assay.

Nevertheless, the analysis demonstrates that none of the genetic manipulation and the experimental procedure (except the simultaneous action of PCCys supplement and UV illumination) on the worms affects their response to gentle touch. At least for touches at the tail, impaired behavioural outputs can be statistically related to PLM ablations. This fits with the previous observations during the apoptosis assay (in Chapter 3.2.2.1.2) that PLM neurons were ablated the most easily after whole-worm uncaging.

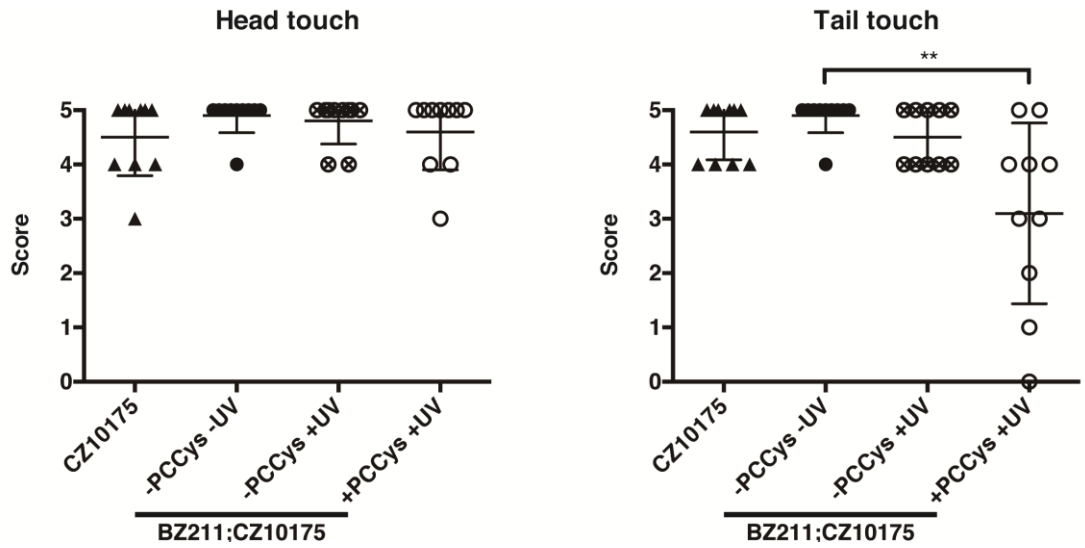

Figure 3.28 Significantly different reactions to gentle touch in worm groups without and with light-activated apoptosis.

Strain CZ10175 was used as the negative control of the response to gentle touch both at the head and tail of an animal. The worms used in all other groups were of Strain BZ211;CZ10175. The sample sizes of all groups are 10. The distribution of the scores (representing the numbers of normal responses) of each experimental group was analysed by Shapiro-Wilk normality test and then all data sets of either head touch or tail touch were subject to Kruskal-Wallis test with post hoc Dunn's multiple comparisons. Error bars show standard deviation of the data set. ** is for a p-value < 0.01.

In order to investigate whether or not there is systematic difference in the sensitivity of these worms to head touches and tail touches, I also chose Wilcoxon matched-pairs signed rank test to compare the scores of touches at

both sites for each treatment (plotted as Figure 3.29). Corresponding to the results above, significant difference (p-value < 0.01) only exists between the head- and tail- touching in the group of UV-illuminated PCCys-incorporated worms.

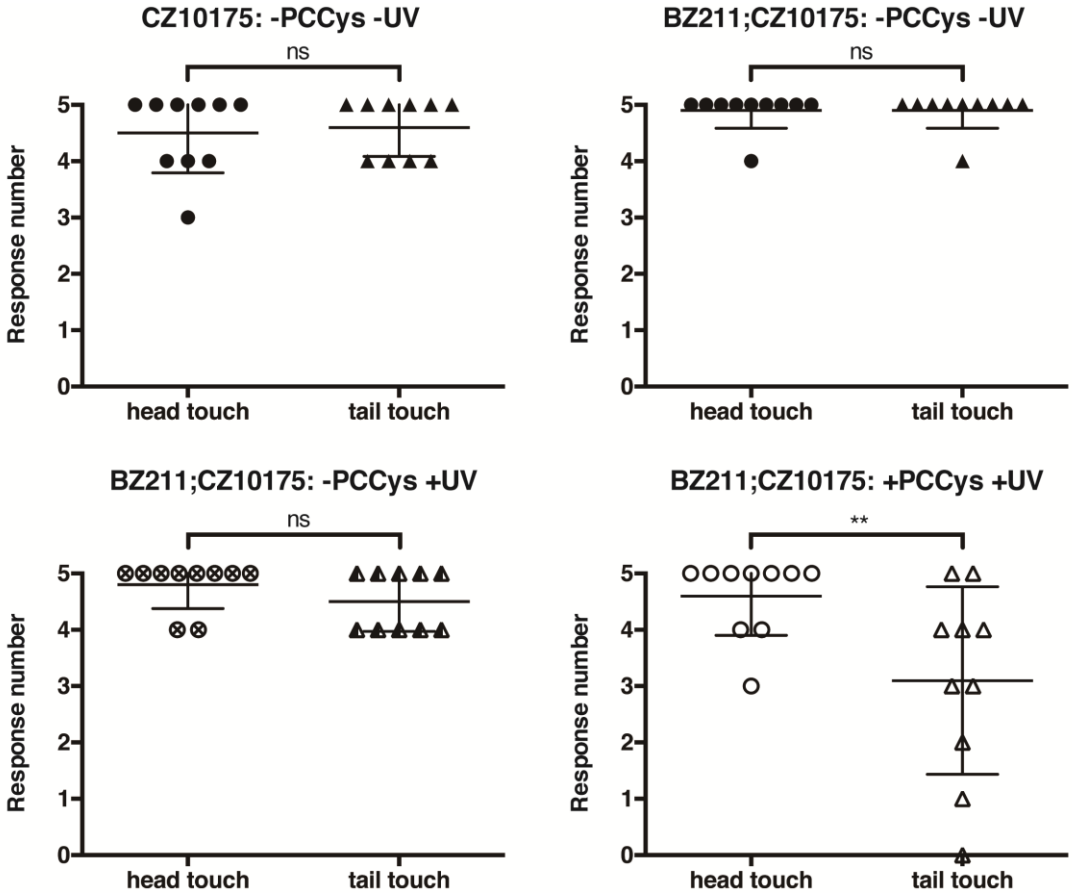

Figure 3.29 The PCCys-incorporated and UV-illuminated worm response differently to the gentle touches at different position.

The same data as in Figure 3.28 was analysed according to each strain and experimental condition for comparing possible discrepancy in the respective scores of head touch and of tail touch. The non-parametric Wilcoxon matched-pairs signed

rank test has been adopted. Error bars show standard deviation of the data set. “Ns” stands for no significance and ** is for a p-value < 0.01.

By combining the results from the touch assays with apoptosis assays in touch receptor neurons, I found a discrepancy between effective ablation of the neurons (especially ALM and AVM) and the observed behavioural outcomes. I noticed that the processes from both AVM and one ALM innervate the head stroke site, whereas for only one PLM neuron would receive mechanical force applied to the tail stroke site[415] (Figure 3.27). In terms of the relationship between neuron apoptosis and mechanosensation, I speculate that only when both the top-side ALM and AVM have been well removed, would the worm fail to react to head touches. In contrast, it seems that to fail the behaviour response to tail touch only requires the ablation of the top-side PLM cell. Besides, the result from my previous apoptosis assay suggests that AVM and one neuron of ALMs are less likely to die as compared to the PLM neurons, which may explain the observation that I could only detect an apoptotic effect on the PLM-mediated tail-touch response. In other words, I may have to ablate at least one ALM as well AVM to eliminate the head-touch response, so only few animals of “+PCCys +UV” group will lose response to head touch. Contrarily, to cancel the reflex to tail touch possibly only requires the removal of one PLM, and PLMs can be relatively

easy to ablate. Altogether this may explain why I could not statistically detect a corresponding behavioural alteration to the light-activation of caspases in anterior touch receptor neurons with the 10-worm sample size. In the meanwhile, the behavioural effect of PLM ablation is probably strong enough to overcome the slight compromise of sensitivity of the gentle touch assay.

In conclusion, the proof-of-principle mechanosensory assay of touch receptor neurons has validated the two assumptions I raised: 1) in the absence of PCCys, Strain BZ211;CZ10175 that expresses the engineered caspase has no difference in responding to gentle touch from the N2-based background strain CZ10175 (the negative control), whether or not being illuminated by 365-nm UV in the whole-worm mode; 2) the optically activated apoptosis following PCCys incorporation and UV uncaging of touch receptor neurons leads to an impaired avoidance to the site of a gentle touch.

3.2.4 High-precision laser-mediated activation of caspase in single neurons

The success and effects of ablating specific groups of cells using my light-activatable caspase have been demonstrated in the previous sections.

To reach a single-cell precision as in my objective, I planned to perform uncaging experiments on individual neurons using a microscope-mounted UV laser (MicroPoint). The laser generator has the energy of its pulse adjustable within a 90-step (from 0.1% to 100%). Energy 27-29 have been shown sufficient for uncaging PCLys in cells of living worms, while barely bleach the fluorescence in the illuminated cell (Lloyd Davis, personal communication). I would therefore use the laser pulses within this energy range for activating caged-caspase in individual neurons. I also designed 365-nm laser-uncaging experiments on cells in the absence of PCCys to test the physiological tolerability of the energy range.

As shown in Table 3.4, whole-worm uncaging via UV crosslinker is compared with other two global uncaging methods (Methods B[241] and C[240], while the technical parameters of UV-laser uncaging via MicroPoint and two other laser microbeam ablation methods (E and F[173]) are placed together for comparison. Figures 3.6-3.8 and Figure 3.24 in previous sections demonstrate that global illumination of 365nm light can be physically benign to the cells and animals targeted, if the total energy delivered is not high. This is consistent with the results from the two instances of globally uncaging mammalian cells by 365nm illumination of similar intensities[240, 241]. The pulse-energy measurements of the modes (Energy 27-29) of the

microscope-mounted UV laser I would use, are much lower than the level of the cell-damaging Methods E, by at least 5 orders of magnitude[173]. Though the single pulse energy of femtosecond laser (Method F) is quite low, it is fired in a minimal frequency of 1000 Hz, which gives a power of 30 μW[173], which is more than 10 times higher than the power of my UV laser. The femtosecond laser has been adopted in laser axotomy with relatively less damage to surroundings tissues reported, so the 365-nm laser set for my uncaging experiments is very unlikely have detrimental effects on the cells illuminated.

Table 3.4 Technical parameters of illumination for NCAA uncaging and laser surgery in eukaryotic cells and *C. elegans*.

Among the listed methods, A and D are what I conducted and referred to elsewhere respectively as “whole-worm ablation” and “single-cell ablation/UV laser ablation”, and others have been described in the literature (B: by Nguyen *et al.*[241]; C: by Ren *et al.*[240]; E and F: Fang-Yen *et al.*[173]). “*In vivo*” means the illumination is on cells of anaesthetised worms, while “*in vitro*” refers to the illumination on mammalian cell cultures. Methods A-D all use 365nm light. The parameters of the listed three modes of Method D are calculated from the maximum pulse energy (50 μJ) and average power at 15 Hz of the laser apparatus (calculation details in Chapter 4.14.1). The energy figures of Method D-F are of a single pulse. Blanks are left for the

unavailable or unnecessary parameters.

Methods	Apparatus	Mode	Energy (μJ)		Average power (μW)	Intensity (μW/cm^2)
A	UV lamp	In vivo, whole-worm				5000
B	UV lamp	In vitro				4000
C	UV lamp	In vitro				600
D	Microscope-mounted nanosecond UV laser (MicroPoint)	In-vivo, single-cell	Range 27	0.25	3.72	
			Range 28	0.27	4.02	
			Range 29	0.29	4.35	
E	Microscope-mounted nanosecond laser	In vivo, single-cell		30000		
F	Microscope-mounted femtosecond laser	In vivo, single-cell		0.03	30	

My laser-uncaging trials were then conducted in two types of neuronal cells, namely oxygen-sensing neurons and touch receptor neurons, still with the use of the promoter of *gcy-32* and the promoter of *mec-7*, respectively.

3.2.4.1 Oxygen-sensing neuron

Strain BZ52.2.8 worms were cultured on NGM plates with or without a supplement of PCCys (5mM) to the L4 stage. Then I mounted them onto slides, placed the slides under fluorescent microscope, and illuminated the

target neurons in worms of the UV-positive groups by repetitive 365-nm pulses (approximately 0.4 μJ per pulse). The worms were imaged after one-day recovery on NGM plates, and Figure 3.30, 3.31 and 3.32 exhibit the consequent morphological states of PQR, AQR and URX neurons, respectively.

PQR was the first oxygen sensing neuron I tried to target, as among the four neurons this cell is the most isolated and hence the technically easiest to illuminate. I included three control groups for the trials below: one group of worms had no exposure to PCCys or UV light, one group was supplied with the amino acid but the PQR of the worms were unexposed to UV, and the third group had no PCCys access but would be UV illuminated. Meanwhile in the microscopy I would still image the untargeted O_2-sensing neurons (i.e., AQR and URX) of the same worms as the internal controls in the experimental group.

As I can observe from the images of the top two rows in Figure 3.30, nuclear mCherry expression reflects the incorporation of PCCys in PQR, which is consistent with the previous incorporation results and has not been influenced by the change in the anaesthetising means coupled with the single-cell uncaging mode. UV-illuminated PQR was nearly completely removed in

apoptosis, leaving debris of its soma and process that emits fluorescence in both CFP and mCherry channels.

The image panel below displays the other three neurons (AQR and two URXs) in the head of the same worm as a comparison. These cells express the incorporation marker mCherry in the nuclei while remain intact as they have not been illuminated by UV laser. I have also excluded the possibility that UV light alone damages the neuron, because in the absence of PCCys supply the PQR stays normal after the same uncaging treatment.

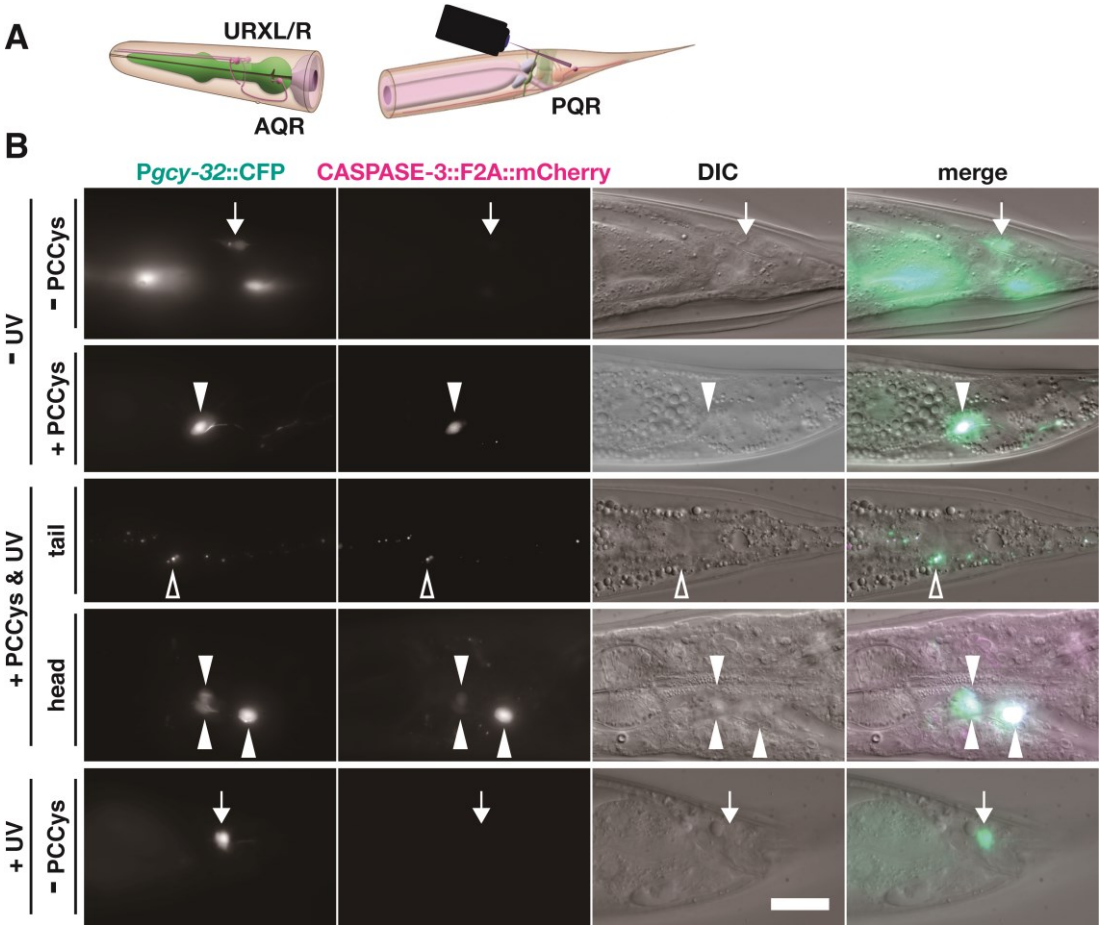

Figure 3.30 Neuron PQR killed by activated PCCys-Caspase upon UV laser illumination.

(A) Schematic of UV-laser specifically illuminating PQR among O_2-sensing neurons in *C. elegans*. Adapted from WormAtlas[264]. **(B)** Without being illuminated by UV, worms from both NGM group (top row) and PCCys-supplied group (second row) have soma and processes of PQR intact. In the meantime, those supplemented with PCCys lose PQR (third row) but not all other O_2-sensing neurons (fourth row), after UV illumination specifically on PQR. Mere the laser illumination conducted does not kill the target, as a healthy non-incorporated PQR is observed by CFP fluorescence after uncaging (bottom line). Arrowheads point intact PQRs without PCCys

incorporation, solid arrowheads mark PCCys-incorporated normal neurons and hollow arrowheads locate the apoptotic remnants of PQR. The second row images were taken from the worm of which URXs were specifically illuminated while PQR remained unexposed. All images are in the same scale and scale bar = 20 μm.

Following the successful uncaging and activation of the photo-caged Caspase-3 in PQR, I performed the experiment on the head neuron AQR. As URX neurons are located in the same region as of AQR, when I searched for AQR through the ocular, URXs were often seen within the field of illumination (with the usage of the 63x objective lens). Consequently, there might be a possibility that off-target uncaging also occurs in URXs during the experimental procedure, which might impair the practicability of the caspase tool.

The results displayed in Figure 3.31 relieve this concern, as I could detect no undesirable side effects in the neighbouring O_2-sensing neurons of AQR after the single-cell UV illumination. Optical uncaging completely removes AQR, while both URX cells nearby are present and one of them clearly show PCCys-incorporation marker mCherry in the nucleus. PQR of the same worm is also unaffected by the procedure, as its CFP and mCherry fluorescence can be observed in the respective channels. Besides, in the images of

unexposed worms, AQR neurons express mCherry only in the presence of PCCys. As I hurried to test whether or not UV laser uncaging was able to uncage only one cell of the neuronal pair URX, I did not set a control group ("-PCCys, +UV") to confirm that illumination from the UV laser alone does not have lethal effects on AQR, for which I should have used 365-nm laser to illuminate AQR of worms grown in absence of PCCys. Again, in the end I was not able to conduct the plan of a "formal" AQR ablation experiment with all the control groups, due to unexpected quarantines during the global pandemic.

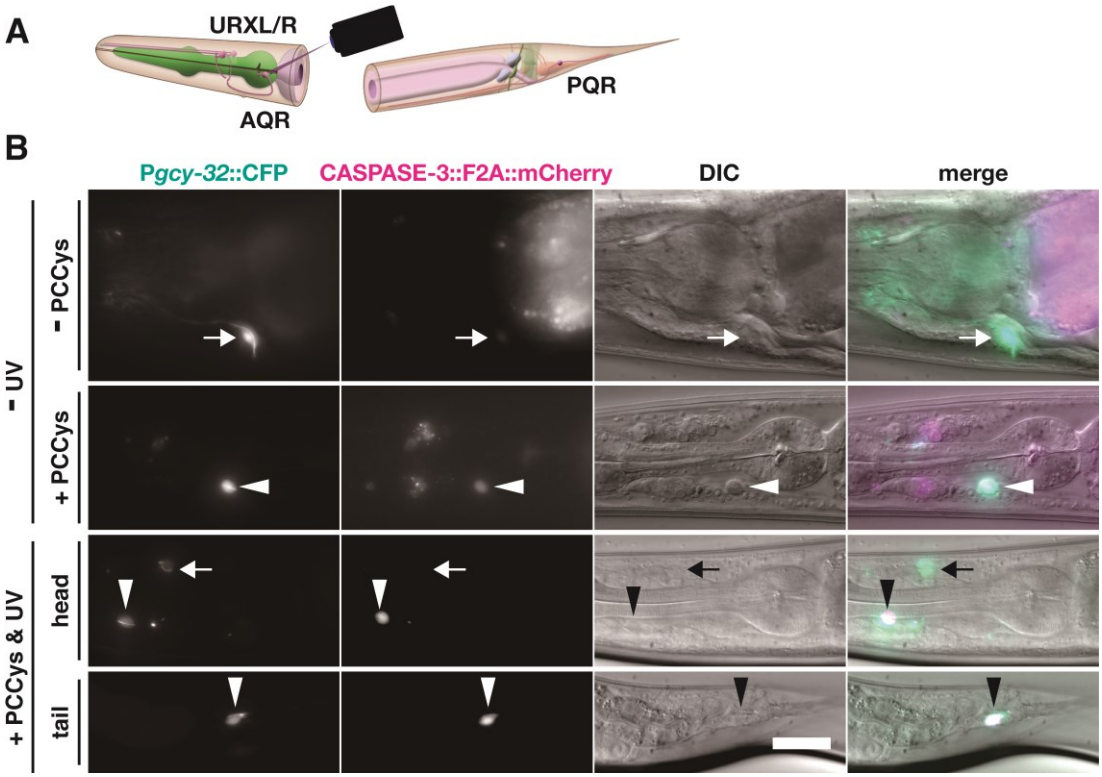

Figure 3.31 Neuron AQR killed by activated PCCys-Caspase upon UV laser illumination.

(A) Schematic of UV-laser specifically illuminating AQR among O_2-sensing neurons in *C. elegans*. Adapted from WormAtlas[264]. **(B)** The top half shows worms grown without UV exposure either on NGM (first row) or PCCys-NGM (second row) plates. Strong mCherry signal is only detected in the PCCys-supplied groups (the bottom three rows). After AQR-specific UV uncaging, only AQR gets ablated while two URX neurons also in the head (third row) and PQR neuron in the tail (bottom row) are intact. Arrowhead marks healthy neurons with PCCys incorporation and arrows point the cell that does not clearly show incorporation. All images are in the same scale and scale bar = 20 μm.

The task of uncaging one URX cell of the neuronal pair is the most challenging among single-cell ablation of any O_2-sensing neurons, particularly in nematodes lying on the side (the most usual and the natural posture for the animal) when mounted, because the left-right symmetrical neurons would have an overlap on the x-y plane, the exact plane of my visual field where I select the region to deliver the laser pulses. I attempted to address this problem by gently moving the coverslip to roll the worm underneath, however, this operation usually press a worm enough to burst meanwhile can not improve different postures of several worms altogether mounted on the same slide. Then I found better means as listed below, and their combination in

experimental practice has allowed me to perform the ablation of one cell of a neuronal pair in adult animals (shown as Figure 3.32):

1) In the many cases that two URXs are not entirely superimposed, I targeted the cytoplasm of the target outside the overlap. Activated caspases (of an adequate amount) initially at any part of the cytoplasm are theoretically able to execute apoptosis of the entire cell, so this method requires a sufficient size of the target area and abundant caspases with PCCys incorporation.

2) In the case of nearly complete superimposition, I adjusted the focal plane of the UV laser, restricting it in the z-axis range of the target neuron meanwhile further away from the other neuron.

I conducted laser-uncaging experiments on the L4 worms that have been cultured in either presence or absence of 5mM PCCys since their L1 stage. The “+PCCys +UV” worms was further divided into two subgroups, labelled as “aim at both” in which both URXs were uncaged by UV laser, and “aim at single” in which I would only fire laser at one URX. Also, I use the laser to illuminate both URX cells of “-PCCys +UV” worms. For each “-UV” control group (“-PCCys -UV” and “+PCCys -UV”), the slide of worms was still replaced on the microscope stage, and the URXs were centred in the field of vision for the same amount of time used for operations on a “+UV” group, but with no pulse fired at the neurons.

After I specifically illuminated two URX neurons, both cells underwent apoptosis (Figure 3.32): when I imaged the worm, the upper cell in the image existed as a disk-like apoptotic body (especially evident in the DIC channel), while the lower one had been entirely removed and was undetectable in any of the channels. In contrast, AQR and PQR of the worm remained normal (Row 5 and 6 of Figure 3.32B), indicating the whole experiment had been well performed with no collateral damage. The successful ablation of both URXs is also the positive control for the next single-cell ablation trial. In comparison, after I UV-illuminated one of the two URX cells, I found only the targeted URX cell got ablated, meanwhile the other URX and AQR remained intact in the head of the animal (see images of “aim at one”). Since both URXs are removable by the optical uncaging method, the presence of the untargeted URX is unlikely to result from failure of the uncaging procedure. The UV laser I used is of the safe intensity, as the URX neurons without PCCys incorporation are not affected.

Up to this stage, by the illumination technique that combines sub-cellular area selection and focal plane aim, I am able to ablate an individual URX without affecting the presence of the nearby oxygen-sensing neurons.

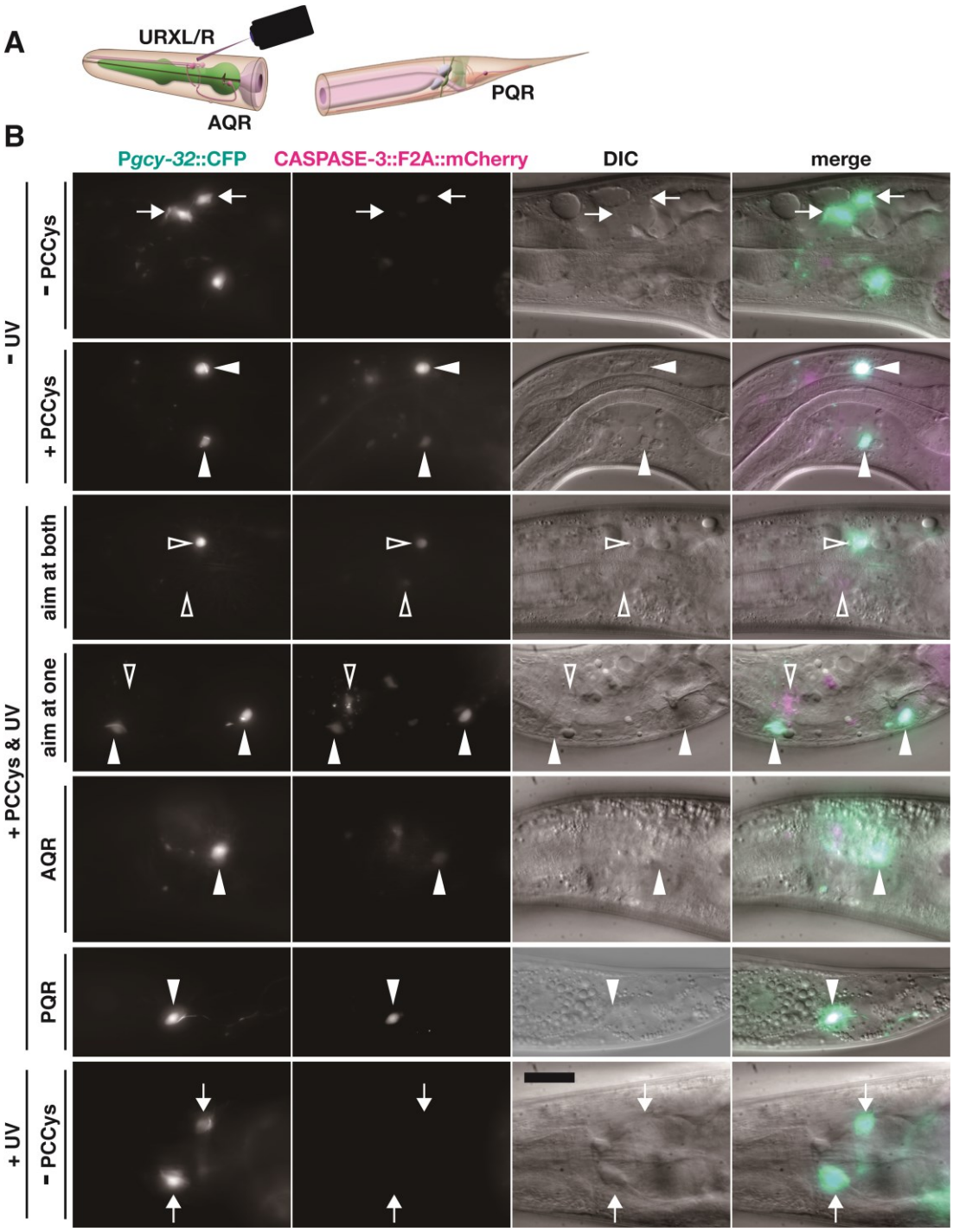

Figure 3.32 Neuron URX killed by activated PCCys-Caspase upon UV laser illumination.

(A) Schematic of UV-laser illuminating the pair of URX among O_2-sensing neurons in *C. elegans*. Adapted from WormAtlas[264]. **(B)** Without the illumination of UV laser,

worms either on NGM plates (top row) or on PCCys-supplemented plates (second row) have both URX neurons present. Optical uncaging at the two PCCys-incorporated URXs activated apoptosis of both cells (third row). Contrarily, aiming UV laser at a single URX only ablated this one neuron of the URX pair (fourth row). Within the same animals, AQR and PQR neurons are not affected (fourth to sixth rows). The implemented energy of the UV laser poses no harm to cells without PCCys incorporation (bottom row). Arrows mark intact neurons that did not incorporate PCCys and solid arrowheads mark the cells with PCCys incorporation. Hollow arrowheads show where the removed URX cells were originally located. All images are in the same scale and scale bar = 20 μm.

Altogether I have conducted UV-laser illumination on more than 100 worms. Although these uncaging experiments were not designed and performed with rigorous technical and physiological controls, with the current results I could still figure out some empirical rules as below. The worm anaesthetising time is an important factor to the survival rate in the post-uncaging recovery of mounted, illuminated worms. In retrospect, I found that only 57.7% (i.e., 40 of 75) animals got recovered after more than 20 minutes' immersion in NaN_3, while the fraction rises to 69.0% (i.e., 20 of 29 worms) for trials with shorter anaesthesia (Table 3.5). Both the rate of successful cell ablation (reflecting the efficacy) and the rate of control cells unaffected (reflecting the specificity)

are consequently raised by this higher chance of worm survival. An even higher recovery rate with a short anaesthesia (of approximately 10 minutes) has been observed in PCLys-uncaging trials (Lloyd Davis, personal communication). Above observations also agree with the rule of thumb in literature that relatively short anaesthesia (≤ 15 minutes) results to better recovery[173]. Thus it is a parameter to consider for improving the performance of laser uncaging in future experiments.

Table 3.5 Summary of all UV-laser uncaging trials.

"Duration" stands for the total time of a sample worm being anaesthetised in NaN_3 solution. The animals recovered from anaesthesia are counted as "survivor". "Target" refers to the PCCys-incorporated and UV-uncaged neurons. "Untargeted cells" include neurons (of Strain ZG610) without caspase expression, transgenic cells (of Strain 52.2.8) with no PCCys supplied, and PCCys-incorporated cells unexposed to UV light.

Trials	**Worm no.**	**Survivor no.**	**Efficacy**			**Specificity**		
			Target no.	Ablated cell no.	Success rate	Untargeted cell no.	Unaffected cell no.	Unaffected rate
All	104	60	111	26	23.4%	39	34	87.2%
Duration > 20 min	75	40	82	18	22.0%	29	25	86.2%
Duration ≤ 20 min	29	20	29	8	27.6%	10	9	90.0%

With above figures and tables, I have demonstrated that the UV illumination via microscope-mounted laser is capable of uncaging and thus activating PCCys containing Caspase-3 in an individual neuron. In conclusion, my light-activatable caspase is validated as a promising research tool for precise ablation of a single neuron.

3.2.4.2 Touch receptor neuron

In the previously shown uncaging experiments via whole-worm illumination (Figure 3.17), the majority of PLM cells were completely ablated within 24 hours after illumination. Besides, they are located close to the cuticle, so not much surrounding tissue should deflect or absorb the laser illumination *en route* to them. Therefore, I attempted single-cell uncaging experiments on the pair of PLM neurons, subsequent to the success of laser-activated apoptosis of individual oxygen-sensing neuron.

I first planned a preliminary trial on worms of Strain BZ205-211, in which I did not set “-PCCys -UV”, “-PCCys +UV” or “+PCCys -UV” group. Worms of each strain were grown in the presence of PCCys for 3 days and divided into two groups. One group of worms were transferred to a test plate and illuminated throughout in a UV crosslinker. Worms of the second group were mounted in

slides for laser uncaging in which both PLM neurons were illuminated. Afterwards, all these worms were cultured for another day and then checked under fluorescent microscope.

Representative images exhibited in Figure 3.33 were all obtained from Strain BZ207.1.1 worms. After being uncaged in UV crosslinker (whole-worm mode), PLMs are still present with CFP expressed in their somas but the processes are not visible, which is a common manifestation of the touch receptor neurons of all whole-worm uncaging groups. Meanwhile, laser uncaging only took effect in one worm in which only one PLM, instead of both, had been removed. The remaining PLM appeared in normal shape and showed CFP signal at a similar level to the cyan fluorescence of the ALM and the PVM (of the same animal) that had not been illuminated. Besides, no mCherry signal has been detected in all touch receptor neurons of any imaged worm.

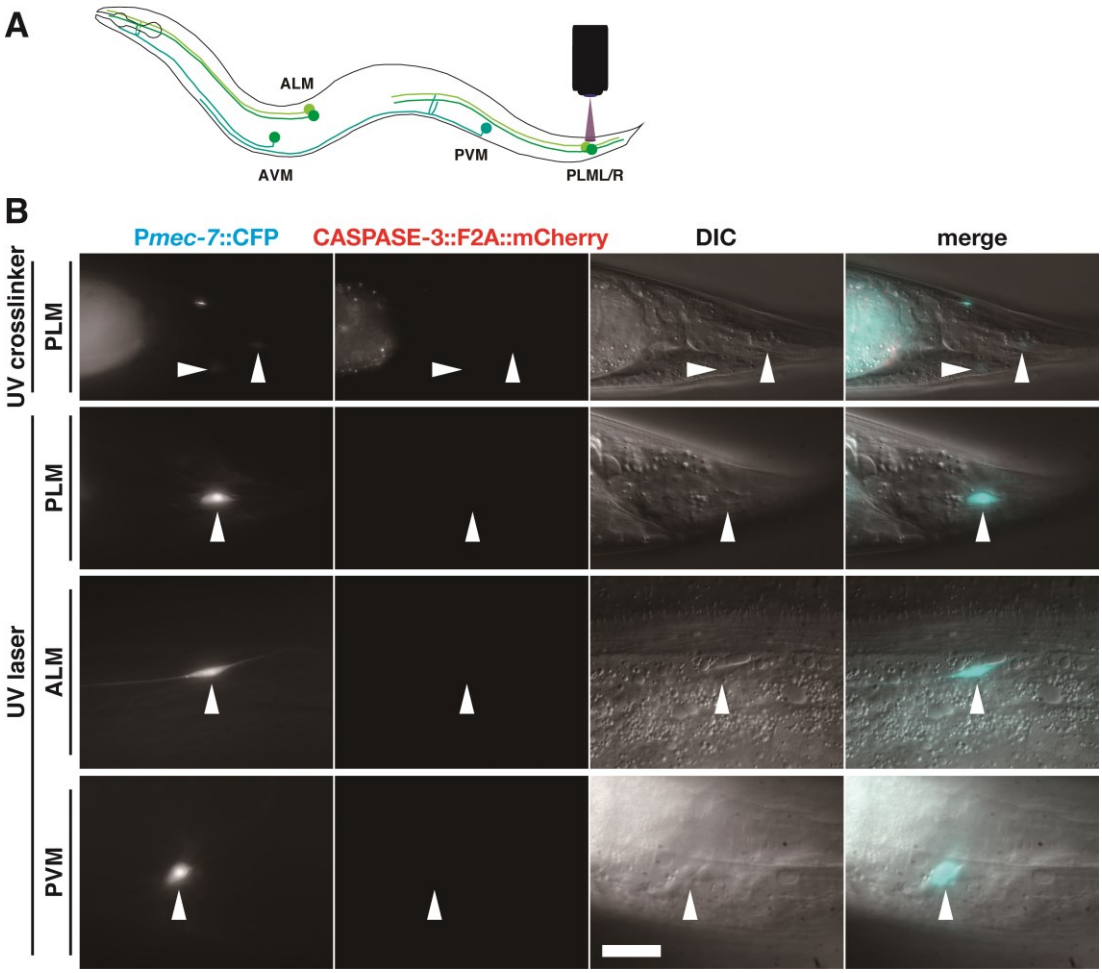

Figure 3.33 Neuron PLM killed by activated PCCys-Caspase upon UV laser illumination.

(A) Schematic of touch receptor neurons in *C. elegans*, with PLML and PLMR illuminated by UV laser. Adapted from Chalfie *et al.*[415]. **(B)** Worms feeding on PCCys-NGM plates with the UV illumination delivered either via UV crosslinker on the entire worm (the first row) or UV laser at the two PLM neurons (the bottom three rows) were imaged. In worms of the crosslinker group, the neurons show CFP fluorescence only in the normal-shaped somas but not processes. In comparison, cell-specific illumination of laser pulses leads to the death of one PLM (the second

row). PCCys-incorporated neurons unexposed to laser uncaging, for instance ALM (the third row) and PVM (the last row), remain intact. Normal touch receptor neurons are marked by solid arrowhead. All images are in the same scale and scale bar = 20 μm.

The results of the whole-worm uncaging (by crosslinker) is inconsistent with my previous observation in BZ210 worms that light-activating the PCCys-caged Caspase-3 ablated the touch receptor cells including PLMs (Figure 3.15 - Figure 3.18). This batch of worms seem to have relatively low expression of the caspase construct as reflected by their barely invisible CFP, which subsequently leads to an inefficient incorporation of PCCys, for the incorporation marker mCherry is unobservable. The removal of only one neuron of the PLM pair after laser uncaging, on the other hand, suggests different efficiencies of PCCys incorporation between the left and right cell of this neuronal pair. More data is required to test the speculation, however, I did not investigate this further with these strains. I rather integrated the array into the genome for BZ211 worms to reduce the expression variability, and proceeded to conduct uncaging experiments as well as behaviour assays on the integrants (Chapter 3.2.2.1.2).

3.2.5 Attempted indirect control of protein via intein activation

By the example of caspase, I have demonstrated that genetic codon expansion of PCCys enables unprecedentedly precise *in-vivo* caging and subsequent reactivation of proteins with a critical cysteine residue. Hundreds of other proteins contain a cysteine of functional importance, with neurobiological instances like various innexins (for assembling gap junctions)[424, 425] and synaptotagmin (for chemical synapses)[426], but I would still like to broaden the applicable range of the genetic encoding of PCCys beyond the above category, to reach more cysteine-containing proteins.

Inteins are proteins that naturally mediate post-translational splicing, in which it self-excises from flanking polypeptides (specially named as N-terminal extein and C-terminal extein) to generate a new polypeptide (i.e. full-length extein) and the excised intein[408] (Figure 3.34A). Among a range of different inteins, the one identified from the *DnaE* gene of cyanobacterium *Nostoc punctiforme* (*Npu*) is the most active and has already been shown to work in worms[408, 409, 427]. Its self-cleavage relies on two cysteine residues: one at its very beginning (C1) and the other is just downstream to its C-terminus (denoted as C+1) and starts the sequence of C-terminal extein[428]. The underlying reaction has several steps as the following[428, 429]. Firstly, C1

participates in an N-to-S (nitrogen-to-sulphur) acyl shift, which generates a thioester intermediate. Then trans-thioesterification happens between C1 and C+1 of the intermediate, forming a succinimide linked N- and C-exteins. Afterwards, S-to-N acyl shift of C+1 transforms the succinimide to the final extein splice[428, 429].

As long as being inserted upstream to a cysteine of the target protein, this intein could excise itself from the nascent polypeptide to possibly restore the structure and function of the host protein[427]. The indispensability of C1 in the self-cleavage activity[430] may potentially render *Npu DnaE* intein a handle for switching on an *in-vivo* protein via a proper external stimulus like light. More specifically, a photo-caged intein can possibly be generated with a photolabile derivative of cysteine incorporated at the first residue (in replacement of C1), and its installation within a cysteine- bearing protein of interest would be very like to allow a light-activated reassemble of the split protein.

This possibility has been validated in mammalian cells by Ren *et al.* with the use of DMNPE-caged cysteine[240]. In the designed construct, the intein gene was put amid the coding region of an mCherry, with the codon of C1 replaced by an amber codon. After the incorporation of DMNPE-Cys via

amber suppression, the bulky DMNPE moiety stops the excision reaction of the intein and only after the removal of DMNPE upon UV illumination, can the intein excise itself and ligate the flanking exteins to produce the anticipated full-length mCherry[240].

So far optical control of intein has been confined in mammalian cells[240], so the first step of my plan was to verify that *Npu DnaE* intein also functions in *C. elegans*. Although the caging group I worked with is different from the caging group DMNPE of the above study[240], I expect that the intein re-activation upon UV uncaging and the subsequent splicing would follow the same course (illustrated in Figure 3.34A). I designed two versions of fluorescent reporter for the wild-type intein and built the plasmids: there is an mCherry::HA::NLS region in which the mCherry sequence is split between the codons of Q142 and C143 by the sequence of *Npu DnaE* intein; the sequence of the transgenic marker GFP is placed either in the upstream (Version 1, Figure 3.34B) or downstream (Version 2, Figure 3.34D) to the mCherry::HA::NLS region. Then I co-bombarded either reporter with the genetic constructs of PCCRS/$tRNA_{CUA}$ to generate transgenic strains.

When imaging the transgenic worms, I observed the co-localisation of green and red fluorescence and all GFP-expressing cells also contain mCherry

signal in their nuclei (Figure 3.34C). The presence of full-length reporters (Version 2) was tested in protein electrophoresis: fusion proteins of approximately 62 kD were detected by both anti-HA and anti-GFP antibodies (Band1 in Figure 3.34E). The results confirmed that *Npu DnaE* intein functions in the cellular context of *C. elegans*.

In the meantime, two extra bands also appeared from anti-HA blotting: Band2 has a size between 38 and 49 kD, while Band3 is between 28 and 38 kD. Zhang *et al.* reported a similar observation from the western blot of an mCherry downstream to an intein[431]. They interpreted the extra band of a size between 14 and 26 kD as a hydrolysed fragment of mCherry from preparing western samples, which seems compatible with my blotting results. Another explanation is that secondary initiation of translation may happen at certain downstream AUG codons of the transcript[432]. A hypothetical reinitiation from Met155 (Met13 of the C-terminal part) of the mCherry of my plasmid will produce a truncate of around 42 kD, matching the size range of Band2. Translation reinitiated from Met231 (Met89 of the C-terminal part) of my mCherry will produce a 34.3 kD protein that matches Band3. A high frequency of the secondary initiation may also explain why some cells only have red fluorescence (Figure 3.34C).

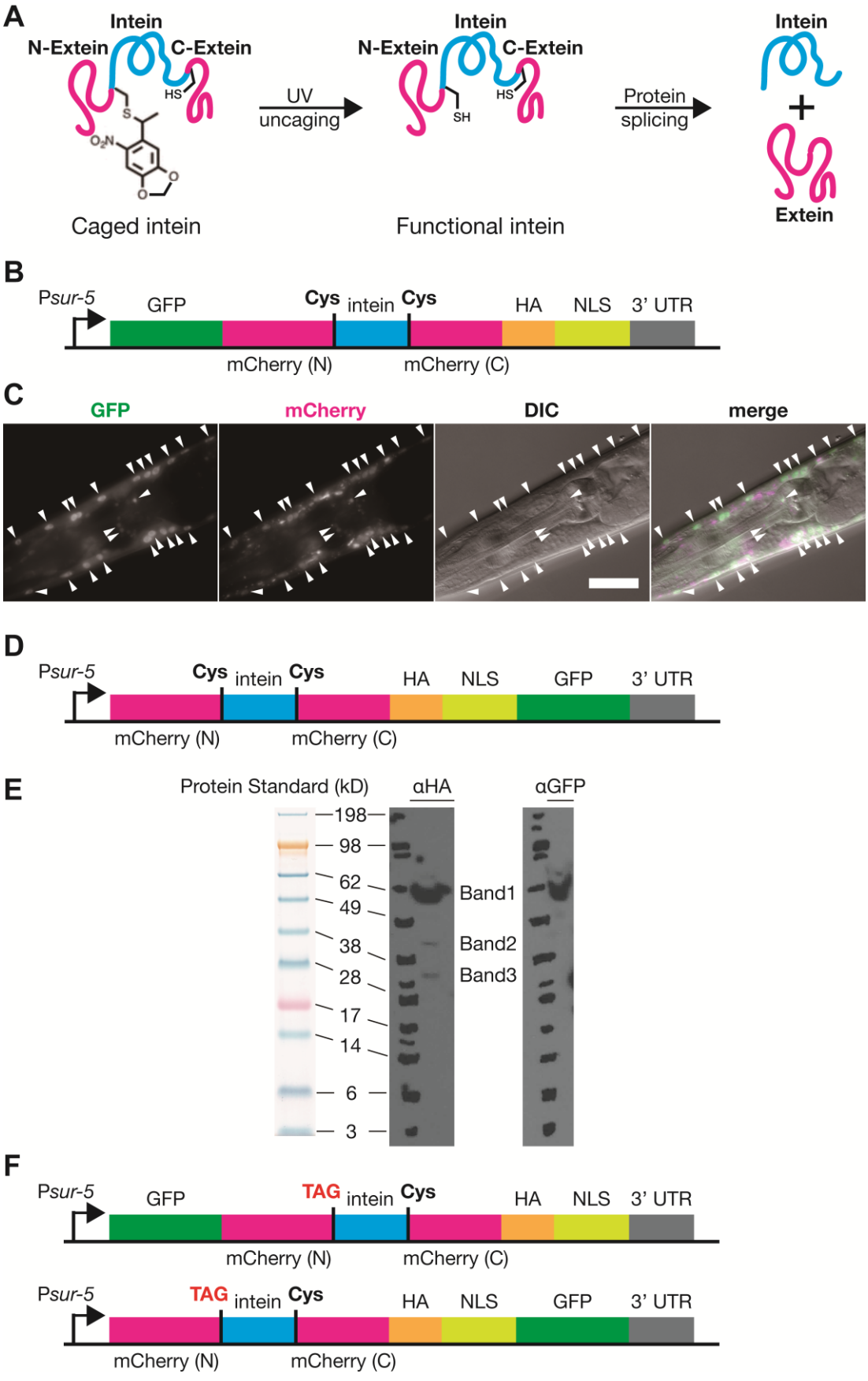
A
Intein
N-Extein
C-Extein
HS
S
O_2N
O
O
UV
uncaging
Intein
N-Extein
C-Extein
HS
SH
Protein
splicing
Intein
+
Extein
Caged intein
Functional intein
B
Psur-5
GFP
Cys
intein
Cys
HA
NLS
3' UTR
mCherry (N)
mCherry (C)
C
GFP
mCherry
DIC
merge
D
Psur-5
Cys
intein
Cys
HA
NLS
GFP
3' UTR
mCherry (N)
mCherry (C)
E
Protein Standard (kD)
αHA
αGFP
198
98
62
49
38
28
17
14
6
3
Band1
Band2
Band3
F
Psur-5
GFP
TAG
intein
Cys
HA
NLS
3' UTR
mCherry (N)
mCherry (C)
Psur-5
TAG
intein
Cys
HA
NLS
GFP
3' UTR
mCherry (N)
mCherry (C)

Figure 3.34 Wild-type *Npu DnaE* intein retains its self-cleavage activity in *C. elegans*.

(A) Diagram of the anticipated light-activated reassembly of a target protein with the pre-installation of PCCys-incorporated intein. The wild-type *Npu DnaE* intein (blue, of "Functional intein") with a Residue 1 cysteine (C1) excises itself from the polypeptide and splices the flanking N- and C-exteins (magenta) to generate a full-length extein[408]. The substitution of C1 with a photo-caged cysteine is anticipated to prevent the intein (blue, of "Caged intein") from self-excision, until the caging group gets removed by UV uncaging. **(B)** Plasmid schematic of the Version 1 fluorescent reporter for the wild-type intein. Ubiquitous P*sur-5* leads expression of the transgenic marker GFP. The downstream mCherry is split between Q142 and C143 by the wild-type intein and the resultant two parts are shown as "mCherry (N)" and "mCherry (C)". Vertical black line represents the cysteines of Residue 1 and Residue 143. HA tag and *egl-13* NLS are attached to the 3' of mCherry (C) region for the biochemical detection and the nuclear localisation of mCherry. **(C)** Images of Strain BZ20.4.1 worms that express the reporter Version 1. Co-localisation of GFP and mCherry is shown in the nucleus of most cells (by arrowhead). The images all follow the same scale bar of 20 μm. **(D)** Plasmid schematic of the Version 2 fluorescent reporter that swaps the GFP region and the mCherry::HA::NLS region of Version 1. **(E)** Western blot image of the reporter Version 2. The reporter-expressing worm (of BZ25.7.1) samples were blotted by either anti-HA ("αHA") or anti-GFP ("αGFP")

antibody. A protein standard was added on the left of both sample lanes and its stains on the membrane (illustrated in the leftmost column) were traced as black bars onto X-ray film. Bands 1-3 were from anti-HA blotting, whereas mere Band 1 was from anti-GFP. **(F)** Plasmid schematics of both versions of the fluorescent reporter for the PCCys-incorporable intein. The intein has its first residue replaced by amber stop codon (TAG, in red). The other genetic elements are the same as those in **B** or **D**, respectively.

As the results confirm a normal self-excision of wild-type intein in *C. elegans*, I then replaced C1 codon of the intein sequence in both versions of the reporter plasmid with an amber stop codon for incorporating PCCys (maps in Figure 3.34F). I transformed constructs of either reporter version with the orthogonal *Mm* PCCRS/$tRNA^{Pyl}_{CUA}$ pair into worms in four bombardment trials (no. 26, 27, 58, 59) and obtained several transgenic strains to test the optical control of intein. As the mCherry region has been split in my design, no red fluorescence would appear before UV uncaging; hence there is no fluorescent indicator of PCCys incorporation. In this case, it is after UV uncaging I began to screen the anticipated transgenic worms whose inteins are light-activatable, by dual fluorescence. To my surprise, not a single cell of any worm express detectable mCherry (no images taken). Considering that I had not used the optimised tRNA synthetase or $tRNA_{CUA}$ for these pilot experiments, inefficient

incorporation of PCCys at the amber site probably leads to this failure. In addition, relatively low temperature (25℃) of worm culture to the optimal 37℃ for *Npu DnaE* intein might reduce the activity of uncaged inteins[430]. In summary, future improvements towards a practical intein tool should include both incorporation boost and intein sequence optimisation.

3.3 Discussion

3.3.1 Amber codon read-through without incorporation

Near the end of Chapter 3.2.2.1.1 I displayed that without the presence of PCCys, the mCherry fluorescence normally indicative of incorporation was still observed in the nuclei of O_2-sensing neurons of some transgenic worms from bombardments no. 194-202. Besides, no mutation has been detected in the neighbouring sequences of the amber codon in the bombarded caspase plasmid (that contains the mCherry) as well as in the transgenes expressed by the worms. Here I will discuss a few possibilities (each in one section) that may contribute to this phenomenon (named as “no-PCCys mCherry” below), and in the end of each section, I will propose corresponding experiments for further investigation in the future.

3.3.1.1 Compromised orthogonality

I noticed that different bombardment groups have different occurrence rate of the no-PCCys mCherry (summarised in Table 3.6). For bombardments using the same tRNA version, be it PylT (no. 194, 197, 200), M15 (no. 195, 198, 201) or C15 (no. 196, 199), bombardments using the Smad4 NES-PCCys tend to generate strains with the false-positive mCherry signal the most easily, followed by the PKIα NES groups. As for the bombardments using the same synthetase, i.e. the original version (no. 194-196), the Smad4 NES version (no. 197-199) or the PKIα NES version (no. 200-201), the M15-using groups have the no-PCCys mCherry by the highest proportion. Unsurprisingly, bombardment no. 201 with the combined use of the Smad4 NES-PCCRS and the M15 has all its transgenic strains expressing the mCherry in the absence of PCCys incorporation. Comparing all the bombardments, I obtain the rough rank (Δ denotes the increment of the occurrence rate of the no-PCCys mCherry): Δ(Smad4 NES, M15) > Δ(PKIα NES, M15) > Δ(PKIα NES, PylT) > Δ(Smad4 NES, PylT) ≈ Δ (M15) > Δ (PKIα NES, C15) > Δ(PylT) > Δ(C15). This seems indicative of the possible synergy or antagonism between orthogonal components that has not been seen in the previous report[290].

Table 3.6 A positive correlation between the uses of NES-PCCRS and $tRNA^{M15/C15}_{CUA}$ with the no-PCCys expression of mCherry in O_2-sensing neurons.

Plasmid details of the bombardment no. 194-202 are in Appendix. Nuclear export signals (NES) from either one of the human genes (PKIα and Smad4)[282, 283, 401] had been attached N-terminally to PCCRS to optimise the synthetase variants. In the column of tRNA, PylT stands for the original $tRNA^{Pyl}_{CUA}$, M15 stands for $tRNA^{M15}_{CUA}$ and C15 is for $tRNA^{C15}_{CUA}$. "PCCys- mCherry+ Strains" column contains the numbers of strains that have red mCherry fluorescence without PCCys supplement. "Fraction of PCCys- mCherry+ Strains" shows the proportion of the PCCys- mCherry+ strains to all transgenic strains.

Bombardment No.	NES of PCCRS	tRNA	Transgenic Strains	PCCys- mCherry+ Strains	Fraction of PCCys- mCherry+ Strains
194	/	PylT	4	1	25%
195	/	M15	6	3	50%
196	/	C15	7	1	14.3%
197	Hu PKIα	PylT	7	4	57.1%
198	Hu PKIα	M15	6	4	66.7%
199	Hu PKIα	C15	7	2	28.6%
200	Hu Smad4	PylT	6	3	50%
201	Hu Smad4	M15	3	3	100%
202	Hu Smad4	C15	5	N/A	N/A

In the results presented by Nguyen *et al.*, mis-incorporation might still happen at a low frequency in the worms expressing the orthogonal PCCRS/$tRNA_{CUA}$, in spite of the extensive laboratory selection of the PCCRS for tRNA acylation specificity[241]. More importantly, my bombardment no. 194 used the same PCCRS/$tRNA_{CUA}$ pair as this study. Compromised orthogonality as I speculated may allow the introduction of amino acids other than photo-caged cysteine at the amber stop site (phrased as amber read-through), which produces a full-length Caspase-3 that mostly remains inactive (except for the specific mis-incorporation of cysteine), and an mCherry::HA::NLS fusion. This deduction agrees with my observation on transgenic worms of bombardment no. 194 (Figure 3.11). And if the assumption holds true, the occurrence of possible mis-incorporations will increase with the use of the NES-PCCRS and/or the modified tRNA variant M15/C15 (as those of bombardment no. 195-201), in corresponding to the incorporation efficiency rise by the increased cytoplasmic concentration of the synthetase by NES attachment and/or tRNA modification, respectively. However, NES or the modified tRNAs *per se* is highly unlikely to impair the orthogonality of tRNA acylation, because of two reasons: 1) N-terminal NES is localised far away from the amino acid binding site of the synthesis, and 2) no NES-related or M15/C15-related read-through has been found by others[166, 169, 290] as well as in my other experiments.

Limited by the amount and the form of the data I collected, I am unable to draw conclusions on the hypothesised mis-incorporation. To further investigate this possibility, I need to generate a test strain whose oxygen-sensing neurons express the caspase construct but not any component of PCCys-encoding aaRS/tRNA pair, and then to check neuronal fluorescence. In the absence of PCCys, if mCherry signal still appears in the CFP-expressing neurons, then the no-PCCys mCherry appearance results from other factors but not an orthogonality loss of any *Mm* PCCRS/tRNA pair that I previously used.

3.3.1.2 Secondary initiation in translation

From the literature I know there is a theoretical possibility that after a normal stop at the amber, protein translation can re-initiated somewhere downstream[432]. As mCherry was observed in the worms, the nucleotide range for the exact re-initiation site is quite limited. The most likely case would be that a ribosome pauses briefly at the amber codon 123-bp upstream of the start codon of mCherry, then starts a secondary initiation at the beginning of mCherry coding region.

For secondary initiation to occur, the 40S subunit of eukaryotic ribosome should remain on the mRNA after it meets the termination codon of the first open reading frame (ORF) and continue to scan for the next AUG codon of this mRNA[432]. When it encounters an AUG, chance exists for a downstream re-initiation with methionyl-tRNA^{Met} ready to bind[433], which requires concentrated initiation factors[434-437]. That correlates with the empirical rule that re-initiation usually occur after a small first ORF (upstream ORF/uORF) under 13 codons, due to gradual dissociation of initiation factors during the elongation phase[432]. And the longest peptide-encoding uORF with secondary initiation that I can find from the literature, is the 93-codon second ORF in the human glucocorticoid receptor 1A transcript[433, 438]. Still, it is far shorter than the 322-codon ORF of my Caspase-3. Therefore, the possibility of secondary initiation resulting in the strong mCherry signal is very small.

3.3.1.3 F2A-mediated titration effect on release factor

2A sequence is used in my caspase construct for multi-cistronic expression [396]. During eukaryotic translation of a single ORF, a 2A sequence facilitates the release of nascent peptide translated from the upstream ORF and allows

the same ribosome to immediately resume translating the downstream sequence. This unusual “intra-ribosomal cleavage” or “StopGo” reaction was discovered in a positive-strand RNA aphthovirus species named foot-and-mouth disease virus (FMDV)[392, 396, 397], so the particular 18-codon mediating sequence is named as F2A. The last three amino acids as well as the first downstream residue (namely Asn16-Pro17-Gly18-Pro19, or NPGP) of F2A are strictly conserved among all the viral 2A sequences that result in the same translational outcome[439]. From a range of experiments, a rough model of the 2A reaction has deduced and is described as below[392, 440].

The translation of a 2A sequence proceeds normally until the incorporation of Gly18 in the ribosomal A site and the translocation of peptidyl(2A)-$tRNA^{Gly}$ to the P site. The working ribosome then unusually stalls for a while, which leaves the codon of Pro19 exposed in the A site and a prolyl-$tRNA^{Pro}$ therefore enters the site. The generation of a peptidyl bond between Gly18 and Pro19 requires a nucleophilic attack by the amino nitrogen of prolyl-$tRNA^{Pro}$ on the carbonyl carbon of peptidyl(2A)-$tRNA^{Gly}$, in which the proline acts as a nucleophile[392, 440]. However, proline has lower nucleophilicity than primary amino acids, as its amine is sterically constrained within a five-member ring[392, 441]. The rigid ring also results in a weaker

interaction of proline with the acceptor site of peptidyl-transferase[441]. In the meantime, the interaction of the nascent chain (especially the conserved DXEXNPGP sequence) with the ribosomal exit tunnel transforms the conformation of the ribosome and brings two consequences[440]: 1) the peptidyl-transferase centre becomes less able to promote the formation of the peptidyl bond between Gly18 and Pro19; 2) ribosomal conformation may become more favorable for the binding of release factors. These changes lead to either one of two possible but incompatible sequences of events as the following[392, 440].

In the first hypothetical situation, peptidyl(2A)-tRNA^{Gly} may be hydrolysed without the involvement of release factors[392]. The water molecule required for the hydrolysis comes from the dipole moment when the ester linkage of peptidyl(2A)-tRNA^{Gly} is positioned beneath the α-helix of the nascent peptide. Subsequently, there no longer exists the conformational effect that expels the prolyl-tRNA^{Pro} and the latter is now able to enter the A site. Then proline becomes the start residue for the new polypeptide upon the translocation of prolyl-tRNA^{Pro} to the ribosomal P site. The hypothesised hydrolysis of peptidyl linkage between the glycine and proline at the C-terminus of 2A should produce the upstream and downstream proteins in a strict 1:1 stoichiometry[392].

In the model opposed to the first one, the failure of the peptide bond formation promotes the cognate prolyl-tRNA^{Pro} to leave from the complex, allowing the binding of release factors and consequent release of the tRNA^{Gly}[440]. Following the termination, the nascent chain gets released from the hydrolysis of its ester linkage with tRNA^{Gly}. Chances exist that a ribosome will not progress along the mRNA after the release of the upstream peptide, producing a higher amount of the upstream protein than the downstream one[440].

To examine the first model, Donnelly *et al.* built an artificial poly-protein construct in which two polypeptides were linked by a F2A sequence[392]. They expressed the construct in an *in-vitro* system of coupled transcription/translation and measured the stoichiometry of the resultant upstream and downstream peptides. The result shows a 2- to 5-fold molar excess of the upstream peptide over the downstream one[392]. Therefore, the most likely mechanism is the second one in which release factors are involved.

The relation between the F2A-mediated ribosomal “StopGo” and release factors has been further investigated in *S. cerevisiae*[440]. The experiments built a reporting system that co-expresses two plasmids. Plasmid1 had a

tandem arrangement of yeast pro-α-factor, F2A and GFP within one ORF. In Plasmid2, a stop codon was placed in-frame between the coding regions of a ß-galactosidase and a luciferase (*lacZ*::*STOP*::*luc*). A negative control of Plasmid2 (here referred as Plasmid2') replaces the stop codon with GCA, a codon of alanine. The activities of ß-galactosidase and luciferase were measured in yeast lysates, and luciferase activity was normalised to ß-galactosidase activity. Then the stop codon read-through was presented as a ratio of the normalised luciferase activity in the Plasmid2-expressing strain to that of the Plasmid2'-expressing strain. It turned out that the frequency of stop codon read-through was increased by the F2A co-expression, which can be further enhanced by over expressing 2A or knocking out a release factor (i.e. in strain [PSI$^+$] which lacks eRF1)[440]. The phenomenon is particularly significant for the amber stop codon followed by a cytidine or guanosine (UAGC or UAGG). As 2A has shown no direct interaction with release factors, the researchers therefore hypothesised a titration model that the *trans*-participation of release factors in the 2A reaction delays its recycling for catalysing translational termination[440].

Notably, Doronina *et al.* reported that of all the translation events of the read-through reporter that bears UAGC (in-frame amber stop codon followed by a cytosine), the sequence also exists in my caspase construct, nearly 30%

read through the amber codon[440]. A noticeable difference in my construct, however, is that the F2A region is also placed in the caspase plasmid and only a few amino acids downstream (i.e. a *cis* arrangement) to the amber codon. Recent work by Cesaratto *et al.* expressed a secretory reporter-stop-F2A construct of stop read-through indication in human cells and analysed the cell supernatant by western blot[442]. However, tno trace of the anticipated read-through protein was observed. Several possibilities may explain the discrepancy: 1) the downstream nucleotide to the amber in the mammalian cell experiment was adenosine, but not cytidine or guanosine that was correlated with high frequency of stop read-through[440]; 2) the assumed titration does not take effect when the amber stop codon is in vicinity of the F2A region; 3) the titration effect is too small to be detected by Western blotting the supernatant; and 4) the mechanism in mammalian cells differs from the one in yeast cells.

Although unsure about the exact reason of such experimental contradiction, I infer that 2A-mediated titration of eRF1 may contribute to the prevalent amber read-through observed in oxygen-sensing neurons of the PCCys- mCherry+ strains across different bombardments. Additionally, my initial screen for strains with high-level expression of CFP (in the caspase plasmid) probably selected worms that over-express F2A as well, which may further enhance

the titration effect on amber read-through. Further investigation is required to test these hypotheses. For instance, I can express in new worm strains a read-through reporting system that consists of a CFP::F2A fusion construct and a GFP::amber::mCherry::HA construct, so that the prospective amber read-through can be observed and measured by fluorescent imaging and western blot. Also, I can bombard the caspase plasmid into worms of an eRF1 RNAi/knockout background, and compare their mCherry expression with the counterpart animals that have a normal level of the release factor.

An important premise for my optical control of caspase is that amber suppression only happens in the presence of PCCys-charged orthogonal $tRNA_{CUA}$. Since false-positive signal of mCherry (i.e., no-PCCys mCherry) may spontaneously emerge and subsequently interfere with its reporting of PCCys incorporation, and I have demonstrated cell ablation by light-activating PCCys-caged Caspase-3, the multi-cistronic expression via F2A of a mCherry reporter concurrently with the engineered caspase becomes no longer necessary. Therefore, to ablate a neuron(s) of interest via this light-activatable caspase in future, I will only have to include the coding region of my caspase variant in a plasmid and co-bombarded it along with the constructs of the orthogonal PCCRS/tRNA pair into worms of a background strain that genomically expresses a fluorescent marker in the neuron(s) to be

targeted. Accordingly, I will only need to screen the candidate worms that have been cultured on PCCys and illuminated in UV crosslinker for the absence of the target neurons.

3.3.2 Comparisons between light-activatable caspase and other ablation tools

Despite the aforementioned aspects that need further improvement, my caspase tool has advantages over other established ablation methods for *C. elegans*. In Chapter 1.2.2, I have proposed five criteria for evaluating cell ablation methods. None of the available tools fulfill all required criteria, which is the main motivation for me to develop a light-activatable caspase. In this section I will compare my caspase tool with the major ablation tools, in respect of the criteria I set earlier.

Nanosecond and femtosecond lasers are two types of laser pulses used for cell ablation and they have different working mechanisms and effects[173]. The shear stress impulse and the ionisation avalanche from nanosecond irradiation usually leads to bursting of plasma membrane and necrosis of the target cell[173, 175, 176, 443]. The high-energy pulse employed by

nanosecond ablation can generate disruptive mechanical effects via shock wave and cavitation on the tissue beyond the energy deposition region[174, 175]. Cell ablation by femtosecond irradiation, however, is a result of chemical bond breakages within biomolecules by low-energy electron resonation, which is more focal and has less collateral damage[173, 174]. In comparison, the optical activation step of my ablation approach does not have these drawbacks. That is because I use low-energy UV laser pulses and short illumination to uncage the caspase.

Besides the above side effects, physical ablation by laser microbeam is of limited applicability from a developmental perspective: practically neuronal ablation has to be performed in L1 worms to eliminate cellular function completely[206]. The nervous system of worms at this stage, however, is not fully developed and not all synaptic connections are formed[116]. Thus, eliminating cells early at the L1 stage may affect the post-embryonic neuronal development of the irradiated worms. During the long interval between the ablation experiment and the behavior assay, maturation and/or plasticity of larval neurons can bring about compensations or adaptations[444, 445], compromise and even mislead the investigation. Furthermore, to ensure the removal of a target cell, in practice it is often necessary to ablate its precursor cell, which usually leads to the loss of additional daughter cells and

consequent side effects. On the contrary, by activating the caged caspase in adult worms, I am able to ablate neurons within a fully matured nervous system.

Genetic ablation works with the ectopic expression of pro-necrotic degenerins, toxins or pro-apoptotic proteins[178-180]. Since this approach relies on promoters to express the exogenous protein in certain cell groups, it has only very limited specificity: the majority of worm cells cannot be individually targeted by a promoter, and using two overlapping promoters only slightly expands the applicable scope to 97 of all 118 neuronal types[110, 116, 180]. Temporal control of the ablation requires the combinatory use of a heat-shock promoter and a tissue-specific promoter, which would further reduce the specificity of this method[180].

Compared with the above methods, optogenetic ablation means have displayed much higher precision in targeting a promoter-specific group of cells at a certain time point. However, the specificity and temporal control, as well as their illumination requirements and the killing effects, seem to be highly influenced by the cellular localisation chosen to express the photo-sensitiser. Table 3.7 compares the experimental schemes of photo-sensitiser variants in respect of localisation. For all these proteins,

mitochondrial expression correlates with the lower light intensity and total energy of the exciting irradiation. This may be because the oxidative respiration within mitochondria is a vital cellular activity that is particularly sensitive to ROS[446]. And the decisive difference of the mitochondrial localisation from the cytosolic presence of ROS generators is that a majority of all the free radicals produced act on the more sensitive mitochondria. One downside of this subcellular concentration, however, is the potential risk of mitotoxicity, which has been observed in worms both acutely (e.g. impaired locomotion[183]) and chronically (reflected by reduced lifespan even for animals kept in the dark[203]). The conundrum is rooted in the mechanism of optogenetic ablation: localising the photo-sensitiser to the most sensitive cellular compartment (i.e., mitochondrion) simultaneously increases the probability of producing oxidative damage in both desired and undesired circumstances, meanwhile to ablate cells with the damaging efficiency reduced means to adopt a harmful length of irradiation (especially with blue light) to worms[201, 202].

The most effective killing has however been achieved by localising ROS generators to plasma membrane[189, 203]. Correspondingly, it is the destruction of the cell membrane and cell integrity that eventually kills the target. The death is non-apoptotic and possibly much more pro-inflammatory

than targeting the ROS generators to mitochondria or to the cytoplasm[189, 203]. In addition, the ROS-based killing of an individual neuron with a targeted laser would require extensive irradiation on the soma and all processes of the target but no other cells, which may go beyond the technical capability of the experimental facility. Hence, ROS generators are generally used only with whole-field instead of cell-specific illumination. Indeed, to my knowledge no published reports have used ROS generators for the targeted ablation of single cells.

In contrast, I have not been troubled by such practical difficulty in my experiments. Illuminating a fraction of cytosolic area can activate a sufficient amount of caspase proteins and then these proteases diffuse throughout the targeted neuron to trigger the “chain reaction” of caspase activation and cellular content degradation that spread over the entire cell, including all branching synapses. As demonstrated in several figures, my method has a spatial resolution high enough to target a single cell for precise ablation.

Prolonged exposure to light of blue (470±25 nm), violet (435±10 nm), or long ultraviolet (350±25 nm, UVA) can induce paralysis and even lethality of *C. elegans*[202]. To activate the caged caspases for inducing apoptosis in my laser-uncaging experiments, I only illuminated the sample for around 10

seconds. In contrast, the durations of illumination required for cell ablation via photo-sensitisers are approximately 25-600 times longer (Table 3.7). Also, the apoptotic cells will get engulfed and degraded by their neighboring cells, which avoids uncontrolled content spillage from dying cells into the surroundings and leaves minimal side effects to the tissue and the animal.

Table 3.7 Comparing parameters of different optogenetic ablation methods and the light-activatable caspase.

Typical parameters for the different ablation methods based on phototoxic proteins are listed. The illumination is continuous unless a pulse frequency is additionally provided. Rows in bold are the case of ablation trials on touch receptor neurons. *Whole worms within the objective field were illuminated and the resultant cell-specific effects were promoter-dependent. **The ablation attempt had no observable effect.

Protein	Localisation	Sensitising light (nm)	Duration (min)	Reference
KillerRed	**Cytosol, neuron**	**530-560**	**30**	[188]
	Mitochondria, muscle	530-610	60	[183]
	Cytomembrane*, neuron	540–580	120	[189]
MiniSOG	Cytosol, epidermis**	455-465	12	[203]
	Mitochondria,	455-495	15-30	[182]

	neuron			
	Cytomembrane*, neuron	455-465	**4, 2Hz**	[203]
MiniSOG (Q103L)	Cytomembrane*, neuron	455-465	0.5, 2Hz	[203]
Light-activatable Caspase-3	**Cytosol, neuron**	**365**	**0.1-0.2, 15 Hz**	This study

In summary, my light-activatable caspase successfully combines the spatial precision of laser targeting, the flexible temporal control from photochemistry, and the efficiency as well as a tightly regulated cell death course by the apoptotic pathway.

Chapter 4: Materials and Methods

4.1 Plasmid construction

Genetic fragments were amplified by PCR from various templates, including the genomic DNA of worms, existent plasmids, and synthetic sequences. DNA fragments (up to 3 fragments) were assembled via overlap extension PCR. The polymerases/kits for DNA amplifying and assembling DNA(s) were Q5 HotStart High-Fidelity 2X MasterMix (New England Biolabs) and KAPA HiFi HotStart ReadyMix PCR Kit (Roche), used along with oligonucleotide primers (Sigma-Aldrich or IDT) listed in Appendix.

Promoters of *gcy-32* and *myo-3*, and a vector carrying no-ATG green fluorescent protein were gifts from the lab of Emanuel Busch (Busch Lab). Promoter of *dat-1* was a gift from the lab of Doitsidou (Doitsidou Lab). Synthetic coding sequences were mainly ordered as gBlocks from IDT. To enhance their expression in *C. elegans*, codons were optimised for via website http://worm-srv3.mpi-cbg.de/codons/ and the following 3 artificial introns were altogether inserted: 5'- gtaag tttaa acata tctat actaa ctaac cctga ttatt taaat tttca g -3' (SynIVS.R), 5'- gtaag tttaa acagt tcggt actaa ctaac catac atatt taaat tttca g -3' (SynIVS.S) and 5'- gtaag tttaa acatg atttt actaa ctaac taatc tgatt taaat tttca g -3' (SynIVS.T)[243, 419, 447, 448].

Plasmid vectors, recombinases and protocols of Gateway Cloning Technology (Life Technologies) were the scaffold on which the plasmid construction of this thesis based. A sequence of interest was inserted into a vector of the 3-fragment platform (either pDONR P4-P1r, pDONR 221, or pDONR P2r-P3) in Gateway recombination mediated by BP clonase II, forming a module for further assembly. Complex genetic components amplified from its original plasmids were constructed altogether into one of the above vectors. The resultant plasmid was further assembled with entry clones of the other 2 types of pDONR vectors and an empty pDEST R4-R3 II vector (backbone) or its derivative in LR recombination via LR clonase II plus, producing the destination plasmid.

Genes of interest were built into the pDEST backbone via DNA restrictive digestion and ligation. For inserting the zipGFP, GC3AI or VC3AI cassette into the vector plasmid IR99, the DNA molecules were each digested by the restriction enzyme pair *Age*I and *Nde*I (New England Biolabs) and then ligated head-to tail by T4 DNA ligase (New England Biolabs). Similar steps were taken iteratively for multiplying the copy number of tRNA expression cassettes in the backbone, with the help of the pre-introduced restriction sites of enzymes *BamH*I, *Bgl*II and *Sal*I (New England Biolabs) in the plasmid backbone.

The products of Gateway reactions, DNA ligation, etc. were mostly transformed into cells of the bacterial strain NEB 5α Competent *E. coli* High Efficiency (New England Biolabs) for plasmid proliferation. The strain NEB stable Competent *E. coli* High Efficiency (New England Biolabs) was used for proliferating multi-copy tRNA plasmids, while OneShot ccdB Survival 2 T1R Competent Cells (ThermoFisher) were for vectors that contain the ccdB gene.

The core regions of plasmids created for the first time were sent out for Sanger sequencing (Source BioScience). The DNA electrophoresis fingerprints obtained subsequent to restrictive digestions were checked for plasmids that were further proliferated.

4.2 *C. elegans* strains and maintenance

N2 (from Bristol, UK) was the wild type strain adopted[103], and the used N2-based genetic variants include: the strain "*smg-2*" carrying the allele *smg-2(e2008)*, strain AX204 with *npr-1*(*ad609*) allele, strain ZG610 carrying a genomically integrated P*gcy-37*::GFP and, strain CZ10175 carrying a genomically integrated P*mec-4*::GFP. AX204, ZG610 were kindly gifted to me from Busch Lab and CZ10175 was a gift from Doitsidou Lab. Most transgenic

strains were generated from biolistic bombardment (in Chapter 4.4), and on some of the transformants, genomic integrations (in Chapter 4.6 and 4.7) were performed.

Unless otherwise stated, worms of all strains were cultured on bacteria-seeded 1x NGM agar[103] at 15-25℃. The bacterial strain for seeding was *E. coli* OP50 of either a liquid culture from clones on LB plates[449], or a suspension of freeze-dried bacteria powder (batch v.2, LabTIE) in sterile water. For mating plates, only 10 µl of overnight OP50 culture was seeded onto the centre of NGM plates. Solution of hygromycin B (Formedium) was added onto plates to a final concentration of 0.3 mg/ml for maintaining the transgenic animals. Gentamicin (Formedium) and nystatin (Merck/Sigma-Aldrich) were also used for inhibiting bacterial and fungal contamination on plates, respectively.

Worm bleaching was performed 1) to eliminate contaminations, and 2) to synchronise a mixed population of worms to the L1 stage. For the first purpose, a 1:1 mixture of 5M NaOH and 5% sodium hypochlorite (Sigma-Aldrich) was freshly made and added onto a NGM plate in a drop, into which gravid adults were picked[450]. The bleaching solution for the latter purpose was an 8:3:2 mixture of H_2O, sodium hypochlorite (5%) and

NaOH (5M). Gravid adults were washed off plates by M9 (supplemented with octylphenol ethoxylate/Triton X-100) and transferred into tubes. Subsequent to a brief wash of M9 and spinning-down, supernatant above the worm pellet was removed with a vacuum pump. The pellet was suspended and shaken in 1.5-2 ml of bleach for around 5 min till most adults broke into two halves. Another three M9 washes were performed and the obtained clean eggs were shaken in M9 for hatching at room temperature over night.

To backcross the lines from genomic integration, integrant worms (Strain1) were crossed with N2 or CZ10175 (both denoted as Strain2) as the following: first, hermaphrodites of Strain1 and males of Strain2 were crossed (P0 generation) on mating plates. In the progeny of P0, the F1 males showing traits of both strains were picked to new mating plates and crossed with hermaphrodites of Strain2. The obtained F2 males of the required traits were again crossed with hermaphrodites of Strain2 and the subsequent hermaphrodite progeny (F3) displaying all the traits needed were singled out and later genotyped. For each crossing, only worms of L4 stage were selected and the ratio of male to hermaphrodite in each crossing was 5:1.

4.3 Worm liquid culture

The protocol is adapted from Praitis *et al.*[450]. When worms grown on NGM plates depleted food and the plates were full of L1 larvae, these worms were washed through an autoclaved Premier polypropylene filter bag (5 micron pore size) with S Medium. The obtained synchronised population of L1 was grown in around 400 ml S Medium supplemented with 400 µl 1000X kanamycin (final concentration 50 µg/ml, Kanamycin-disulfate, Merck/Sigma-Aldrich), 400 µl 1000X gentamicin (final concentration 100 µg/ml, Formedium), 4 ml 100X penicillin/streptomycin/Fungizone (final concentration 10/10/0.4 µg/ml, Lonza Verviers), and 4 ml 100X or 400 µl 1000X nystatin suspension (10,000 or 100,000 units/ml, respectively; final concentration 100 units/ml, Merck/Sigma-Aldrich). The flask of medium was incubated with shaking at 140-180 rpm for dissolving adequate oxygen in the liquid. Worm larva can survive in shaking liquid culture for weeks. 10 ml of concentrated *E. coli* HB101 food was added to initiate worm growth. For biolistic bombardment, the population should be mainly composed of late L4 larvae and young adults.

S Medium was freshly made via mixing the following: S Basal (100mM NaCl, 50mM pH6.0 potassium phosphate buffer and 5 mg/ml cholesterol in ethanol), trace metals solution (1 µM $CuSO_4 \cdot 5H_2O$, 25 µM $FeSO_4 \cdot 7H_2O$, 10 µM

$MnCl_2 \cdot 4H_2O$, 10 µM $ZnSO_4 \cdot 7H_2O$, and 50 µM disodium EDTA), 3mM $CaCl_2$, 3mM $MgSO_4$, and potassium citrate (10 mM, pH6.0).

4.4 Biolistic bombardment

The protocol is adapted from Praitis *et al.*[134, 167, 247, 416]. A Bio-Rad Biolistic PDS-1000/He particle delivery system with Hepta adaptor was used. Accessory parts are 900-1100 psi rupture discs (Bio-Rad), biolistic macrocarriers (Bio-Rad) and stopping screens (Bio-Rad). Empirically, 70,000-200,000 worms are required for a single bombardment. Best transformation yield was obtained from a synchronised population of worms staged between late L4 and young adult grown in liquid culture.

The worms in the liquid-culture flask were pelleted on ice for approximately 40 minutes. The pellet was taken and divided into 15 ml Falcon tubes for washing with M9 until the supernatant was clear. Then 0.5-1.5 ml settled worms were spread onto one unseeded 9-cm NGM plate that had been cooled in fridge and were kept in the fridge for bombarding. A spare plate of worms was also prepared as the negative control of a batch of bombardments.

Approximately 6 mg of 0.3-3 μm gold beads (ChemPur) were coated with 10-15 μg DNA in 50 μl for each bombardment. The DNAs were a combination of freshly linearised plasmids of the orthogonal tRNA synthetase, orthogonal tRNA(s) and gene of interest[451]. Before being coated with the DNA mixture, 60 mg beads per tube were cleaned as the following: first, they were vortexed in 70% ethanol (HPLC-grade, Sigma-Aldrich, the same for the rest) for at least 30 min, then pelleted briefly and washed 3 times with sterile water (beads were vortexed for 1 min, soaked for 1 min, and pelleted briefly for each time, the same for ethanol-wash below) and finally re-suspended in 1 ml 50% glycerol (for dividing and also for at most 2-week storage at 4℃).

After 10-min vortex, 100 μl of gold beads in 50% glycerol (equivalent to 6 mg bead) were transferred to a DNA LoBind tube (Eppendorf) for each bombardment. After a brief spin, the supernatant was discarded. Then the beads were re-suspended in the prepared DNA mixture and the tubes were kept shaking for the remaining steps unless otherwise stated. The tubes were being vortexed while 50 μl ice-cold $CaCl_2$ (2.5M) and 20 μl freshly made spermidine solution (100mM in H_2O, Sigma-Aldrich) were added. The new mixture was incubated on ice for at least 30 min for DNA precipitation and the supernatants of a brief spin-down were discarded. The coated beads were then washed with 300 μl 70% ethanol and 1000 μl 100% ethanol,

re-suspended in 140-170 µl 100% ethanol and kept shaking in a cold block (0℃) until they were used in bombardments. Alternatively, the DNA-coated beads in 100% ethanol can be stored at -20°C for weeks without observable reduction in transformation efficiency.

Macrocarriers and rupture discs were briefly washed in isopropanol and air-dried. Then the macrocarriers were placed into slots of the Hepta adaptor and the DNA-coated beads in ethanol were evenly spread onto each macrocarrier. The partially loaded Hepta adaptor was then placed into the chamber of the Biolistic PDS-1000/He apparatus for evaporating the ethanol by vacuuming. A rupture disc was then placed onto the top of a 7-way gas divider (Figure 4.1B) and the latter was installed onto the ceiling of the chamber. Then a stopping screen (Figure 4.1E) was assembled with the macrocarrier holder (Figure 4.1C), and then both were together placed into the chamber, inverted, so the screen was beneath the upside-down macrocarriers. Up to this step, the Hepta-adapter (Figure 4.1A) has been fully assembled. The worm-spread cold plate (Figure 4.1F) was placed onto the target shelf (Figure 4.1G) below the adapter and the worms were bombarded in the vacuum chamber (of 27.5 inchHg).

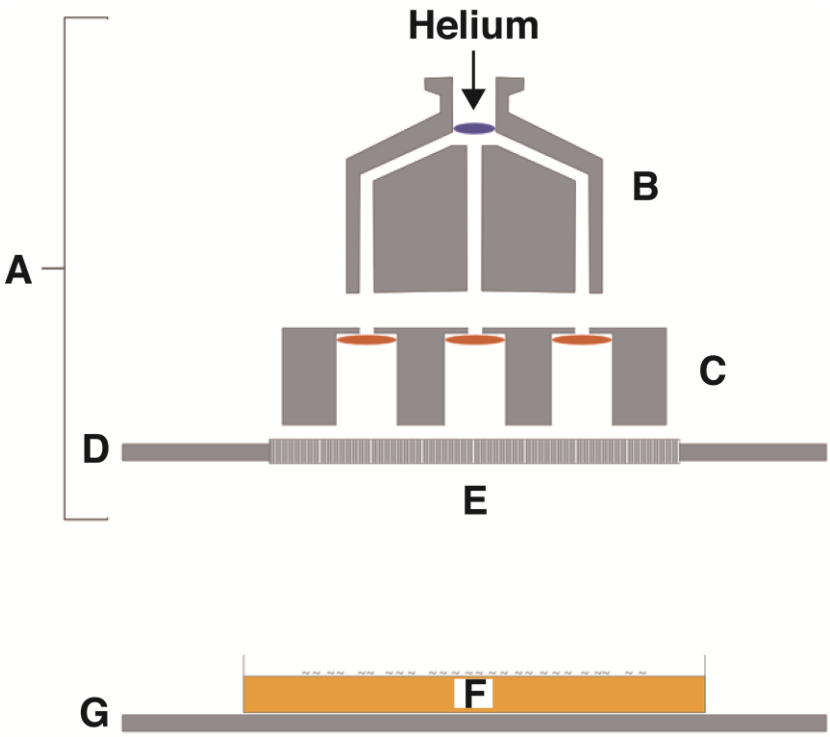

Figure 4.1 Illustration of the vacuum chamber part of the biolistic bombardment system.

Adapted from a figure by Lloyd Davis and Sebastian Greiss (personal communications). Helium gas is delivered via a tube (PEEK Plastic Tubing, 1/16" OD x .010" ID tubing and fittings) into the top of the vacuum chamber. **(A)** The Hepta-adapter consists of four components **B**-**E**. **(B)** The seven-way gas manifold (in cross section) holds a rupture disk (blue oval) and divides the helium stream. **(C)** The holder of macrocarriers (brown oval) that has top openings matching the manifold. **(D)** The holder of a stopping screen (grey gratings, **E**) that allows the pass of DNA-coated microcarriers (i.e., gold beads) while preventing the macrocarriers from hitting the target. **(F)** A plate (with no lid) with worms spread on its agar surface. **(G)** The target shelf sits beneath the Hepta-adapter for holding the plate.

4.5 Post-bombardment selection

The protocol is adapted from Radman *et al.*[159]. Worms on each plate after bombardment were left to recover at room temperature for at least 30 min, before being evenly chunked onto ten 9-cm 3x NGM plates (3 times the peptone added) seeded with *E. coli* HB101 (selection plates, limiting the number of obtained independent transgenic strains at most to ten). The worms were left to lay eggs for 1-2 days so that their progeny (in L1 stage) covered the plates.

Then 100 µl hygromycin B (100 mg/ml in H_2O, Invivogen) was added onto the selection plates to a final concentration of 0.3 mg/ml (hygB powder mass/volume of NGM agar) for selecting transformants. In order to spread hygromycin B over the entire surface, the chemical was diluted with the following antibiotics for a selection solution (the same composition throughout Chapter 4 protocols unless otherwise stated): 30 µl 1000X gentamicin (final concentration 100 µg/ml), 300 µl 100X penicillin/streptomycin/Fungizone (final concentration 10/10/0.4 µg/ml), and 300 µl 100X nystatin (final concentration 100 units/ml). The selection solution was added also to the negative control plate for reference.

To allow the growth of transgenic worms, 50-100 µl bacteria food was added

the next day. Worms were then left to grow for 7-14 days before being screened for fluorescence. Plates with mainly dauer larvae or only very few transgenic worms were re-fed and re-screened.

4.6 UV integration of extrachromosomal arrays

The protocol was adapted from Evans *et al.*[416]. 100 worms of late L4 stage were picked onto 20 unseeded 6-cm NGM plates (5 worms per plate). Then the plates (with the lids off) were irradiated by 254-nm UV with a radiant exposure of 30 mJ/cm^2 in a Stratagene UV Stratalinker 2400 apparatus. After irradiation, 300-400 µl OP50 and hygromycin B with the antibiotic mix of 1000X gentamicin (final concentration 100 µg/ml), 100X penicillin/streptomycin/Fungizone (final concentration 10/10/0.4 µg/ml), and 100X nystatin (final concentration 100 units/ml) was added to each plate. Approximately 1-2 days later, plates were checked for dead eggs as a sign of successful irradiation.

Half of the plates were incubated at 20℃ and the other half at 15℃. The 20℃ half of worms were grown for 2 weeks and the 15℃ half for 3 weeks until worm starvation. A chunk of 1-1.5 cm^2 from each starved plate of the 20℃

half was then transferred onto a fresh plate (master plate) followed by the addition of hygromycin B selection solution. The next day, averagely 10-15 transgenic worms were singled out from each master plate (for a total of 100-150 plates) and transferred to individual fresh plates without hygromycin B selection solution added. After 2 generations, F3 progeny of the single-out animals were scored under microscope for 100% transmission. If no integrant has been found in the initial screen, the experimental steps from plate chunking were taken again for the half of plates incubated at 15℃.

4.7 Gamma integration of extrachromosomal arrays

The protocol was adapted from Evans *et al.*[416]. Synchronised populations of L4/young adult stages were cultured on eight 9-cm NGM plates with food for each strain containing the planned extrachromosomal array. Each of the plates was irradiated by 4000 rads of X-rays (carried out by Michael Fasseas, Invermis) and then chunked to a fresh plate with hygromycin B selection solution added. After 1 day (Day 1), 200 young adults (carrying no egg or only one row of eggs) were singled out to seeded 6-cm plates (P0 plate) and incubated at 25°C. On Day 7, the worms on the plates should be F3 generation and starved. A chunk of approximate 0.5 cm^2 was taken from

each plate and transferred to a fresh seeded plate. Worms showing expected traits were singled out to new 6-cm plates on Day 9 and Day 10, and were scored microscopically for homozygous integration 3 days later.

4.8 Worm lysis

The lysate of sample worms (of BZ194.2.1 and BZ197.6.1) was first made for sequencing the amber-bearing part of Caspase-3 coding region in their extrachromosomal arrays with primers S9 and Z12 (sequence details in Appendix). PCR lysis buffer as a mix of the following reagents was prepared by Feng Xue of Doitsidou Lab and stored at 4°C: (for 100 ml aliquot in sterile water) 5 ml KCl (1M), 250 µl $MgCl_2$ (1M), 10ml Tris (0.1M, pH8.3), 450 µl IGEPAL-CA630 (NP-40 as its used name, Sigma-Aldrich), 450 µl Tween-20 (Sigma-Aldrich), and 500 µl Gelatin (2% in H_2O). A lysis solution was made based on the lysis buffer just prior to the experiment by adding Proteinase K (Life Technologies) to a final concentration of 5 µl/ml.

Individual worms were picked into each PCR tube containing 20-25 µl of the lysis solution. The samples were frozen at -80℃ for 30 min or at -20℃ over night and then incubated at 65℃ for 1-1.5 hr, followed by a final 15-min heat

inactivation at 95℃. After being cooled down to room temperature, the samples were sent out for Sanger sequencing (Source BioScience).

4.9 PCCys-NGM plate preparation

Photo-caged cysteine, PCCys (*S*-[(*R*,*S*)-1-{4',5'-(methylenedioxy)-2'-nitrophenyl}ethyl]-L-cysteine, $C_{12}H_{14}N_2O_6S$) used in this thesis was custom synthesised by ChiroBlock GmbH. The estimated molecular weight of the chemical was 314.31436 g/mol.

The plate for incorporation was the same in composition as an NGM plate except for the dissolved PCCys and the optional addition of Triton X-100 (Merck/Sigma-Aldrich). For a total of 1L NGM agar, 17 g agar, 2.5 g peptone and 3 g NaCl were added into an accordant amount of distilled water and then the mixture was autoclaved. The following ingredients were added into every 25 ml aliquot of the above mixture after it cooled to around 60℃: 250 µl 2.5M- or 625 µl 1M- phosphate buffer (pH6.0), 25 µl $CaCl_2$ (1M), 25 µl $MgSO_4$ (1M), and 25 µl cholesterol (filtered, 5 mg/ml in ethanol). The mixture before the autoclave step was usually cast in big volume, autoclaved and stored at room temperature as a stock. The mixture was microwaved to melt before the

addition of other components.

The initial dissolve of PCCys powder was in an aliquot, which requires an acid solvent. For instance, when making a 25-ml bottle of 5mM PCCys-supplemented NGM agar, 40 mg PCCys powder was weighed in tube by electronic scale. 500 µl sterile water and at least 350 µl HCl (1M) were used to dissolve PCCys. If PCCys powder still precipitates, 50 µl HCl was further added followed by a violent vortex of the PCCys suspension until it became clear, which is usually required for making solutions of PCCys concentration higher than 5mM. A same amount of NaOH as that of HCl was added into the NGM agar for balancing the acid and then the PCCys-HCl solution was added into the NGM agar to finally make PCCys-NGM agar for plate pouring. For the permeabiliser-containing type of PCCys plates, 250 µl of 10% Triton X-100 (for a final concentration of 0.1%) was gently added, avoiding producing bubbles.

For each 3.5-cm plate, 1-1.5 ml of PCCys-NGM agar was used. Extra care should be taken in pipetting, to make sure the agar cover the entire bottom of the plate without any bubble on the surface. To prevent halfway solidification of the agar, the mixture in glassware was kept at around 60℃ in a hybridisation oven till the end. Subsequent to agar solidification, the plates

were placed into opaque boxes for complete darkness and stored at 4℃ for up to 2 months.

The illumination of the lab area is from rare-earth fluorescent lamps (T5 FHO Luxline Plus 24W 840, Sylvania) on the ceiling, and a fraction of the total energy can be emitted as light of wavelengths around 365 nm to which PCCys is reactive (Figure 4.2). Therefore, the handling and storage of PCCys powder, PCCys solution or PCCys-NGM agar (either as liquid or solid) should avoid UV exposure, unless otherwise stated.

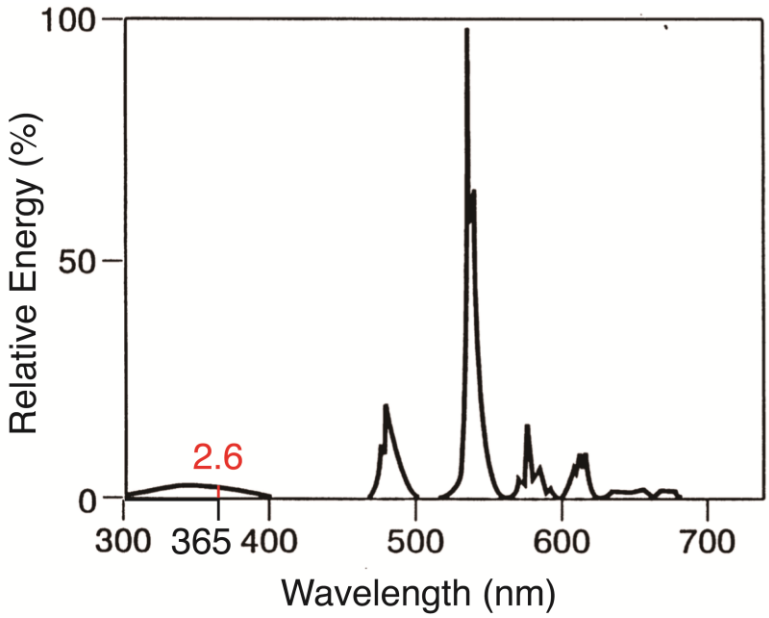

Figure 4.2 Possible emission spectrum of the green light phosphor of lab illumination lamp.

Adapted from Ronda, et al., 1998[452]. (Ce,Tb)$MgAl_{11}O_{19}$ is used for the green light emission in quality fluorescent lamps (CRI > 80) and approximately 2.6% of the full emission energy is at 365 nm.

4.10 Delivery of PCCys to worms

A worm population were grown on seeded NGM plates until the plates were filled with mostly L1 larvae that just starved. Animals of each strain were then either chunked or washed with M9 buffer (50-100 µl, supplemented with 0.001% - 0.01% Triton X-100) onto one PCCys-NGM plate. The same procedure was taken for a negative control group except the use of an unseeded 6-cm NGM plate. A drop of OP50 food was added onto plates after at least 40 min incubation and then worms were grown at 15-25℃ for 24-72 hours prior to subsequent experiments.

4.11 Western blots for reporter protein detection

Synchronised L1 population (around 200-600 individual worms) of target strains were cultured on either PCCys-NGM plates or NGM plates to mostly

young adults. Then the worms were washed off each plate by M9 (supplemented with 0.001% - 0.01% Triton X-100) and collected in a 1.5 ml tube for each group. After worms settled, the supernatant of each tube was removed and the pellet was weighed for estimating its approximate volume (the density of water is used for converting). 30-80 µl 4x Bolt LDS sample buffer (ThermoFisher) and 10-20 µl 10x Bolt Reducing Agent (ThermoFisher) was added to the worm pellet accordingly to reach 1x. The sample tubes were frozen at -20°C over night or longer until the day of protein electrophoresis. Just prior to protein gel running, the samples were boiled at 95°C with violent shaking for 10-15 min. The Bolt LDS sample buffer was also used for sample dilution.

Prepared samples and protein ladder SeeBlue Plus2 Pre-stained Protein Standard (ThermoFisher) were loaded onto Bolt 4-12% Bis-Tris Plus polyacrylamide gel (17-well, ThermoFisher). The electrophoresis ran in freshly diluted Bolt MES SDS running buffer (ThermoFisher) at 200V for 22 minutes. Afterwards, the gel was placed onto the membrane of an iBlot2 Nitrocellulose Transfer Stack (ThermoFisher) installed in the iBlot 2 Gel Transfer Device (ThermoFisher) and the proteins were transferred onto the nitrocellulose membrane by 20V for 7 min.

The blots on the membranes were probed by the following antibodies: Primary ones were rat anti-HA (clone 3F10, Roche) and mouse anti-GFP (clones 7.1 and 13.1, Roche), in a dilution ranging from 1:1,000 to 1:5,000. Respectively, goat anti-rat IgG(H+L)-HRP (Thermo Fisher Scientific) and horse anti-mouse IgG-HRP (Cell Signaling Technology) were used as secondary antibodies, in 1:3,000 to 1:10,000 dilutions. All dilutions were in a blocking buffer (PBS supplemented with 0.1% Tween 20 and 5% milk powder). Blocking was performed at room temperature for 1 hr. Binding of the primary antibody was at 4 °C over night and secondary antibody binding was at room temperature for 1 hr. After the binding of either antibody, the membranes were washed 4 times (10 min per time) in a washing buffer (PBS supplemented with 0.1% Tween 20 and 2.5 % milk powder). Prior to the detection step, extra washes (4X 10 min) in PBS were performed.

To detect proteins with HA tag, the membrane was incubated in 1:1 mixture of Stable Peroxide Solution and Luminol/Enhancer Solution of SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific) for 5 min at room temperature. For GFP detection, incubation of the membrane was in 1:1 mixture of Detection Reagents 1 and 2 of Pierce ECL(enhanced chemiluminescent) Western Blotting Substrate (Thermo Scientific) for 1 min at room temperature. X-ray films (18 x 24 cm, ThermoFisher) exposed to the

chemiluminescence were then developed and fixed in an X-ray developer to show protein bands. Grey values of the protein bands were measured via Fiji/ImageJ (version 2.0.0-rc-68/1.52h, NIH) and the corresponding bar graphs were created using GraphPad Prism version 6.0 for Mac (GraphPad Software, La Jolla California USA, www.graphpad.com).

4.12 Fluorescence microscopy

Zeiss Axio imager.M2 microscope with objective lens of 10x/0.3 NA air immersion, 40x/1.3 NA oil immersion, 63x/1.4 NA oil immersion and 100x/1.4 NA oil immersion was used for imaging worms. The immersion oil with n_D = 1.518 is used accordingly (Immersol, Zeiss). The microscope is also equipped with light sources of a halogen lamp and a mercury vapor short-arc lamp. Differential interference contrast (DIC), GFP, CFP and mCherry images were taken and analysed with the respective use of ZEN software (Zeiss) and Fiji/ImageJ. In the figures presented, images taken through any monochromatic channel or DIC were first given the grey pseudo-colour, and before the merge step were allocated appropriate pseudo-colours (e.g., green for GFP, cyan for CFP, and magenta or red for mCherry).

To prepare samples for imaging, worms were anaesthetised in a 10µl drop of levamisole (5mM in H_2O, Sigma-Aldrich) on an agarose pad (3% in H_2O, ThermoFisher) on glass slide (76 × 26 mm, VWR) and covered by a high-precision glass cover slip (22 × 22 mm, Marienfeld).

4.13 Whole-worm uncaging of PCCys by UV crosslinker

Worms grown on each host plate (either PCCys-NGM or NGM) were washed off by approximately 50-100 µl of M9 supplemented with 0.001% - 0.01% Triton X-100, divided into two groups, one for uncaging and the other unexposed to UV as the negative control, and separately transferred onto two unseeded 6-cm NGM plates.

The plate for uncaging was placed with the lids off into a CL-1000L UV crosslinker (UVP) for 365-nm ultraviolet light illumination (5 mW/cm^2). The range of illumination time used in the thesis was from 6s to 5 min. A drop of OP50 food was added to the plate after illumination. All steps were performed in the dark and the plates were kept in opaque boxes to avoid light exposure.

4.14 Single-cell uncaing of PCCys by microscope-mounted UV laser

The experiments were performed on the Zeiss microscope (Axio imager.M2) with a 63x/1.4 NA oil immersion objective, a UV MicroPoint Galvo module, 365-nm dye resonator cell (MP-27-365-DYE, Andor) and an iQ3 Live Cell Imaging Software (Version 81D6, Andor) pre-installed. The entire apparatus is shown in Figure 4.3.

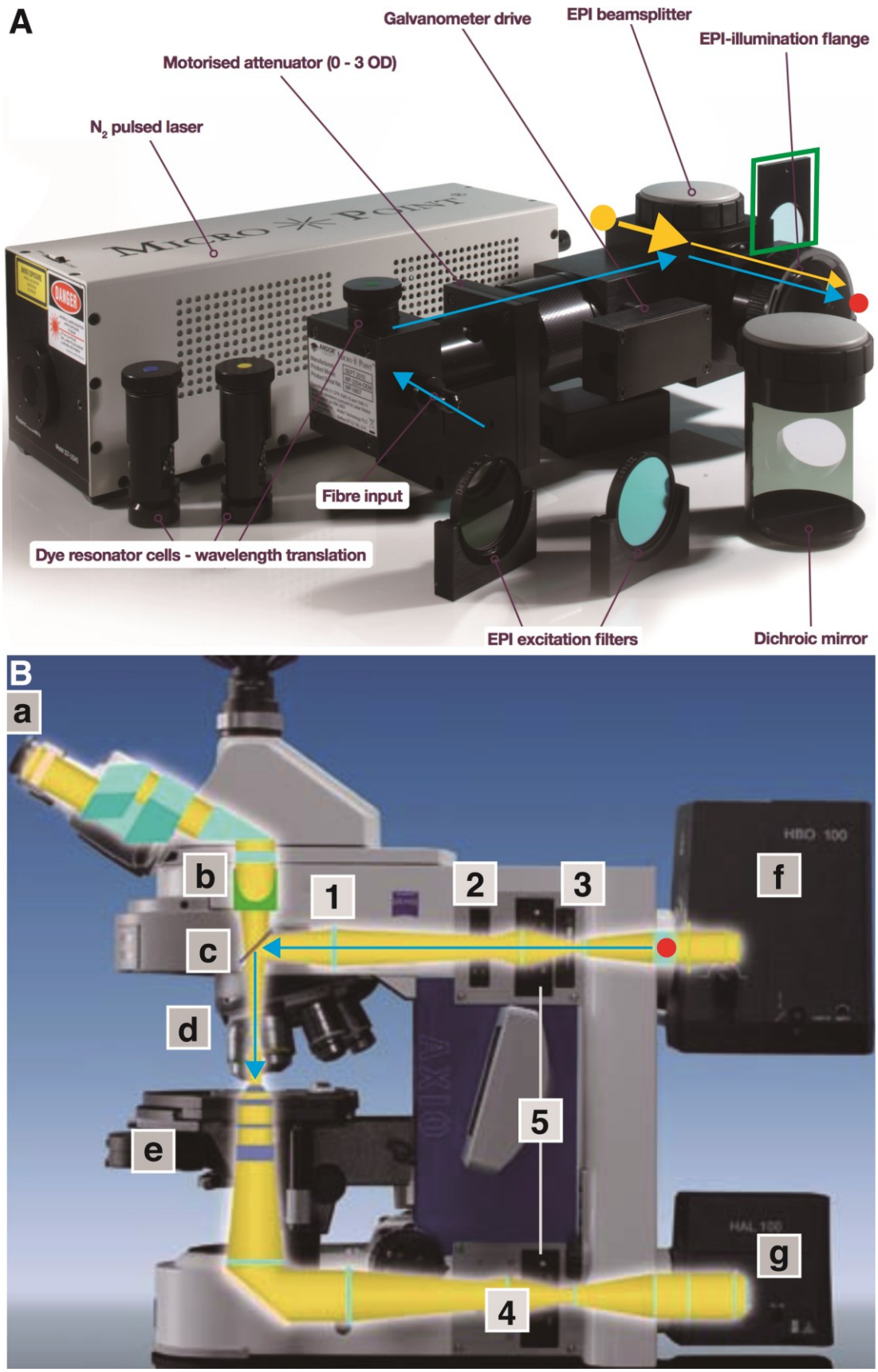
A
Galvanometer drive
EPI beamsplitter
EPI-illumination flange
Motorised attenuator (0 - 3 OD)
N_2 pulsed laser
Fibre input
Dye resonator cells - wavelength translation
EPI excitation filters
Dichroic mirror
B
a
b
c
d
e
f
g
1
2
3
4
5

Figure 4.3 Diagram of the laser beam path of UV-laser uncaging apparatus.

The piece of equipment (MicroPoint) for generation and control of UV laser (**A**) is mounted on an upright fluorescent microscope (**B**). Adapted from an illustration in MicroPoint Specifications sheets that can be downloaded from https://andor.oxinst.com/products/photostimulation/micropoint#. The yellow dot shows the joint of the mercury lamp for MicroPoint and fluorescent excitation. Red dots show the joint of MicroPoint and the main part of Zeiss microscope, which is also the entry of UV light into the microscopic light path. Cyan arrows altogether represent the path of UV beams within the uncaging apparatus, while yellow arrows and shades represent beams from the sources of fluorescent light, transmitted light and reflected light. **(A)** Light beam emitted from N_2 pulsed laser transmits through fibre into a dye resonator cell (MP-27-365-DYE, Andor), where its original wavelength is translated to 365 nm. Then the UV light passes through a motorised attenuator to EPI (Epipolis) beamsplitter with a dichoric mirror, and then the laser joins the attenuated light from the mercury vapour short-arc lamp in HBO 103 W/2 illuminator (**f**). Light from both sources will pass through an EPI excitation filter (GFP exciter filter) marked by green frame and then together enter the microscope. **(B)** Adapted from an illustration of Axio Imager 2 brochure downloaded from https://www.zeiss.com/microscopy/int/products/light-microscopes/axio-imager-2-for-biology.html#downloads. The light from HAL halogen 100 illuminator (**g**) passes

through a transmitted light-field diaphragm (**4**) before being reflected upwards through the condensers (**e**) to the sample on the mechanical stage. Meanwhile, the entry light from MicroPoint sequentially travels through a FL attenuator (or "Slot A", **3**) that has been set to "1" (no attenuation), a reflected light-field diaphragm (**2**) and an accessible interface to the ∞ FL-space (**1**), to reach a reflector (**c**). Then the light is reflected downwards through a 63x objective (**d**) to illuminate the selected area of the sample.

Prior to uncaging, the coordination of MicroPoint aiming centre within the field of view (also displayed as the live image in the iQ3 Live Cell Imaging Software) was firstly calibrated: A reflective slide was shot at for defining the threshold for making a puncture and the x-y coordinates.

Slides of sample worms were prepared in a same way as in Chapter 4.12, except that 2 µl NaN_3 (20mM in M9 buffer, Merck) was used for worm immobilisation. The installation of GFP exciter filter in the beam path allows green fluorescence of the target neurons to indicate their positions for targeting. Before the laser for experimental uncaging was fired, the energy range of laser pulse was calibrated to be just below the intensity that bleaches GFP. In each firing, a central cytosolic area of the target cell was circled and each point within would be shot at in the "fill" mode with the

frequency of 15 Hz, repetitively for 10 times. The firing of an individual cell took approximately 10 seconds. For the negative control/mock-treated groups, the procedure was the same except that no UV energy was delivered.

The uncaged worms were immediately recovered in a drop of M9 buffer (supplemented with 0.001% - 0.01% Triton X-100) on a seeded NGM plate. Anaesthesia of animals of less than 20 minutes ensures acceptable chance of survival.

4.14.1 Calculations of optical parameters

According to the technical manual of the UV laser (MicroPoint Galvo, Andor), the 90-step attenuation of its emission energy follows an exponential curve. I hence have the equation as

$$f(r) = kb^r$$

in which $f(r)$ represents the attenuation coefficient, r represents the attenuation step i.e. the “energy range” changeable in uncaging, b is the exponential base, and the k is a constant.

For $r = 1,\ f(1) = 0.001$, I have $0.001 = kb$ (1)

For $r = 90,\ f(90) = 1$, I have $1 = kb^{90}$ (2)

Combine (1)(2), I have $k = 0.00092532, b = 1.080706753$

Thus I obtain the attenuation formula

$$f(r) = 0.00092532 \times 1.080706753^{r}$$

The energy ranges used in this thesis are 27-29.

For $r = 27$, I have $f(27) = 0.007523233$

For $r = 28$, I have $f(28) = 0.008130408$

For $r = 29$, I have $f(29) = 0.008786587$

Beside, the initial wavelength (440nm) of the laser was tuned to 365nm through the dye cell, which compromises the efficiency by an additional (tuning) coefficient of 0.66. Therefore, the pulse energy and average power of each step/range (r) can be calculated by multiplying the according attenuation coefficient $f(r)$ and tuning coefficient 0.66 with the known full pulse energy (50 μJ) and full average power (750 μW), respectively. The results are used for comparisons in Table 3.4 (Chapter 3.2.4).

4.15 Apoptosis assay

4.15.1 Oxygen-sensing neuron

A synchronised population of newly hatched larva of Strain BZ52.2.8 were transferred onto either NGM plates (negative control) or 5mM PCCys plates on Day 0. Around 40 min after worm transference, 50 µl of OP50 solution was supplied to both groups. Worms were kept at 20℃ for approximately 72 hours and on Day 3 they were picked to two sets of unseeded NGM plates, one for 5-min UV illumination (whole-worm mode by crosslinker) and the other was kept unexposed. Then bacterial food was added to the plates. After approximately 24-32 hours, the worms were scored under fluorescent microscope. Replication experiments were completed for another four times for the same strain. The scoring step of three replicas was performed via Zeiss Axio imager.M2 microscope with an objective lens of 40x/1.3 NA oil immersion on sample worms on glass slides, meanwhile the rest two used a Leica M165FC fluorescent stereomicroscope. The Day-3 phase of the experiments were kindly performed by Ailish Tynan in an effort to reduce subjective bias in scoring. Additionally, independent randomisation was conducted for worm grouping and deciding the scoring order, for minimising possible systemic biases.

The neuron presented with normal morphology (regarding the size and shape of soma and dendrites) were scored “1”, the neuron observed in abnormal shape or of smaller size were scored “0.5”, and neuron disappeared from their original locations were scored “0”. Averages were taken from the two scores of URX neurons of one sample worm. For each neuron type, the scores of all worms (exact worm numbers in Table 4.1) in the group with either absence or presence of PCCys was firstly averaged for each replicate assay. Then the set of five numbers was tested for normality by Shapiro-Wilk test, of which the null-hypothesis is that the population is normally distributed.

As the corresponding p-values to all the sets are well above 0.05 (shown in Table 4.2), the data of each group follows a normal distribution. Therefore, student’s t test with Welch correction was chosen to evaluate the significance of the differences between groups with and without PCCys supply. The p-values of all neuronal types are below 0.05 (Table 4.1), the threshold of statistical significance.

To assess the possible differences in the apoptosis activation by neuronal types, the scores of PQR, AQR and URX in either PCCys-negative or PCCys-positive group were compared by the following procedure. As the data of all groups have normal distributions, so for each PCCys supply

condition (without or with), the homogeneity of variances among the data of different neuronal types were subject to Bartlett test. The resultant two p-values are both below 0.05, so the null hypothesis that the data sets have the same variance was rejected. Bartlett tests were also conducted on the data counterparts from logarithmic transformation by base 2. In this case, the homogeneity of variances was achieved (with p-values higher than 0.05). Consequently, one-way ANOVA (pair-matched, with Geisser-Greenhouse correction) was adopted for each PCCys supply condition. As shown in the table, both p-values are larger than the significance threshold.

Shapiro-Wilk normality tests and Bartlett homogeneity of variances tests were conducted using R language in R Studio (version 0.99.903) and student's t tests with Welch correction and Repeated measures one-way ANOVA tests with Geisser-Greenhouse corrections were performed using GraphPad Prism6.0.

Table 4.1 Sample size of each assay group.

Treatment group	Replicate no.	Sample size
- PCCys	1	19
	2	24
	3	31

	4	30
	5	30
+ PCCys	1	25
	2	35
	3	30
	4	30
	5	30

Table 4.2 P-value results of the statistical tests for oxygen-sensing neurons.

As the calculated p-value in Shapiro-Wilk test is exact for *n = 3* only, the values listed are approximations[399]. The asterisk signs of significance are also used in Figure 3.9 and 3.10.

Major group	Subject for test	Test	P-value	Significance
PQR	-PCCys group	Shapiro-Wilk normality test	0.1474	No
	+PCCys group		0.4113	
AQR	-PCCys group		0.1428	
	+PCCys group		0.3902	
URX	-PCCys group		0.5271	
	+PCCys group		0.31	
PQR	Comparing two groups of the neuronal type	Student's t test	0.0202	*
AQR			< 0.0001	****
URX			0.0006	***
-PCCys	PQR vs. AQR vs. URX	Bartlett test	< 0.0001	****
+PCCys			< 0.0001	****
-PCCys	PQR vs. AQR vs. URX, logarithmicly		0.08922	No

+PCCys	transformed by base 2		0.1905	No
-PCCys	PQR vs. AQR vs. URX, logarithmicly	One-way	0.6523	No
+PCCys	transformed by base 2	ANOVA	0.5619	No

4.15.2 Touch receptor neuron

To begin with (Day 0), synchronised L1 larvae were grown on either 5mM PCCys plates or NGM plates. Around 40 minutes later, 50 µl of OP50 solution was added onto each plate and then the worms were cultured in 25℃ incubator for about 72 hours. On Day 3, worms of each plate were washed off by M9 solution (supplemented with 0.001% Triton), divided into two groups and separately transferred onto two floxuridine (FUDR)-supplied NGM plates (final concentration is 400 µM): one group would be illuminated in UV crosslinker for 5 minutes while the other was kept in the dark. Afterwards, OP50 solution was added and all plates were stored at 20℃ for another 48 hours. On Day 5, neurons of interest in worms were scored under Zeiss Axio imager.M2 microscope with a 40x/1.3 NA oil immersion objective. The agarose pad (3%) on each slide was split into 4 quadrants and onto each quadrant approximately 2.5-µl levamisole (5mM in H_2O) drop was added for immobilising at most 3 worms. Images were taken through GFP and DIC channels of the microscope.

Based on the previously observed apoptotic manifestations of well-incorporating neurons, the morphological changes on the apoptotic course of touch receptor neurons were graded as the following (Figure 4.4): Grade 0 is for a generally invisible cell including a non-structural point; Grade 1 represents a neuron has either no process or severely deformed soma with fragmented processes; Grade 2 denotes a cell of unhealthy soma and fragmented processes; Grade 3 represents the case where either the soma is in slightly abnormal shape or the neuronal processes have nodes/knots, and Grade 4 is for a completely healthy neuron as that of Strain CZ10175.

In the grading phase, randomisations were applied to reduce systematic bias and Ailish Tynan kindly helped giving grades to the neurons through microscope for minimising possible subjective bias.

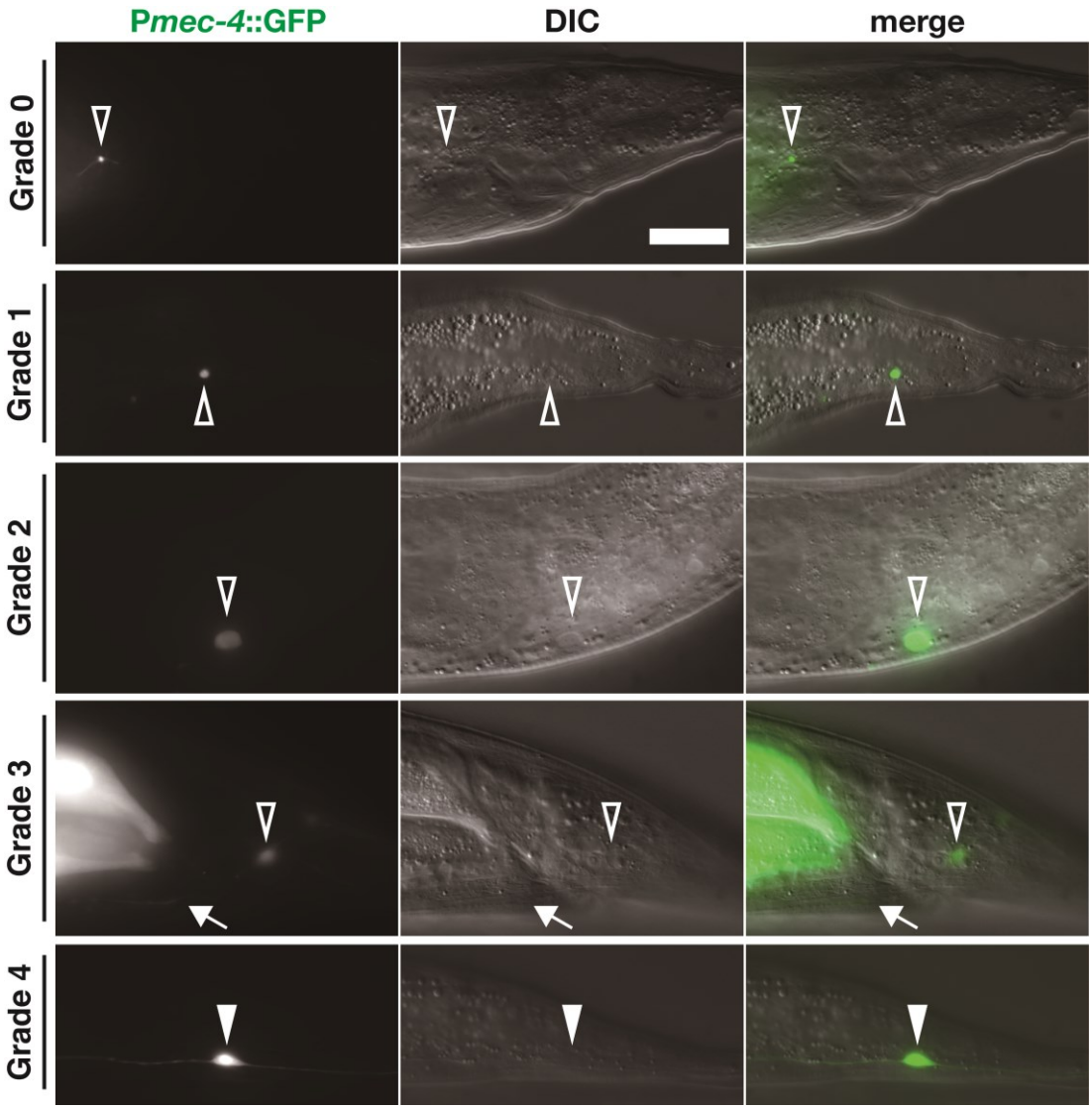

Figure 4.4 Morphological grading of a dying touch receptor neuron via fluorescence microscopy.

Adult worms of BZ211;CZ10175 fed on PCCys-NGM plates were uncaged 2 days before being imaged through the GFP channel of a fluorescent microscope. Posterior touch receptor neurons PLM and PVM are taken as examples for the grading. Compared with the intact PLM of Grade 4 (marked by solid arrowhead), neurons of Grade 3 to Grade 0 (marked by hollow arrowheads) show increasing severity of structural damage. Arrows point to the fragmentation of PLM dendrite. All images are

in the same scale and scale bar = 20 μm. All images display one PLM neuron, except for PVM of Grade 2.

The grading data of each neuronal type in worms without or with PCCys supplement was first subject to Shapiro-Wilk normality test. All resultant p-values are far below 0.05, so the data distributions are non-parametric. All the treatment groups of each neuronal type were statistically compared by Kruskal-Wallis test with post hoc Dunn's multiple comparisons. The accordant p-values are below the significance threshold 0.05 (shown in Table 4.3).

Whether or not the apoptotic manifestations of different neurons differ in each PCCys supply condition (with/without PCCys) was tested through the procedure below. First, the two grades of each neuronal pair (ALM or PLM) in one worm were averaged. The unreplicated data sets of each PCCys supply condition were subject to Friedman rank sum test, the non-parametric counterpart of one-way ANOVA. None of the p-values obtained is smaller than 0.05, so the null hypothesis is accepted: within the same animal, the state of one neuron is related to the state(s) of other neuron(s).

Shapiro-Wilk normality tests were conducted using R language in R Studio

and Kruskal-Wallis tests with post hoc Dunn's multiple comparisons and Friedman rank sum tests were performed using GraphPad Prism 6.0.

Table 4.3 P-value results of the statistical tests for touch receptor neurons.

"NaN" stands for Not a Number. The p-values from Friedman rank sum test are approximate[453]. The asterisk signs of significance are also used in Figure 3.22 and 3.23.

Groups	Sub-groups	Test	P-value	Significance
ALM	Altogether	Kruskal-Wallis test	< 0.0001	****
AVM	Altogether		< 0.0001	
PLM	Altogether		< 0.0001	
PVM	Altogether		< 0.0001	
ALM	-PCCys -UV vs. -PCCys +UV	post hoc Dunn's multiple comparisons	> 0.9999	No
	-PCCys -UV vs. +PCCys +UV		< 0.0001	****
	-PCCys +UV vs. +PCCys +UV		< 0.0001	****
AVM	-PCCys -UV vs. -PCCys +UV		> 0.9999	No
	-PCCys -UV vs. +PCCys +UV		< 0.0001	****
	-PCCys +UV vs. +PCCys +UV		< 0.0001	****
PLM	-PCCys -UV vs. -PCCys +UV		0.6065	No
	-PCCys -UV vs. +PCCys +UV		< 0.0001	****
	-PCCys +UV vs. +PCCys +UV		0.0003	***
PVM	-PCCys -UV vs. -PCCys +UV		> 0.9999	No
	-PCCys -UV vs. +PCCys +UV		< 0.0001	****
	-PCCys +UV vs. +PCCys +UV		< 0.0001	****
-PCCys -UV	ALM vs. AVM vs. PLM vs. PVM	Friedman rank sum	NaN	No

-PCCys +UV		test	0.1834	No
+PCCys +UV			0.5934	No

4.16 Gentle touch assay

10 BZ211;CZ10175 worms were randomly picked from the Day-5 counterpart treatment groups prepared in the serial experiment described in Chapter 4.15.2 for gentle touch assay, along with a control group of 10 young adults of Strain CZ10175. All groups were acclimatised on fresh unseeded NGM plates for at least 15 minutes.

The tool for generating gentle touch was made from an eyebrow hair and a capillary glass tube, by sticking the hair shaft parallel to the glass tube. The mechanosensory force generated from the bend of the hair was calibrated by analytical balance with the calculation:

$$Force\ (\mu N) = Weight\ (mg)\ \times\ g\ (9.8\ m \cdot s^{-2})$$

In order not to stimulate PVD neuron that is responsible for harsh touch, the force delivered should be maintained between 10 μN (the threshold of gentle touch response) and 100 μN [423].

To avoid stimulating other mechanosensory neurons, the anterior stroke point was set just posterior the terminal pharynx bulb and the posterior point was just anterior to the anus of *C. elegans*. In each stroke, the hair shaft should sweep perpendicularly to the body axis, across the entire side of the worm cuticle. A normal reaction to the head touch is reverse, while being touched at the tail should stimulate the animal either to accelerate during forward movement or to switch from going backward to forward.

Each worm was touched 10 times in a row, with the stroke point alternating between head and tail to minimise habituation. The positive responses to the 5 head or tail touches were counted and recorded as the scores. Before switching to assay a new worm, the eyebrow hair of the tool was dipped into 70% ethanol for sterilisation and air dried.

All eight sets of scores of four treatment groups (i.e., the head scores and tail scores of “CZ10175”, “-PCCys -UV”, “-PCCys +UV” and “+PCCys +UV”) were tested for normality by Shapiro-Wilk test. As only the p-value of “+PCCys +UV” tail scores is above 0.05 (namely, 0.3254), Kruskal-Wallis test, the non-parametric counterpart of two-way ANOVA, was selected to compare all four data sets of either head touch or tail touch. As shown in Table 4.4, the p-value of head touch is above the significance threshold, while that of tail

touch is smaller than 0.05. Therefore, multiple Dunn's post hoc tests were performed for comparing the mean of each set with the mean of every other set.

To investigate the effect of stroke position in all these conditions, the data was subject to Wilcoxon matched-pairs signed rank test. Except that the group of the PCCys-incorporated and UV-illuminated worms reported significance, all other p-values are above 0.05 (Table 4.4).

Table 4.4 P-value results of multiple statistical tests for gentle touch assays.

All the p-values of Dunn's tests are multiplicity-adjusted for multiple comparisons. The asterisk signs of significance are also used in Figure 3.28 and 3.29.

Groups	Sub-groups	Test	P-value	Significance
Head touch	Altogether	Kruskal-Wallis test	0.4177	No
Tail touch	Altogether		0.0038	**
Tail touch	CZ10175 vs. -PCCys -UV	Post hoc Dunn's multiple comparisons	> 0.9999	No
	CZ10175 vs. -PCCys +UV		> 0.9999	No
	CZ10175 vs. +PCCys +UV		0.0836	No
	-PCCys -UV vs. -PCCys +UV		0.8063	No
	-PCCys -UV vs. +PCCys +UV		0.0021	**
	-PCCys +UV vs. +PCCys +UV		0.2224	No
CZ10175	Head touch vs. tail touch	Wilcoxon matched-pairs	0.2500	No

-PCCys -UV		signed rank test	0.500	No
-PCCys +UV			0.1250	No
+PCCys +UV			0.0078	**

Chapter 5: Future directions

From the basis of the work displayed in the above four chapters, there are several directions of further investigation.

As has been discussed in Chapter 3.3.2, ablating a defined neuron within the mature nervous system of an adult worm is not attainable via previous ablation tools. I demonstrated in Chapter 3 that oxygen-sensing neurons can be individually ablated upon optical uncaging of the PCCys incorporated into the active site of Caspase-3 that has been expressed in the cells. With this method, it becomes possible to selectively kill individual neurons in a neuronal circuit of interest to dissect various functions of different cells in behavioural assays. For example, single cell ablation will help distinguish between subtly different roles of the left cell and the right cell of a bilaterally symmetrical neuronal pair that share the same genetic profile, therefore facilitating investigation into how functional laterality observed in worm behaviour emerges from certain symmetrical circuits of neurons[454, 455].

Although my thesis only demonstrates successful light-activation of engineered Caspase-3, it may very likely be possible to endow CED-3 with photo-controllability via applying the same engineering scheme used for Caspase-3. In addition to apoptosis mediation, the *C. elegans* caspase

CED-3 was shown involved in the early stage of healing of neural process that were severed by laser axotomy: CED-3 inactivation delays the initiation of axonal repair and also slows the regeneration process[456]. It is however still unclear how activity of CED-3, the core executor of the *C. elegans* apoptotic pathway, is regulated so that it promotes axon reconstruction but not apoptosis of the severed neuron. The question is difficult to address with the traditional mutant-rescue approach, as expressing the *ced-3* transgene specifically in the target neuron of *ced-3* deficient worms usually results to the death of the cell[456]. It may be possible to tune caspase activity in a target neuron with the introduction of light-activatable CED-3, via two layers of control: 1) modulating the concentration of PCCys delivered in worm culture to regulate the amount of activatable caspase proteins; 2) adjusting the area of UV illumination to confine the activity of caspases to certain cellular compartments like a process or its distal severed segment.

Similarly, low-level caspase activity has been discovered in a range of developmental events of mammalian cells[457, 458]. In mouse olfactory sensory neurons for example, loss of function of their endogenous Caspase-9 impairs the growth of axons and synapses in late embryonic stage[458]. It is still unknown how Caspase-9 activation is regulated in the pathways of neuronal development and apoptosis with contrary cellular outcomes[458].

With the level of caspase activity tunable via PCCys incorporation and uncaging, it may be technically easier to build corresponding neuronal models in either cell culture or organismic brain slices, which may facilitate deeper investigation into this question.

Several pro-inflammatory caspases of mammalian immune cells (mainly dendritic cells, monocytes and macrophages) have been demonstrated to mediate protein cleavage in pyroptosis, a type of programmed cell death that is initiated by contact with pathogens or their products[459]. For instance, human Caspase-4, Caspase-5 and mouse Caspase-11 are found able to bind lipopolysaccharide (LPS) upon infections with gram-negative bacteria to mediate inflammatory programmed death (or pyroptosis) of the infected cell[460]. Spatiotemporally controlled activation of corresponding caspase proteins as demonstrated by Caspase-3 may assist future research into the role of non-apoptotic cell death in innate immune mechanisms and exploration of potential targets for immune therapy.

Besides caspases, protein control via site-specific installation of photo-caged cysteine has application in a potentially broad range of target proteins. Critical cysteine residues have been identified in various classes of proteins, such as proteins forming gap junctions (innexins in invertebrates, connexins in

vertebrates), ubiquitin ligases, deubiquitinases and inteins. For example, the two gap junction hemichannels formed by innexins or connexins (innexons/connexons) on adjacent cytoplasmic membranes are locked together by conserved and critical cysteine residues to assemble into a electrical synapse[461-463]. *C. elegans* has 25 innexin genes in total and each class of neuronal cell contains several different types of innexin[259, 464]. Moreover, almost all worm neurons are interconnected by the approximately 6000 electrical synapses made from various combinations of different innexins[117, 464]. Among them are heterotypic gap junctions whose two hemichannels are different from each other and heteromeric types of which at least one hemichannel contains various innexin subunits[464]. Therefore, it is extremely difficult to decipher the innexin composition and dissect the function of an individual electrical synapse within the *C. elegans* connectome, with current unspecific chemical inhibitors and gene knockdown/knockout methods. However, through PCCys incorporation in the target innexin type and then optical uncaging of specific neuronal process, it may be possible to inhibit and rescue individual gap junctions for subsequent electrophysiological or behavioural recordings.

Ubiquitination is an important type of post-translational protein modification. According to the number of ubiquitin attached to its substrate, ubiquitination

can be classified into three types, monoubiquitination, multi-monoubiquitination and polyubiquitination: monoubiquitination and multi-monoubiquitination regulate signal transduction of various cell activities while polyubiquitination promotes proteolysis of proteins, and all are vital to proteostasis of a cell[465-469]. The ubiquitin-proteasome system mediates one of the two major pathways of protein degradation. In this system, ubiquitin ligases tag undesired proteins with a polyubiquitin chain, so that the proteasome can recognise these substrates and degrade them into short peptides by proteolysis[469]. Each ubiquitin addition round of protein polyubiquitination consists of three sequential steps and each step is mediated by a different ubiquitin ligase: ubiquitin-activating enzyme (E1) activates a glycine at the C-terminus of ubiquitin and this activated glycine then forms a thioester bond with a catalytic cysteine of E1, generating a E1-ubiquitin thioester; ubiquitin-conjugating enzyme (E2) catalyses a transthiolation by which the activated ubiquitin in the thioester intermediate gets transferred to a conserved cysteine residue in the E2 active site; ubiquitin-protein enzyme (E3) transfers the ubiquitin to a protein substrate or the ubiquitin chain on a target protein already recruited[467, 469]. A critical cysteine is indispensable to the respective catalytic function of E1, E2 and the HECT-domain class of E3, which provides a site for photo-caging and its subsequent uncaging for spatiotemporal controllability. In addition,

deubiquitinases reverse protein ubiquitination by removing ubiquitin(s) from tagged substrates (deubiquitination)[470]. The catalytic activity of the cysteine protease class of deubiquitinases also relies on a key cysteine[471], allowing it to be optically regulated via site-specific incorporation of PCCys.

C. elegans has one ubiquitin-activating ligase UBA-1, 22 ubiquitin-conjugating enzymes, multiple ubiquitin-protein enzymes in four different classes and 45 enzymes for deubiquitination[469, 472, 473]. Particularly, mutating UBA-1 impairs ubiquitin-proteasome-mediated protein degradation and different E2s each activate a special set of E3s[469, 473]. By installing PCCys into the active site of these cysteine-containing enzymes, it should become possible to *in vivo* regulate protein ubiquitination and deubiquitination in specific cells of an animal. Furthermore, both polyubiquitination and deubiquitination are highly conserved through evolution[469], so the knowledge gained from the envisaged PCCys application in these worm enzymes is likely to facilitate research into the enzyme homologues in other species.

There are still many proteins whose cysteine residues are neither functionally characterised nor evidently essential. With the facility of current artificial intelligence-based models for protein structure prediction like

AlphaFold[474], it is now possible to *in silico* preview the potential effects of photo-caging a cysteine residue in proteins of interest thus to potentially endow a wider range of target proteins with photo-controllability, via PCCys incorporation and uncaging in experimental practice. In addition, as introduced in Chapter 3.2.5, light-activatable inteins can be used to reversibly inactivate proteins of interest in mammalian cells[240]. Despite of my unsuccessful attempts of intein engineering, with optimisation means applied to the genetic code expansion system (i.e., NES usage and tRNA hybrids[169, 281, 290]), the self-cleavable peptide remains a promising candidate for PCCys incorporation in various cysteine-bearing targets.

Potential applications of PCCys incorporation for controlling various proteins in future *C. elegans* research require quantifiable and stable expression of relevant transgenes across generations of worms. Although it is convenient to adopt biolistic bombardment for transgenesis, unstable inheritance of the transgenes in extrachromosomal arrays can result in variable expression and therefore mosaic transgenic worms[133, 157]. My attempt of array integration via UV or gamma irradiation, however, is labour-intensive and inefficient and lacks control of the copy number of transgenes eventually integrated. Meanwhile as demonstrated in Chapter 3, optimisations of the orthogonal PCCRS/tRNA system enhances the efficiency of PCCys incorporation.

Therefore, single-copy integration of the encoding components may still suffice PCCys incorporation while eliminating copy-number variability of the transgenes. The techniques include MosSCI that integrates transgenes at a specific Mos1 transposon site[136] and miniMos with random genomic insertion sites[475]. Genetic editing tool CRISPR/Cas9 can be used for efficiently integrating small components (of 1-2 kb)[476-478].

In addition to the above tools, recombinase-mediated cassette exchange (RMCE) methods have also been developed for stable transgenesis[479-481]. One such example combines two different recombinase systems, Flp/FRT and Cre/loxP[481]. First, Flp recombinase integrates into the landing site (with Mos1 and one loxP site pre-introduced) the sequence of interest and a self-excising cassette (SEC) that contains a hygR for antibiotic selection, heat-shock promoted Cre recombinase and one of loxP sties[476]. Then Cre recombinase excises the sequence between the two loxP sites so that only one loxP site, two FRT sites and sequence of interest remain in worm genome. The approach shows high rate of integrant generation in installing a single copy of bipartite reporter system of several different types into *C. elegans* chromosome[481], including the *E. coli*-originated tetR/tetO system[482], the yeast-originated Gal4/UAS system[483-485], LexA/lexO system[486], and the QF/QUAS system that can be made ligand-inducible[487-489]. As the

resultant strain of each system expresses a much higher level of GFP than an integrated simple promoter-GFP construct, the RMCE-bipartite reporter system combination is promising to realise stably high expression with clear transgenic architecture. This is particularly useful for building an application platform of light-activatable caspase and other prospective target proteins. A transgenic strain for this purpose could for example contain UAS-driven orthogonal PCCRS together with a P*rpr-1*-driven tRNA for PCCys incorporation, and the engineered caspase (or other target proteins), so tissue-specific expression of this system could be easily achieved by crossing the strain with another integrant strain bearing GAL-4 with a promoter active in the target tissue.

Another major source of variability of transgene expression in worms is germline silencing, which is not prevented by genomic integration[490]. Contrary to their persistent expression in somatic tissues[491], transgenes as simple arrays only express for the first few generations in germline cells and are then silenced by dsRNA-mediated interference[158, 492].

To inhibit germline silencing, a range of transformation methods have been tried, including genomic DNA co-injection[158], biolistic bombardment[134], single-copy transgenesis [136, 475] and CRISPR/Cas9 insertion[140].

However, none of these is sufficient to avoid silencing. Besides, sequence adaptation means, such as codon optimisation[243], piRNA removal[140, 493, 494], the removal or restrictive digest of the vector backbone[451] and intron addition[390], only promote expression to a limited extent.

Recently, periodic A_n/T_n clusters (PATCs) in *C. elegans* genome[495] were extensively used in designed introns of transgenes for single-copy integration, which substantially prevents germline silencing in the generated transgenic worms[496-500]. The enhancement of germline expression via PATC-rich introns in transgenes is also compatible with intergenic operon and 2A sequence[397, 490, 501].

My previous plasmid design included artificial introns in transgene coding regions[390] and further optimisation should consider the prevalent embedment of various tRNA sequences in endogenous introns of worm protein-coding genes[260] and PATC-related antagonising effects to germline silencing shown as above. Therefore, future designs of the transgene-expressing plasmids should have the synthetic introns inserted with the orthogonal tRNA sequences and then optimised to be PATC-rich via an online builder[490], thereby establishing a reliable platform of PCCys

incorporation that is promising to benefit various biological studies based on *C. elegans*.

Appendix

Table A1. List of strains used in this thesis

Below are the strains obtained from other labs (as Chapter 4.2 stated): AX204, *npr-1*(*ad609*) X; CZ10175, *zdIs5 [Pmec-4::GFP + lin-15(+)]*; *smg-2*, *smg-2(e2008)* allele; and ZG610, *iaIs25 [Pgcy-37::GFP + unc-119(+)]*.

For the example numbering “a.b.c” of a transgenic strain generated in this study: “a” is the number of biolistic bombardment experiment (details in Table A2); “b” is the number of independent selection plate from one bombardment, meaning an independent strain; “c” is for distinguishing different single-out worms from the same selection plate.

Strain no. (BZ)	**Description**
2.4.2	*Ex [Pwars-1::MmPCCRS + Psur-5::hygR; Prpr-1::MmPylT; Prps-0::GFP(amber)::mCherry::HA::egl-13 NLS]*
20.4.1	*Ex [Psur-5::MmPCCRS + Pwars-1::hygR; Prpr-1::MmPylT; Psur-5::GFP::mCherry(N)::Intein(wt)::mCherry(C):NLS]*
25.7.1	*Ex [Psur-5::MmPCCRS + Pwars-1::hygR; Prpr-1::MmPylT; Psur-5::mCherry(N)::Intein(wt)::mCherry(C):NLS::GFP]*
52.2.8	*Ex [Pgcy-32::MmPCCRS + Pwars-1::hygR; Prpr-1::MmPylT; Pgcy-32::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
84.1.2	*Ex [Pgcy-32::MmPCCRS + Psur-5::hygR; Pgcy-32::Prpr-1::MmPylT; Pgcy-32::rev-ced-3::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
91.4.1	*Ex [Pgcy-32::MmPCCRS + Prps-0::hygR; Pgcy-32::Prpr-1::MmPylT;*

	Pgcy-32::GFP(amber)::mCherry::HA::egl-13 NLS]
92.3.1	*Ex [Pdat-1::MmPCCRS + Prps-0::hygR; Pdat-1::Prpr-1::MmPylT + NeoR; Pdat-1::GFP(amber)::mCherry::HA::egl-13 NLS]*
132.1.1	*Ex [Pncs-1::MmPCCRS + Prps-0::hygR; Pncs-1::Prpr-1::MmPylT; Pncs-1::GFP(amber)::mCherry::HA::egl-13 NLS]*
138.1.1	*Ex [Pgcy-32::MmPCCRS + Prps-0::hygR::SL2::VC3AI; Pgcy-32::Prpr-1::MmPylT; Pgcy-32::rev-ced-3::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
194.2.1	*Ex [Pgcy-32::MmPCCRS + Prps-0::hygR; Prpr-1::MmPylT; Pgcy-32::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
197.6.1	*Ex [Pgcy-32::Hu PKIa NES::PCCRS + Prps-0::hygR; Prpr-1::MmPylT; Pgcy-32::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
207.1.1	*Ex [Pmec-7::Hu PKIa NES::PCCRS; Prpr-1::tRNA(M15); Pmec-7::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
210.1.1	*Ex [Pmec-7::Hu Smad4 NES::PCCRS; Prpr-1::tRNA(M15); Pmec-7::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
210.5.2	
211.2.1	*Ex [Pmec-7::Hu Smad-4 NES::PCCRS + Prps-0::hygR; Prpr-1::tRNA(C15); Pmec-7::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
211.2.1.7.1	*Is [Pmec-7::Hu Smad-4 NES::PCCRS + Prps-0::hygR; Prpr-1::tRNA(C15); Pmec-7::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
212.5.1	*Ex [Pmyo-3::Hu PKIa NES::PCCRS + Prps-0::hygR; Prpr-1::MmPylT; Pmyo-3::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*
212.9.2	
213.8.1	*Ex [Pmyo-3::Hu Smad-4 NES::PCCRS + Prps-0::hygR; Prpr-1::MmPylT; Pmyo-3::rev-Caspase-3(amber)::F2A::mCherry::HA::egl-13 NLS::SL2::CFP]*

Table A2. Bombardments for all mentioned strains

Bombardment No.	Plasmid 1	Plasmid 2	Plasmid 3	Plasmid 4	Background
2	4 μg SG88	1 μg ZX36	2.5 μg IR11	2.5 μg IR9	*smg-2 (e2008)*
3	4 μg SG88	1 μg ZX37	2.5 μg IR11	2.5 μg IR9	*smg-2 (e2008)*
4	4 μg SG88	1 μg ZX40	2.5 μg IR11	2.5 μg IR9	*smg-2 (e2008)*
5	4 μg SG88	1 μg ZX43	2.5 μg IR11	2.5 μg IR9	*smg-2 (e2008)*
20	4 μg ZX107	1 μg ZX43	2.5 μg IR11	2.5 μg IR9	*smg-2 (e2008)*
25	4 μg ZX106	1 μg ZX43	2.5 μg IR11	2.5 μg IR9	*smg-2 (e2008)*
26, 58	4 μg ZX104	1 μg ZX43	2.5 μg IR11	2.5 μg IR9	*smg-2 (e2008)*
27, 59	4 μg ZX105	1 μg ZX43	2.5 μg IR11	2.5 μg IR9	*smg-2 (e2008)*
52	4 μg ZX122	1 μg ZX130	2.5 μg IR11	2.5 μg IR9	ZG610
64	4 μg ZX123	1 μg ZX130	2.5 μg IR11	2.5 μg IR9	ZG610
84	4 μg ZX125	1 μg ZX129	5 μg ZX154	/	ZG610
91	4 μg ZX156	1 μg ZX141	5 μg ZX154	/	*smg-2 (e2008)*
92	4 μg ZX175	1 μg ZX177	5 μg ZX180	/	*smg-2 (e2008)*
132	4 μg ZX225	1 μg ZX223	5 μg ZX224	/	N2
133	3 μg ZX225	2 μg ZX223	5 μg ZX224	/	N2
136	4 μg ZX122	1 μg ZX213	5 μg ZX154	/	N2
137	4 μg ZX122	1 μg ZX214	5 μg ZX154	/	N2
138	4 μg ZX122	1 μg ZX215	5 μg ZX154	/	N2
139	4 μg ZX225	1 μg ZX223	5 μg ZX224	/	N2
140	4 μg ZX228	1 μg ZX226	5 μg ZX227	/	N2
141	3 μg ZX228	2 μg ZX226	5 μg ZX227	/	N2
144	4 μg ZX229	1 μg ZX212	5 μg ZX210	/	N2
166	3 μg ZX251	2 μg ZX246	5 μg ZX250	/	N2
180	4 μg ZX233	1 μg ZX235	5 μg ZX234	/	N2
182	4 μg ZX259	1 μg ZX257	5 μg ZX258	/	N2
184	4 μg ZX271	1 μg ZX272	5 μg ZX273	/	N2

186	4 µg ZX269	1 µg ZX267	5 µg ZX268	/	N2
194	3 µg ZX122	2 µg ZX141	5 µg SG322	5 µg LD49	AX204
195	3 µg ZX122	2 µg ZX141	5 µg SE149	5 µg ZX295	AX204
196	3 µg ZX122	2 µg ZX141	5 µg SE150	5 µg ZX297	AX204
197	3 µg ZX122	2 µg KB136	5 µg SG322	5 µg LD49	AX204
198	3 µg ZX122	2 µg KB136	3 µg ZX122	2 µg KB136	AX204
199	3 µg ZX122	2 µg KB136	5 µg SE150	5 µg ZX297	AX204
200	3 µg ZX122	2 µg KB137	5 µg SG322	5 µg LD49	AX204
201	3 µg ZX122	2 µg KB137	5 µg SE149	5 µg ZX295	AX204
202	3 µg ZX122	2 µg KB137	5 µg SE150	5 µg ZX297	AX204
203	3 µg ZX229	2 µg ZX212	5 µg SG322	5 µg LD49	N2
204	3 µg ZX229	2 µg ZX212	5 µg SE149	5 µg ZX295	N2
205	3 µg ZX229	2 µg ZX212	5 µg SE150	5 µg ZX297	N2
206	3 µg ZX229	2 µg KB138	5 µg SG322	5 µg LD49	N2
207	3 µg ZX229	2 µg KB138	5 µg SE149	5 µg ZX295	N2
208	3 µg ZX229	2 µg KB138	5 µg SE150	5 µg ZX297	N2
209	3 µg ZX229	2 µg KB139	5 µg SG322	5 µg LD49	N2
210	3 µg ZX229	2 µg KB139	5 µg SE149	5 µg ZX295	N2
211	3 µg ZX229	2 µg KB139	5 µg SE150	5 µg ZX297	N2
212	3 µg ZX238	2 µg KB134	5 µg SG322	5 µg LD49	N2
213	3 µg ZX238	2 µg KB135	5 µg SG322	5 µg LD49	N2
214	3 µg ZX229	2 µg ZX212	5 µg SG322	5 µg LD49	*smg-2 (e2008)*
225	3 µg ZX238	2 µg ZX147	5 µg SG322	5 µg LD49	N2

Table A3. Plasmids for all mentioned bombardments

Plasmid Name	Construct carried	Parent Vector
IR9	*Prpr-1::MmPylT::sup-7 3'UTR*	pDONR p2R-p3
IR11	*4 x Prpr-1::MmPylT::sup-7 3'UTR*	pDONR p2R-p3
IR98	*Prps-0-sh::HygR::unc-54 3' UTR in plasmid backbone*	pDEST R4-R3 II
IR99	*Prps-0-sh::HygR::gpd-2/gpd-3 SL2::GFP::unc-54 3'UTR in plasmid backbone*	pDEST R4-R3 II
IR244	*NeoR in plasmid backbone*	pDEST R4-R3 II
KB134	*Pmyo-3::Hu PKIa NES PCC2RS::let858 3'UTR::HygR*	IR98
KB135	*Pmyo-3::Hu Smad4 NES PCC2RS::let858 3'UTR::HygR*	IR98
KB136	*Pgcy-32::Hu PKIa NES PCC2RS::let858 3'UTR::HygR*	IR98
KB137	*Pgcy-32::Hu Smad4 NES PCC2RS::let858 3'UTR::HygR*	IR98
KB138	*Pmec-7::Hu PKIa NES PCC2RS::let858 3'UTR::HygR*	IR98
KB139	*Pmec-7::Hu Smad4 NES PCC2RS::let858 3'UTR::HygR*	IR98
LD49	*4 x Prpr-1::MmPylT::sup-7 short 3'UTR*	pDONR p2R-p3
SE80	*Psur-5::HygR::unc-54 3' UTR in plasmid backbone*	pDEST R4-R3 II
SE82	*Pwars-1::HygR::unc-54 3' UTR in plasmid backbone*	pDEST R4-R3 II
SE114	*let858 3' UTR in plasmid backbone*	pDEST R4-R3 II
SE149	*Prpr-1::PylT M15::sup-7 short 3'UTR*	pDONR p2R-p3
SE150	*Prpr-1::Bt mttRNA Ser C15::sup-7 short 3'UTR*	pDONR p2R-p3
SG88	*Prps-0::GFP_amber_mCherry_HA_NLS::unc-54 3' UTR*	pDEST R4-R3 II
SG322	*Prpr-1::MmPylT::sup-7 short 3'UTR*	pDONR 221
SG606	*let858 3' UTR*	pDONR p2R-p3
ZX36	*Pwars-1::MmPCC2RS CeOpt::let858 3'UTR::HygR*	SE80
ZX37	*Psur-5::MmPCC2RS CeOpt::let858 3'UTR::HygR*	SE80
ZX40	*Pwars-1::MmPCC2RS CeOpt::let858 3'UTR::HygR*	SE82
ZX43	*Psur-5::MmPCC2RS CeOpt::let858 3'UTR::HygR*	SE82
ZX104	*Psur-5::mCherry_Intein(amber)_NLS_GFP::let858 3'UTR*	pDEST R4-R3 II

ZX105	*Psur-5::GFP_mCherry_Intein(amber)_NLS::let858 3'UTR*	pDEST R4-R3 II
ZX106	*Psur-5::mCherry_Intein(wt)_NLS_GFP::let858 3'UTR*	pDEST R4-R3 II
ZX107	*Psur-5::GFP_mCherry_Intein(wt)_NLS::let858 3'UTR*	pDEST R4-R3 II
ZX114	*gpd-2/gpd-3 SL2::CFP::let858 3'UTR*	pDONR p2R-p3
ZX122	*Pgcy-32::rev-Casp-3(CeOpt) amber_F2A_mCherry_NLS::gpd-2/gpd-3 SL2::CFP::let858 3'UTR*	pDEST R4-R3 II
ZX123	*Pgcy-32::rev-Casp-3(CeOpt) wt_F2A_mCherry_NLS::gpd-2/gpd-3 SL2::CFP::let858 3'UTR*	pDEST R4-R3 II
ZX125	*Pgcy-32::rev-ced-3 wt_F2A_mCherry_NLS::gpd-2/gpd-3 SL2::CFP::let858 3'UTR*	pDEST R4-R3 II
ZX129	*Pgcy-32::MmPCC2RS CeOpt::let858 3'UTR::HygR*	SE80
ZX130	*Pgcy-32::MmPCC2RS CeOpt::let858 3'UTR::HygR*	SE82
ZX141	*Pgcy-32::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX147	*Pmyo-3::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX154	*Pgcy-32::Prpr-1::MmPylT::let858 3'UTR*	pDEST R4-R3 II
ZX156	*Pgcy-32::GFP_amber_mCherry_HA_NLS::unc-54 3' UTR*	pDEST R4-R3 II
ZX175	*Pdat-1::GFP_amber_mCherry_HA_NLS::unc-54 3' UTR*	pDEST R4-R3 II
ZX177	*Pdat-1::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX180	*Pdat-1::Prpr-1::MmPylT::let858::NeoR*	IR244
ZX202	*Prps-0-sh::HygR::gpd-2/gpd-3 SL2::GC3AI::unc-54 3'UTR in plasmid backbone*	IR99
ZX205	*Prps-0-sh::HygR::gpd-2/gpd-3 SL2::ZipGFP::unc-54 3'UTR in plasmid backbone*	IR99
ZX206	*Prps-0-sh::HygR::gpd-2/gpd-3 SL2::VC3AI::unc-54 3'UTR in plasmid backbone*	IR99
ZX210	*Pmec-7::Prpr-1::MmPylT::4 x (Prpr-1::MmPylT)::let858 3'UTR*	SE114
ZX212	*Pmec-7::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX213	*Pgcy-32::MmPCC2RS CeOpt::let858 3'UTR::HygR::GC3AI*	ZX202
ZX214	*Pgcy-32::MmPCC2RS CeOpt::let858 3'UTR::HygR::ZipGFP*	ZX205
ZX215	*Pgcy-32::MmPCC2RS CeOpt::let858 3'UTR::HygR::VC3AI*	ZX206

ZX223	*Pncs-1::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX224	*Pncs-1::Prpr-1::MmPylT:: 4 x (Prpr-1::PylT)::let858 3'UTR*	SE114
ZX225	*Pncs-1::GFP_amber_mCherry_HA_NLS::unc-54 3'UTR*	pDEST R4-R3 II
ZX226	*Pttx-3::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX227	*Pttx-3::Prpr-1::MmPylT:: 4 x (Prpr-1::PylT)::let858 3'UTR*	SE114
ZX229	*Pmec-7::rev-Casp-3(CeOpt) C135amber_F2A_mCherry_NLS::gpd-2/gpd-3 SL2::CFP::let858 3'UTR*	pDEST R4-R3 II
ZX233	*ser-2prom2::GFP_amber_mCherry_HA_NLS::unc-54 3'UTR*	pDEST R4-R3 II
ZX234	*ser-2prom2::Prpr-1::MmPylT:: 4 x (Prpr-1::MmPylT)::let858 3'UTR*	SE114
ZX235	*ser-2prom2::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX238	*Pmyo-3::rev-Casp-3(CeOpt) C135amber_F2A_mCherry_NLS::gpd-2/gpd-3 SL2::CFP::let858 3'UTR*	pDEST R4-R3 II
ZX246	*hen-1promA::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX250	*hen-1promA::Prpr-1::MmPylT:: 4 x (Prpr-1::MmPylT)::let858 3'UTR*	SE114
ZX251	*hen-1promA::GFP_amber_mCherry_HA_NLS::unc-54 3'UTR*	pDEST R4-R3 II
ZX257	*Pflp-18::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX258	*Pflp-18::Prpr-1::MmPylT:: 4 x (Prpr-1::MmPylT)::let858 3'UTR*	SE114
ZX259	*Pflp-18::GFP_amber_mCherry_HA_NLS::unc-54 3'UTR*	pDEST R4-R3 II
ZX268	*Pinx-7 long::Prpr-1::MmPylT:: 4 x (Prpr-1::MmPylT)::let858 3'UTR*	SE114
ZX269	*Pinx-7 long::GFP_amber_mCherry_HA_NLS::unc-54 3'UTR*	pDEST R4-R3 II
ZX271	*Pinx-7 long::MmPCC2RS CeOpt::let858 3'UTR::HygR*	IR98
ZX272	*Pinx-7 short::Prpr-1::MmPylT:: 4 x (Prpr-1::MmPylT)::let858 3'UTR*	SE114
ZX273	*Pinx-7 short::GFP_amber_mCherry_HA_NLS::unc-54 3'UTR*	pDEST R4-R3 II
ZX295	*4 x Prpr-1::PylT M15::sup-7 short 3'UTR*	pDEST R4-R3 II
ZX297	*4 x Prpr-1::Bt mttRNA Ser C15::sup-7 short 3'UTR*	pDEST R4-R3 II

Table A4. Primers for strain generation

Primer Name	Usage	Type	Sequence
Z6	make SL2::CFP::let-858 3'UTR	Forward	GGGGACAGCTTTCTTGTACAAAGTGGCTGTC TCATCCTACTTTCACC
Z7	make SL2::CFP::let-858 3'UTR	Reverse	AGCGTAGTCTGGGACATCATATGG
Z8	make SL2::CFP::let-858 3'UTR	Forward	CCATATGATGTCCCAGACTACGCTTAACGTGA AGTGGAATCGGATGATCG
S9	sequence mCherry	Forward	AAAGATGGAGGACATTACGATG
Z13	sequence GFP	Forward	TTAAATTTTCAGGAGGACGG
Z14	amplify Pwars-1	Reverse	GGGGACTGCTTTTTTGTACAAACTTGGCTGT GTTGAACCCTGAAAAAAATAAATTGGGG
S195	amplify Pwars-1	Forward	GGGGACAACTTTGTATAGAAAAGTTGactagtA AGAGCCACCACCGAAATAGATG
Z18	make let-858 3'UTR	Reverse	GGGGACAACTTTGTATAATAAAGTTGATACGG ATTCGCATTTGCCAAG
S203	sequence Pwars-1	Forward	CAATAATTTCAATTTTTGTCGCTC
S204	sequence Psur-5	Forward	CTCCGAATTGTTGTGTACCTAC
Z21	sequence Pwars-1	Forward	GCTTCGTAAAATTCACCCC
Z24	sequence Mm PCCRS	Reverse	CATTTGggcGAActgGAGCATGGTGAACTCCTC GAGGTGC
S58	make mCherry::HA	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTTT AAGCGTAGTCTGGGACATCATATG
S41	sequence Prps-0	Forward	GTTTTCTTGGTTATGACTGATCTG
S192	make Psur-5::HygR	Reverse	TCAACTCGGGTTTTTTCATTCTGAAAACAAAA TGTAAAGTTCAAAGGTTTATCCACAGA
S191	make Psur-5::HygR	Forward	CTTTACATTTTGTTTTCAGAATGAAAAAACCC GAGTTGACCG
Z33	sequence GFP	Reverse	CAAGGACCCAAACGAGAAGCG
Z34	sequence mCherry	Reverse	CCCATAGTCTTCTTCTGCATAAC
Z40	make mCherry::HA	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTC

			GTTAAGCGTAGTCTGGGAC
S23	sequence Mm PCCRS	Forward	GCTCcagTTCgccCAAATGGGATCCGGATGCACCCGTGAG
Z50	amplify Pmyo-3	Forward	GGGGACAACTTTGTATAGAAAAGTTGGTATAGAAAAGTTGGTGCAGGC
Z51	amplify Pmyo-3	Reverse	GGGGACTGCTTTTTTGTACAAACTTGCTTGTGGATCTAGTGGTCGTGGG
Z52	sequence Pmyo-3	Forward	GTCTAATTTTCAGGGCAGGG
Z54	make let-858 3'UTR	Forward	CGTGAAGTGGAATCGGATGATCGACGC
S107	sequence unc-54 3'UTR	Reverse	ACTTACTCTGATGACAGCGG
Z60	sequence mCherry	Forward	CAAAGGGTGAAGAAGATAAC
Z64	sequence Pmyo-3	Forward	CTCCGCTGACTTTAGAACCC
Z65	sequence Pmyo-3	Forward	GGTGCAAAACAGTCAAGTCG
Z66	sequence Pmyo-3	Forward	CATCGGAAAACGAAGGAGGG
Z74	amplify mCherry	Forward	GGGGACAGCTTTCTTGTACAAAGTGGGAGTCTCAAAGGGTGAAGAAGATAACATGG
Z75	amplify unc-54 3'UTR	Reverse	GGGGACAACTTTGTATAATAAAGTTGAGGAAACAGTTATGTTTGGTATATTGGGAATG
Z76	make HA::elg-13 NLS::unc-54 3'UTR	Forward	CCATATGATGTCCCAGACTACGCTAATTCCAACTGAGCGCCGGTCGCTACCATTAC
Z82	sequence GFP	Forward	CCAAGGACCCAAACGAGAAG
Z86	make HA::elg-13 NLS::unc-54 3'UTR	Reverse	GCGACCGGCGCTCAGTTGGAATTTTAATTTTCAACTTCCTTGGCAAGCTTCTTC
Z87	sequence unc-54 3'UTR	Forward	AATTCCAACTGAGCGCCGGTCGCTACCATTAC
S313	make an N-terminal part of mCherry	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTaaaaATGGTCTCCAAGGGAGAGGAG
Z96	make a C-terminal part of mCherry	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTTGTTCTCGACCTCCTTGGCG
Z97	make an N-terminal part of mCherry	Forward	GGGGACAGCTTTCTTGTACAAAGTGGTAATGGTCTCCAAGGGAGAGGAGGACAAC

Z98	make elg-13 NLS::let-858 3'UTR	Reverse	GCGTCGATCATCCGATTCCACTTCACGTTAG TTCTCGACCTCCTTGGCGAGCTTCTTG
Z99	sequence GFP	Forward	GGGGACAGCTTTCTTGTACAAAGTGGGCTCC AAGGGAGAGGAGCTCTTC
Z100	sequence mCherry::intein	Forward	GACCAGTCATGCAATAGCTCTCCTAC
Z101	sequence intein	Reverse	CAGTACTCGAAGACCTCTTG
Z102	sequence intein	Forward	CCTCGGAAAGCAAAACGTCTAC
S314	mutate Cys1amber of mutant intein back to Cys	Forward	CGACGGACCAGTCATGCAATGCCTCTCCTAC GAGACCGAGATC
S315	mutate Cys1amber of mutant intein back to Cys	Reverse	GATCTCGGTCTCGTAGGAGAGGCATTGCATG ACTGGTCCGTCG
Z149	mutate the premature amber of reversed Caspase-3 (NoOpt) back to Cys codon	Forward	CATTATTCAGGCCtgcCGTGGTACAGAACTG
Z150	mutate the premature amber of reversed Caspase-3 (NoOpt) back to Cys codon	Reverse	CAGTTCTGTACCACGgcaGGCCTGAATAATG
Z151	mutate the premature amber of reversed Caspase-3 (CeOpt) back to Cys codon	Forward	CTCTTCATCATCCAAGCCTgcCGTGGAACCGA GCTCGAC
Z152	mutate the premature amber of reversed Caspase-3 (CeOpt) back to Cys codon	Reverse	GTCGAGCTCGGTTCCACGgcAGGCTTGGATG ATGAAGAG
Z153	mutate the catalytic Cys codon of ced-3 to amber	Forward	GTTTTTGTGCAGGCTTagCGAGGCGgttcg
Z154	mutate the catalytic Cys codon of ced-3 to amber	Reverse	cgaacCGCCTCGctaAGCCTGCACAAAAAC
Z155	fuse HL to ced-3 SS	Reverse	CTCCTCCGGCTCCGGCGGTGGCTCCTCCTC CGGATGGGGCTCCGAGGATGACGGCAGAGT TTCGTGCCTCCGGCCAAAAGTAGAACTTTTT GAGCAGGCGGGATGTCATc

Z156	fuse HL with a caspase cleavage site to ced-3 LS	Forward	GAGCCACCGCCGGAGCCGGAGGAGCCGGAGGACCAGCCGGACTCATCaactttgtcgatGCACCAACCATAAGCCGTGTTTTCGACGAGAAAACCATGTACAGAAACTTCTCGAG
S320	insert F2A in-between caspase subunits	Forward	CCAggatccggaGTCAAGCAAACCCTC
S321	amplify elg-13 NLS	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAATTTTCAACTTCCTTGGCAAGCTTCTTC
S322	sequence ced-3	Forward	tttaacaaatcgagaaaaagagaatg
S323	sequence ced-3	Forward	ccacttcatcggattggtttg
S324	sequence ced-3	Forward	aagtttgcaaaacgacgc
S325	sequence ced-3	Reverse	attaggaacaaggctgggaaa
S326	amplify ced-3 SS	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTaaaaATGGGAGTTCCTGCATTTCTTCGTC
S327	fuse HL to ced-3 SS	Reverse	CTCCGGATGGGGCTCCGAGGATGACGGCAGAGTTTCGTGCCTCCGGCCAAAAGTAG
S328	fuse HL to ced-3 LS	Forward	GAGGACCAGCCGGACTCATCGCACCAACCATAAGCCGTGTTTTCGACGAG
S329	fuse F2A to ced-3 LS	Reverse	GAGGGTTTGCTTGACtccggatccTGGGTCGACAGAATCCAAGACTGGGAATC
S330	amplify ced-3 LS	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAGTCGACAGAATCCAAGACTGGGAATC
S331	fuse ced-3 cleavage site just upstream to F2A fragment	Forward	GAGGGTTTGCTTGACtccggatccTGGgactccGTCGACAGAATCCAAGACTGGGAATC
S332	insert ced-3 cleavage site after HL	Reverse	GAGGACCAGCCGGACTCATCaactttgtcgatGCACCAACCATAAGCCGTGTTTTCGACGAG
S333	fuse mCherry(CeOpt) to F2A peptide	Forward	GACGTCGAGTCCAACCCAGGACCAGTCTCAAAGGGTGAAGAAGATAACATG
S334	amplify Caspase-3 SS as an insert of Gateway BP reaction	Forward	GGGGACAAGTTTGTACAAAAAAGCAG
S335	sequence Caspase-3(nonOpt)	Forward	TCATACATGGAAGCGAATCAA

Z173	sequence ced-3 LS	Forward	CAAGCTGACATTCTGATTGC
Z174	amplify SL2::CFP	Reverse	CATTTTTTCTACCGGTACAG
Z175	amplify SL2::CFP	Forward	GAAACTGCTGTACCGGTAGAAAAAATGTCAAAGGGAGAAGAACTGTTCAC
Z176	sequence CFP	Forward	GTTCATCTGCACCACCGGAAAG
Z177	insert Cys1amber intein into Casp3(nonOpt) LS	Reverse	CGTACTCGACGGTGAGGATCTCGGTCTCGTAGGAGAGctaGGCCTGAATAATGAAAAGTTTGG
Z178	insert Cys1amber intein into Casp3(nonOpt) LS	Forward	CACAACTTCGCCCTCAAGAACGGATTCATCGCCTCCAACtagCGTGGTACAGAACTGGACTG
Z179	insert Cys1amber intein into Casp3(CeOpt) LS	Reverse	CGTACTCGACGGTGAGGATCTCGGTCTCGTAGGAGAGctaGGCTTGGATGATGAAGAGCTTTGG
Z180	insert Cys1amber intein into Casp3(CeOpt) LS	Forward	CACAACTTCGCCCTCAAGAACGGATTCATCGCCTCCAACTAGCGTGGAACCGAGCTCGACTG
Z181	sequence Pgcy-32	Forward	gacacgccatcttctcgaaac
Z185	sequence Caspase-3(CeOpt) LS	Forward	CTCATGCGTGACGTCTCCAAG
Z188	sequence ced-3 LS	Reverse	TCAGTCGACAGAATCCAAGACTG
Z194	sequence mCherry	Reverse	GCTTTGTAAGTTGTCTTCAC
Z195	amplify SL2	Forward	GGGGACAGCTTTCTTGTACAAAGTGGTAGCTGTCTCATCCTACTTTCAC
Z200	sequence CFP	Reverse	CTCTTGGACGTATCCCTCTG
Z201	amplify GFP	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGTCCAAGGGAGAGGAGCTCTTCACCGGAGTC
Z205	amplify ced-3 SS	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTaaaaATGG
Z206	amplify ced-3	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAaaaatgatgcgtcaagatagaaggagcttg
Z207	amplify ced-3	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTTtt

			agacggcagagtttcgtgcttccggccaaaagtag
Z208	sequence ced-3	Forward	cagcgtctcttcagtgtcatc
Z209	sequence ced-3	Forward	gagagacgcggagaaatacc
Z210	mutate the catalytic Cys codon of ced-3 to amber	Forward	CCAAAGATCGTCTTCGTCCAAGCCTAGCGTGGAGAGCGTCGTGACAACGGATTC
Z211	mutate the catalytic Cys codon of ced-3 to amber	Reverse	CGACGCTCTCCACGctAGGCTTGGACGAAGAC
Z212	sequence ced-3	Reverse	gtgttggcatctgctcaaag
Z216	sequence ced-3 LS	Forward	CGAGCACTTCGAGCAAATGC
Z218	amplify Prps-0	Forward	GGGGACAACTTTGTATAGAAAAGTTGTAATGAAGATGAACGCAAGGAGCCTGATGCAG
Z219	amplify Prps-0	Reverse	GGGGACTGCTTTTTTGTACAAACTTGTATTACCTTAAAATTCAAAAATTAATTTCAGATCAGTC
Z220	sequence Pdat-1	Forward	ctacctcttcccattctttc
Z243	sequence Pmec-7	Forward	GGGGACAACTTTGTATAGAAAAGTTGTAgtttcaagatgaaacgttttgtgtgtagc
Z244	sequence Pmec-7	Reverse	GGGGACTGCTTTTTTGTACAAACTTGTcgacgttcttcctctacacctaca
Z245	sequence Pmec-7	Forward	cccaccagttgtctcttcac
Z252	amplify VC3AI	Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGATCAAGATCGCCACCCGTAAG
Z253	amplify VC3AI	Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTAGTTTGGGAGGTTGTCGACACGC
Z254	amplify Pttx-3	Forward	GGGGACAACTTTGTATAGAAAAGTTGTAatgcataatttttttggaagacctgaac
Z255	amplify Pttx-3	Reverse	GGGGACTGCTTTTTTGTACAAACTTGTcatttgacaccgaagacaattattatg
Z256	sequence Pttx-3	Forward	gctcattgatctgtaacgtg
Z257	sequence Pttx-3	Forward	gttctaacctttaactgtttgg
Z258	sequence Pttx-3	Forward	gagtgagcatctattacagcg
Z259	amplify ser-2prom2	Forward	GGGGACAACTTTGTATAGAAAAGTTGTAgtctgg

			taagttgaacatggagtc
Z260	amplify ser-2prom2	Reverse	GGGGACTGCTTTTTTGTACAAACTTGTtcatttttg caaattacttgaggctgccttg
Z261	amplify ser-2prom2	Forward	GGGGACAACTTTGTATAGAAAAGTTGTAattttat gactttcactagaaatg
Z262	sequence ser-2prom2	Forward	gtcgtcaaaataactgatcg
Z263	sequence ser-2prom2	Forward	gtacatgtgcaaataaggtg
Z264	sequence ser-2prom2	Forward	cctgttttatcatttttccgac
Z265	sequence ser-2prom2	Forward	ctgaaagctaaggcaatttg
Z266	sequence ser-2prom2	Forward	gatgctaggcttcgtcgattg
Z267	sequence ser-2prom2	Forward	gctttttcgggcaacaaatg
L92	sequence Pncs-1	Forward	cgccgaaacacatcg
Z274	sequence ser-2prom2	Reverse	caccaagaccgtaaacacaac
Z275	amplify hen-1promA	Forward	GGGGACAACTTTGTATAGAAAAGTTGTAtaatttt gaattgccgcacaaattcttgg
Z276	amplify hen-1promA	Reverse	GGGGACTGCTTTTTTGTACAAACTTGTggtggg gagatttctgaaaataatttattgtag
Z277	sequence hen-1promA	Forward	ctctataatctatacgcattg
Z278	sequence hen-1promA	Forward	caactgaaattttctaacatc
Z279	sequence hen-1promA	Forward	catgtaatagtcttattactc
Z280	sequence hen-1promA	Forward	gttaatttctgatgcctagag
Z281	sequence hen-1promA	Forward	caaccatttccgcgattagg
Z291	amplify Pinx-7 (long)	Forward	GGGGACAACTTTGTATAGAAAAGTTGTAtaattc agttagatgaccgtataatttagtattccagcac
Z292	amplify Pinx-7 (short)	Forward	GGGGACAACTTTGTATAGAAAAGTTGTAactcat tcaaatcttataagcatccttgcctg
Z293	amplify Pinx-7	Reverse	GGGGACTGCTTTTTTGTACAAACTTGTctcggg agagaatgtgagaacttggaag
Z294	sequence Pinx-7	Forward	gatgaagctcggttgcagtg
Z336	sequence mCherry	Forward	ATGGTCTCAAAGGGTGAAGAAG
Z337	fuse a caspase cleavage site	Forward	CCAGTCCTCGACTCCGTCGACGGAATCCTC

	to HL		GGAGCCCCATCC
Z342	sequence the C15 variant of Pyrrolysyl-tRNA	Forward	CGTCAAGTTGTATGGGAGAG

Reference

1. Schrodinger, E., *What is life*. 1944, Cambridge: Cambridge University Press.
2. Wei, G. and A.P. Mahowald, *The germline: familiar and newly uncovered properties.* Annu Rev Genet, 1994. **28**(309-24).
3. Crick, F., *Central Dogma of Molecular Biology.* Nature, 1970. **227**(5258): p. 561-3.
4. Crick, F.H.C., et al., *General nature of the genetic code for proteins.* Nature, 1961. **192**: p. 1227-32.
5. Söll, D., et al., *Studies on polynucleotides, XLIX. Stimulation of the binding of aminoacyl-sRNA's to ribosomes by ribotrinucleotides and a survey of codon assignments for 20 amino acids.* Proc Natl Acad Sci U S A, 1965. **54**(5): p. 1378-85.
6. Crick, F.H.C., *Codon-anticodon pairing: The wobble hypothesis.* J Mol Biol, 1966. **19**(2): p. 548-55.
7. Holley, R.W., et al., *Structure of a Ribonucleic Acid.* Science, 1965. **147**(3664): p. 1462-5.
8. Ibba, M. and D. Söll, *Aminoacyl-tRNA synthesis.* Annu Rev Biochem, 2000. **69**(1): p. 617-50.
9. Beringer, M. and M.V. Rodnina, *The Ribosomal Peptidyl Transferase.* Molecular Cell, 2007. **26**(3): p. 311-321.
10. Kim, S.H., et al., *Three-Dimensional Tertiary Structure of Yeast Phenylalanine Transfer RNA.* Science, 1974. **185**(4149): p. 435-40.
11. Giegé, R., M. Sissler, and C. Florentz, *Universal rules and idiosyncratic features in tRNA identity.* Nucleic Acids Res, 1998. **26**(22): p. 5017-66.
12. Crothers, D.M., T. Seno, and G. Söll, *Is there a discriminator site in transfer RNA?* Proc Natl Acad Sci U S A, 1972. **69**(10): p. 3063-7.
13. Arnez, J.G. and D. Moras, *Structural and functional considerations of the aminoacylation reaction.* Trends Biochem Sci, 1997. **22**(6): p. 211-6.
14. Francklyn, C.S. and P. Mullen, *Progress and challenges in aminoacyl-tRNA synthetase-based therapeutics.* Journal of Biological Chemistry, 2019.

294(14): p. 5365-5385.

15. Fersht, A.R., et al., *Reconstruction by site-directed mutagenesis of the transition state for the activation of tyrosine by the tyrosyl-tRNA synthetase: a mobile loop envelopes the transition state in an induced-fit mechanism.* Biochemistry, 1988. **27**(5): p. 1581-7.
16. Fersht, A.R., *Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase.* Biochemistry, 1977. **16**(5): p. 1025-30.
17. Nureki, O., et al., *Enzyme Structure with Two Catalytic Sites for Double-Sieve Selection of Substrate.* Science, 1998. **280**(5363): p. 578-82.
18. Crick, F.H.C., *The origin of the genetic code.* J Mol Biol, 1968. **38**(3): p. 367-79.
19. Woese, C.R., *On the evolution of the genetic code.* Proc Natl Acad Sci U S A, 1965. **54**(6): p. 1546-52.
20. Ambrogelly, A., S. Palioura, and D. Soll, *Natural expansion of the genetic code.* Nat Chem Biol, 2007. **3**(1): p. 29-35.
21. Barrell, B.G., A.T. Bankier, and J. Drouin, *A different genetic code in human mitochondria.* Nature, 1979. **282**(5735): p. 189-94.
22. Macino, G., et al., *Use of the UGA Terminator as a Tryptophan Codon in Yeast Mitochondria.* Proc Natl Acad Sci U S A, 1979. **76**(8): p. 3784-5.
23. Sagan, L., *On the origin of mitosing cells.* J Theor Biol, 1967. **14**(3): p. 255-74.
24. Hällberg, B.M. and N.-G. Larsson, *Making Proteins in the Powerhouse.* Cell Metabolism, 2014. **20**(2): p. 226-240.
25. Shokolenko, I.N. and M.F. Alexeyev, *Mitochondrial DNA: A disposable genome?* Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2015. **1852**(9): p. 1805-1809.
26. Youle, R.J., *Mitochondria—Striking a balance between host and endosymbiont.* Science, 2019. **365**(6454): p. eaaw9855.
27. Kurland, C.G., *Evolution of mitochondrial genomes and the genetic code.* BioEssays, 1992. **14**(10): p. 709-14.
28. Gerber, A.P. and W. Keller, *An adenosine deaminase that generates inosine*

at the wobble position of tRNAs. Science, 1999. **286**(5442): p. 1146-9.

29. Yarus, M., *Translational Efficiency of Transfer RNA's: Uses of an Extended Anticodon.* Science, 1982. **218**(4573): p. 646-52.
30. Agris, P.F., *Decoding the genome: a modified view.* Nucleic Acids Research, 2004. **32**(1): p. 223-238.
31. Jacobs, H.T., et al., *Nucleotide sequence and gene organization of sea urchin mitochondrial DNA.* J Mol Biol, 1988. **202**(2): p. 185-217.
32. Weigert, M.G. and A. Garen, *Base composition of nonsense codons in E. coli. Evidence from amino-acid substitutions at a tryptophan site in alkaline phosphatase.* Nature, 1965. **206**: p. 992-4.
33. Brenner, S., A.O. Stretton, and S. Kaplan, *Genetic code: the 'nonsense' triplets for chain termination and their suppression.* Nature, 1965. **206**: p. 994-8.
34. Brenner, S., et al., *UGA: a third nonsense triplet in the genetic code.* Nature, 1967. **213**: p. 449-50.
35. Caskey, C.T., *Peptide chain termination.* Trends Biochem Sci, 1980. **5**(9): p. 234-7.
36. Benzer, S. and S.P. Champe, *A Change from Nonsense to Sense in the Genetic Code.* Proc Natl Acad Sci U S A, 1962. **48**(7): p. 1114-21.
37. Brenner, S., A.O. Stretton, and S. Kaplan, *Genetic code: the 'nonsense' triplets for chain termination and their suppression.* Nature, 1965. **206**(988): p. 994-8.
38. Piper, P.W., et al., *Nonsense suppressors of Saccharomyces cerevisiae can be generated by mutation of the tyrosine tRNA anticodon.* Nature, 1976. **262**(5571): p. 757-61.
39. Wills, N., et al., *The genes sup-7 X and sup-5 III of C. elegans suppress amber nonsense mutations via altered transfer RNA.* Cell, 1983. **33**(2): p. 575-83.
40. Chambers, I., et al., *The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA.* EMBO J, 1986. **5**(6): p. 1221-7.

41. Zinoni, F., et al., *Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from Escherichia coli.* Proc Natl Acad Sci U S A, 1986. **83**(13): p. 4650-4.
42. Hao, B., et al., *A new UAG-encoded residue in the structure of a methanogen methyltransferase.* Science, 2002. **296**(5572): p. 1462-6.
43. Srinivasan, G., C.M. James, and J.A. Krzycki, *Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA.* Science, 2002. **296**(5572): p. 1459-62.
44. Leinfelder, W., et al., *Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine.* Nature, 1988. **331**(6158): p. 723-5.
45. Ganichkin, O.M., et al., *Structure and Catalytic Mechanism of Eukaryotic Selenocysteine Synthase.* Journal of Biological Chemistry, 2008. **283**(9): p. 5849-5865.
46. Driscoll, D.M. and P.R. Copeland, *Mechanism and regulation of selenoprotein synthesis.* Annual Review of Nutrition, 2003. **23**(1): p. 17-40.
47. Mansell, J.B., et al., *A dynamic competition between release factor 2 and the tRNA(Sec) decoding UGA at the recoding site of Escherichia coli formate dehydrogenase H.* EMBO J, 2001. **20**(24): p. 7284-93.
48. Thanbichler, M. and A. Böck, *Functional Analysis of Prokaryotic SELB proteins.* BioFactors, 2001. **14**(1-4): p. 53-9.
49. Longstaff, D.G., et al., *In vivo contextual requirements for UAG translation as pyrrolysine.* Molecular Microbiology, 2007. **63**(1): p. 229-241.
50. Namy, O., et al., *Adding pyrrolysine to the Escherichia coligenetic code.* FEBS Letters, 2007. **581**(27): p. 5282-5288.
51. Yuan, J., et al., *Distinct genetic code expansion strategies for selenocysteine and pyrrolysine are reflected in different aminoacyl-tRNA formation systems.* FEBS Letters, 2010. **584**(2): p. 342-349.
52. de la Torre, D. and J.W. Chin, *Reprogramming the genetic code.* Nature Reviews Genetics, 2020. **22**(3): p. 169-184.

53. Kleina, L.G., et al., *Construction of Escherichia coli amber suppressor tRNA Genes: II 1. Synthesis of additional tRNA genes and improvement of suppressor efficiency.* J Mol Biol, 1990. **213**(4).

54. Kleina, L.G., et al., *Construction of Escherichia coli amber suppressor tRNA genes. III. Determination of tRNA specificity.* J Mol Biol, 1990. **213**(4): p. 719-26.

55. Blattner, F.R., et al., *The Complete Genome Sequence of Escherichia coli K-12.* Science, 1997. **277**(5331): p. 1453-69.

56. Liu, D.R., et al., *Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo.* Proc Natl Acad Sci U S A, 1997. **94**: p. 10092-7.

57. Rould, M.A., et al., *Structure of E. Coli Glutaminyl-tRNA Synthetase Complexed with tRNA Gln and ATP at 2.8 A Resolution.* Science, 1989. **246**(4934): p. 1135-1142.

58. Liu, D.R. and P.G. Schultz, *Progress Toward the Evolution of an Organism with an Expanded Genetic Code.* Proc Natl Acad Sci U S A, 1999. **96**(9): p. 4780-5.

59. Steer, B.A. and P. Schimmel, *Major Anticodon-binding Region Missing from an Archaebacterial tRNA Synthetase.* J Biol Chem, 1999. **274**(50): p. 35601-6.

60. Wang, L., et al., *Expanding the Genetic Code of Escherichia coli.* Science, 2001. **292**: p. 498-500.

61. Wang, L. and P.G. Schultz, *A general approach for the generation of orthogonal tRNAs.* Chem Biol, 2001. **8**(9): p. 883-90.

62. Chin, J.W., et al., *Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli.* Proc Natl Acad Sci U S A, 2002. **99**(17): p. 11020-4.

63. Chin, J.W., et al., *Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli.* J Am Chem Soc, 2002. **124**(31): p. 9026-7.

64. Deiters, A. and P.G. Schultz, *In vivo incorporation of an alkyne into proteins in Escherichia coli.* Bioorganic & Medicinal Chemistry Letters, 2005. **15**(5): p. 1521-1524.

65. Fechter, P., et al., *Major tyrosine identity determinants in Methanococcus jannaschii and Saccharomyces cerevisiae tRNA(Tyr) are conserved but expressed differently.* European journal of biochemistry/FEBS, 2001. **268**(3): p. 761-7.
66. Chin, J.W., et al., *An Expanded Eukaryotic Genetic Code.* Science, 2003. **301**(5635): p. 964-7.
67. Wu, N., et al., *A Genetically Encoded Photocaged Amino Acid.* J Am Chem Soc, 2004. **126**(44): p. 14306-7.
68. Rivera, A.L., et al., *Physical methods for genetic transformation of fungi and yeast.* Physics of Life Reviews, 2014. **11**(2): p. 184-203.
69. Hug, L.A., et al., *A new view of the tree of life.* Nature Microbiology, 2016. **1**(5).
70. Blight, S.K., et al., *Direct charging of tRNA(CUA) with pyrrolysine in vitro and in vivo.* Nature, 2004. **431**(7006): p. 333-5.
71. Polycarpo, C.R., et al., *Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase.* FEBS Letters, 2006. **580**(28-29): p. 6695-6700.
72. Neumann, H., S.Y. Peak-Chew, and J.W. Chin, *Genetically encoding Nε-acetyllysine in recombinant proteins.* Nature Chemical Biology, 2008. **4**(4): p. 232-234.
73. Mukai, T., et al., *Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases.* Biochem Biophys Res Commun, 2008. **371**(4): p. 818-22.
74. Nguyen, D.P., et al., *Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry.* J Am Chem Soc, 2009. **131**(25): p. 8720-1.
75. Nguyen, D.P., et al., *Genetically encoding N(epsilon)-methyl-L-lysine in recombinant histones.* J Am Chem Soc, 2009. **131**(40): p. 14194-5.
76. Kavran, J.M., et al., *Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation.* Proceedings of the National Academy

of Sciences, 2007. **104**(27): p. 11268-11273.

77. Li, W.-T., et al., *Specificity of Pyrrolysyl-tRNA Synthetase for Pyrrolysine and Pyrrolysine Analogs.* Journal of Molecular Biology, 2009. **385**(4): p. 1156-1164.

78. Lee, M.M., et al., *Structure of Desulfitobacterium hafniense PylSc, a pyrrolysyl-tRNA synthetase.* Biochemical and Biophysical Research Communications, 2008. **374**(3): p. 470-474.

79. Suzuki, T., et al., *Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase.* Nat Chem Biol, 2017.

80. O'Donoghue, P., et al., *Near-cognate suppression of amber, opal and quadruplet codons competes with aminoacyl-tRNAPylfor genetic code expansion.* FEBS Letters, 2012. **586**(21): p. 3931-3937.

81. Niu, W., P.G. Schultz, and J. Guo, *An expanded genetic code in mammalian cells with a functional quadruplet codon.* ACS Chem Biol, 2013. **8**(7): p. 1640-5.

82. Galli, G., H. Hofstetter, and M.L. Birnstiel, *Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements.* Nature, 1981. **294**(5842): p. 626-31.

83. Geiduschek, E.P. and G.P. Tocchini-Valentini, *Transcription by RNA polymerase III.* Annu Rev Biochem, 1988. **57**(873-914).

84. Hancock, S.M., et al., *Expanding the Genetic Code of Yeast for Incorporation of Diverse Unnatural Amino Acids via a Pyrrolysyl-tRNA Synthetase/tRNA Pair.* J Am Chem Soc, 2010. **132**(42): p. 14819-24.

85. Summerer, D., et al., *A genetically encoded fluorescent amino acid.* Proceedings of the National Academy of Sciences, 2006. **103**(26): p. 9785-9789.

86. Wang, J., J. Xie, and P.G. Schultz, *A Genetically Encoded Fluorescent Amino Acid.* J Am Chem Soc, 2006. **128**(27): p. 8738-39.

87. Xie, J., et al., *The site-specific incorporation of p-iodo-L-phenylalanine into proteins for structure determination.* Nature Biotechnology, 2004. **22**(10): p. 1297-1301.

88. Pearson, A.D., et al., *Trapping a transition state in a computationally designed protein bottle.* Science, 2015. **347**(6224): p. 863-7.
89. Li, C., et al., *Protein (19)F NMR in Escherichia coli.* Journal of the American Chemical Society, 2010: p. 321-7.
90. Alfonta, L., et al., *Site-Specific Incorporation of a Redox-Active Amino Acid into Proteins.* J Am Chem Soc, 2003. **125**(48): p. 14662-1463.
91. Fleissner, M.R., et al., *Site-directed spin labeling of a genetically encoded unnatural amino acid.* Proc Natl Acad Sci U S A, 2009. **106**(51): p. 21637-42.
92. Chin, J.W., *Expanding and reprogramming the genetic code.* Nature, 2017. **550**(7674): p. 53-60.
93. Chin, J.W. and P.G. Schultz, *In vivo photocrosslinking with unnatural amino acid mutagenesis.* Chembiochem, 2002. **4**(3): p. 1135-7.
94. Xuan, W., et al., *Genetic Incorporation of a Reactive Isothiocyanate Group into Proteins.* Angewandte Chemie International Edition, 2016. **55**(34): p. 10065-10068.
95. Patterson, D.M., L.A. Nazarova, and J.A. Prescher, *Finding the Right (Bioorthogonal) Chemistry.* ACS Chemical Biology, 2014. **9**(3): p. 592-605.
96. Lang, K. and J.W. Chin, *Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins.* Chem Rev, 2014. **114**(9): p. 4764-806.
97. Li, J., S. Jia, and P.R. Chen, *Diels-Alder reaction–triggered bioorthogonal protein decaging in living cells.* Nature Chemical Biology, 2014. **10**(12): p. 1003-1005.
98. Li, J., et al., *Palladium-triggered deprotection chemistry for protein activation in living cells.* Nature Chemistry, 2014. **6**(4): p. 352-361.
99. Greiss, S. and J.W. Chin, *Expanding the genetic code of an animal.* J Am Chem Soc, 2011. **133**(36): p. 14196-9.
100. Bianco, A., et al., *Expanding the genetic code of Drosophila melanogaster.* Nat Chem Biol., 2012. **8**(9): p. 748-50.
101. Chen, Y., et al., *Heritable expansion of the genetic code in mouse and zebrafish.* Cell Res., 2017. **27**(2): p. 294-297.
102. Brenner, S., *The worm's turn.* Curr Biol, 2002. **12**(21): p. R713.

103. Brenner, S., *The genetics of Caenorhabditis elegans.* Genetics, 1974. **77**(1): p. 71-94.
104. Corsi, A.K., B. Wightman, and M. Chalfie, *A Transparent Window into Biology: A Primer on Caenorhabditis elegans.* Genetics, 2015. **200**(2): p. 387-407.
105. Zarkower, D., *Somatic sex determination (February 10)*, in *WormBook, ed.* 2006, The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.84.1: http://www.wormbook.org.
106. Doniach, T., *Activity of the sex-determining gene tra-2 is modulated to allow spermatogenesis in the C. elegans hermaphrodite.* Genetics, 1986. **114**(1): p. 53-76.
107. Cassada, R.C. and R.L. Russell, *The dauerlarva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans.* Dev Biol, 1975. **46**(2): p. 326-42.
108. Chalfie, M., et al., *Green fluorescent protein as a marker for gene expression.* Science, 1994. **263**(5148): p. 802-5.
109. Raj, A., et al., *Imaging individual mRNA molecules using multiple singly labeled probes.* Nature Methods, 2008. **5**(10): p. 877-879.
110. Sulston, J.E. and H.R. Horvitz, *Post-embryonic cell lineages of the nematode, Caenorhabditis elegans.* Dev Biol, 1977. **56**(1): p. 110-56.
111. Kimble, J. and D. Hirsh, *The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans.* Dev Biol, 1979. **70**(2): p. 396-417.
112. Sulston, J.E., D.G. Albertson, and J.N. Thomson, *The Caenorhabditis elegans male: Postembryonic development of nongonadal structures.* Dev Biol, 1980. **78**(2): p. 542-76.
113. Sulston, J.E., et al., *The embryonic cell lineage of the nematode Caenorhabditis elegans.* Dev Biol, 1983. **100**(1): p. 64-119.
114. Gershon, D., *Studies on ageing in nematodes. I. The nematode as a model organism for ageing research.* Exp Gerontol, 1970. **5**(1): p. 7-12.
115. Sammut, M., et al., *Glia-derived neurons are required for sex-specific learning in C. elegans.* Nature, 2015. **526**(7573): p. 385-390.

116. White, J.G., et al., *The structure of the nervous system of the nematode Caenorhabditis elegans.* Philos Trans R Soc Lond B Biol Sci, 1986. **314**(1165): p. 1-340.

117. Jarrell, T.A., et al., *The connectome of a decision-making neural network.* Science, 2012. **337**(6093): p. 437-44.

118. Shaye, D.D., I. Greenwald, and K.M. Iijima, *Ortholist: A compendium of C. elegans genes with human orthologs.* PloS one, 2011. **6**(5): p. e20085.

119. Kaletta, T. and M.O. Hengartner, *Finding function in novel targets: C. elegans as a model organism.* Nature Reviews Drug Discovery, 2006. **5**(5): p. 387-399.

120. Sherrington, R., et al., *Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease.* Nature, 1995. **375**(6534): p. 754-60.

121. Greenwald, I. and D. Levitan, *Facilitation of lin-12 -mediated signalling by sel-12 , a Caenorhabditis elegans S182 Alzheimer's disease gene.* Nature, 1995. **377**(6547): p. 351-4.

122. Sundaram, M. and I. Greenwald, *Suppressors of a lin-12 Hypomorph Define Genes That Interact With Both lin-12 and glp-1 in Caenorhabditis elegans.* Genetics, 1993. **135**(3): p. 765-83.

123. Ruvkun, G., et al., *The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans.* Nature, 1997. **389**(6654): p. 994-9.

124. Nakae, J., et al., *Regulation of insulin action and pancreatic β-cell function by mutated alleles of the gene encoding forkhead transcription factor Foxo1.* Nature Genetics, 2002. **32**(2): p. 245-253.

125. Ranganathan, R., et al., *Mutations in the Caenorhabditis elegans Serotonin Reuptake Transporter MOD-5 Reveal Serotonin-Dependent and -Independent Activities of Fluoxetine.* J Neurosci, 2001. **21**(16): p. 5871-84.

126. Ahringer, J., *Reverse genetics (April 6)*, in *WormBook*. 2006, The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.47.1: http://www.wormbook.org.

127. Stewart, H.I., et al., *Most ultraviolet irradiation induced mutations in the*

nematode Caenorhabditis elegans are chromosomal rearrangements. Mutat Res., 1991. **249**(1): p. 37-54.

128. Bessereau, J.-L., et al., *Mobilization of a Drosophila transposon in the Caenorhabditis elegans germ line.* Nature, 2001. **413**(6851): p. 70-4.
129. Williams, D.C., et al., *Characterization of Mos1-Mediated Mutagenesis in Caenorhabditis elegans.* Genetics, 2005. **169**(3): p. 1779-1785.
130. Doitsidou, M., et al., *C. elegans Mutant Identification with a One-Step Whole-Genome-Sequencing and SNP Mapping Strategy.* PLoS ONE, 2010. **5**(11): p. e15435.
131. Zuryn, S., et al., *A Strategy for Direct Mapping and Identification of Mutations by Whole-Genome Sequencing.* Genetics, 2010. **186**(1): p. 427-430.
132. Mello, C.C., et al., *Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences.* EMBO J, 1991. **10**(12): p. 3959-70.
133. Mello, C. and A. Fire, *DNA Transformation*, in *Methods in Cell Biology*, H.F. Epstein and D.C. Shakes, Editors. 1995, Academic Press inc. p. 451-82.
134. Praitis, V., et al., *Creation of low-copy integrated transgenic lines in Caenorhabditis elegans.* Genetics, 2001. **157**(3): p. 1217-26.
135. Berezikov, E., *Homologous gene targeting in Caenorhabditis elegans by biolistic transformation.* Nucleic Acids Research, 2004. **32**(4): p. 40e-40.
136. Frøkjær-Jensen, C., et al., *Single-copy insertion of transgenes in Caenorhabditis elegans.* Nature Genetics, 2008. **40**(11): p. 1375-1383.
137. Frøkjær-Jensen, C., et al., *Improved Mos1-mediated transgenesis in C. elegans.* Nature Methods, 2012. **9**(2): p. 117-118.
138. Jinek, M., et al., *A Programmable Dual-RNA—Guided DNA Endonuclease in Adaptive Bacterial Immunity.* Science, 2012. **337**(6096): p. 816-21.
139. Chen, C., L.A. Fenk, and M. de Bono, *Efficient genome editing in Caenorhabditis elegans by CRISPR-targeted homologous recombination.* Nucleic Acids Research, 2013. **41**(20): p. e193-e193.
140. Dickinson, D.J., et al., *Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination.* Nature Methods, 2013.

10(10): p. 1028-1034.

141. Friedland, A.E., et al., *Heritable genome editing in C. elegans via a CRISPR-Cas9 system.* Nature Methods, 2013. **10**(8): p. 741-743.

142. Frøkjær-Jensen, C., *Exciting Prospects for Precise Engineering of Caenorhabditis elegans Genomes with CRISPR/Cas9.* Genetics, 2013. **195**(3): p. 635-642.

143. Montgomery, M.K., S. Xu, and A. Fire, *RNA as a Target of Double-Stranded RNA-Mediated Genetic Interference in Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15502-7.

144. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-11.

145. Timmons, L. and A. Fire, *Specific interference by ingested dsRNA.* Nature, 1998. **395**(6705): p. 854.

146. Tabara, H., A. Grishok, and Craig C. Mello, *RNAi in C. elegans: Soaking in the Genome Sequence.* Science, 1998. **282**(5388): p. 430-1.

147. Victor E. Velculescu, et al., *Serial Analysis of Gene Expression.* Science, 1995. **270**(5235): p. 484-7.

148. Kim, S.K., et al., *A Gene Expression Map for Caenorhabditis elegans.* Science, 2001. **293**(5537): p. 2087-92.

149. Hope, I.A., et al., *The C. elegans expression pattern database: a beginning.* Trends in genetics, 1996. **12**(9): p. 370-1.

150. Chen, N., et al., *WormBase: a comprehensive data resource for Caenorhabditis biology and genomics.* Nucleic Acids Research, 2005. **33**(Database issue): p. D383-D389.

151. Kondo, K., J. Hodgkin, and R.H. Waterston, *Differential expression of five tRNA(UAGTrp) amber suppressors in Caenorhabditis elegans.* Mol Cell Biol, 1988. **8**(9): p. 3627-35.

152. Li, L., et al., *Differential expression of individual suppressor tRNA(Trp) gene gene family members in vitro and in vivo in the nematode Caenorhabditis elegans.* Mol Cell Biol, 1998. **18**(2): p. 703-9.

153. Hunt-Newbury, R., et al., *High-Throughput In Vivo Analysis of Gene*

Expression in Caenorhabditis elegans. PLoS Biology, 2007. **5**(9): p. e237.

154. Deng, W., *Organization of the Caenorhabditis elegans small non-coding transcriptome: Genomic features, biogenesis, and expression.* Genome Research, 2005. **16**(1): p. 20-29.

155. Li, T., et al., *In vivo analysis of Caenorhabditis elegans noncoding RNA promoter motifs.* BMC Molecular Biology, 2008. **9**(1): p. 71.

156. Huang, L.S., P. Tzou, and P.W. Sternberg, *The lin-15 locus encodes two negative regulators of Caenorhabditis elegans vulval development.* Mol Biol Cell, 1994. **5**(4): p. 395-411.

157. Stinchcomb, D.T., et al., *Extrachromosomal DNA transformation of Caenorhabditis elegans.* Molecular & Cellular Biology, 1985. **5**(12): p. 3484-96.

158. Kelly, W.G., et al., *Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene.* Genetics, 1997. **146**(1): p. 227-38.

159. Radman, I., S. Greiss, and J.W. Chin, *Efficient and rapid C. elegans transgenesis by bombardment and hygromycin B selection.* PLoS One, 2013. **8**(10): p. e76019.

160. Gritz, L. and J. Davies, *Plasmid-encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in Escherichia coli and Saccharomyces cerevisiae.* Gene, 1983. **25**(2): p. 179-88.

161. Hodgkin, J., et al., *A new kind of informational suppression in the nematode Caenorhabditis elegans.* Genetics., 1989. **123**(2): p. 301-13.

162. Page, M.F., et al., *SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast.* Mol Cell Biol, 1999. **19**(9): p. 5943-51.

163. Longman, D., et al., *Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway.* Genes Dev, 2007. **21**(9): p. 1075-85.

164. Hodgkin, J., *Sex Determination in the Nematode C. elegans: Analysis of tra-3 Suppressors and Characterization of fem Genes.* Genetics, 1986.

114(1): p. 15-52.

165. Parrish, A.R., et al., *Expanding the genetic code of Caenorhabditis elegans using bacterial aminoacyl-tRNA synthetase/tRNA pairs.* ACS Chem Biol, 2012. **7**(7): p. 1292-302.

166. Davis, L., et al., *Precise optical control of gene expression in C. elegans using genetic code expansion and Cre recombinase*, U.o. Edinburgh, Editor. 2020: http://www.bioRxiv.org.

167. Radman, I., *Genetic code expansion in Caenorhabditis elegans*, in *Trinity College*. 2015, University of Cambridge: Cambridge. p. 194.

168. Sternberg, N., et al., *Site-specific recombination and its role in the life cycle of bacteriophage P1.* Cold Spring Harb Symp Quant Biol, 1981. **45**: p. 297-309.

169. Davis, L., et al., *Precise optical control of gene expression in C. elegans using genetic code expansion and Cre recombinase.* Elife, 2021. **10**(e67075).

170. Nagel, G., et al., *Light activation of channelrhodopsin-2 in excitable cells of Caenorhabditis elegans triggers rapid behavioral responses.* Curr Biol, 2005. **15**(24): p. 2279-84.

171. Schild, L.C. and D.A. Glauser, *Dual Color Neural Activation and Behavior Control with Chrimson and CoChR in Caenorhabditis elegans.* Genetics, 2015. **200**(4): p. 1029-34.

172. Sulston, J.E. and J.G. White, *Regulation and cell autonomy during postembryonic development of Caenorhabditis elegans.* Dev Biol, 1980. **78**(2): p. 577-97.

173. Fang-Yen, C., et al., *Laser Microsurgery in Caenorhabditis elegans.* Methods Cell Biol., 2012. **107**: p. 177-206.

174. Vogel, A., et al., *Mechanisms of femtosecond laser nanosurgery of cells and tissues.* Applied Physics B, 2005. **81**(8): p. 1015-1047.

175. Venugopalan, V., et al., *Role of Laser-Induced Plasma Formation in Pulsed Cellular Microsurgery and Micromanipulation.* Physical Review Letters, 2002. **88**(7).

176. Needham, D. and R.S. Nunn, *Elastic deformation and failure of lipid bilayer membranes containing cholesterol.* Biophysical journal, 1990. **58**(4): p. 997-1009.

177. Neumann, J. and R. Brinkmann, *Boiling nucleation on melanosomes and microbeads transiently heated by nanosecond and microsecond laser pulses.* Journal of biomedical optics, 2005. **10**(2): p. 024001.

178. Hall, D.H., et al., *Neuropathology of Degenerative Cell Death in Caenorhabditis elegans.* J Neurosci, 1997. **17**(3): p. 1033-45.

179. Harbinder, S., et al., *Genetically Targeted Cell Disruption in Caenorhabditis Elegans.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13128-33.

180. Chelur, D.S. and M. Chalfie, *Targeted cell killing by reconstituted caspases.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2283-8.

181. White, J.G., et al., *The structure of the nervous system of the nematode Caenorhabditis elegans.* Philos Trans R Soc Lond B Biol Sci, 1986. **314**(1165): p. 1-340.

182. Qi, Y.B., et al., *Photo-inducible cell ablation in Caenorhabditis elegans using the genetically encoded singlet oxygen generating protein miniSOG.* Proc Natl Acad Sci U S A, 2012. **109**(19): p. 7499-504.

183. Shibuya, T. and Y. Tsujimoto, *Deleterious effects of mitochondrial ROS generated by KillerRed photodynamic action in human cell lines and C. elegans.* Journal of Photochemistry and Photobiology B: Biology, 2012. **117**: p. 1-12.

184. Liao, J.C., J. Roider, and D.G. Jay, *Chromophore-assisted laser inactivation of proteins is mediated by the photogeneration of free radicals.* Proc Natl Acad Sci U S A, 1994. **91**(7): p. 2659-63.

185. Bulina, M.E., et al., *A genetically encoded photosensitizer.* Nature Biotechnology, 2006. **24**(1): p. 95-9.

186. Xiong, Y., X. Tian, and H.-w. Ai, *Molecular Tools to Generate Reactive Oxygen Species in Biological Systems.* Bioconjugate Chemistry, 2019. **30**(5): p. 1297-1303.

187. Vegh, R.B., et al., *Reactive oxygen species in photochemistry of the red*

fluorescent protein "Killer Red". Chemical Communications, 2011. **47**(17): p. 4887.

188. Williams, Daniel C., et al., *Rapid and Permanent Neuronal Inactivation In Vivo via Subcellular Generation of Reactive Oxygen with the Use of KillerRed.* Cell Reports, 2013. **5**(2): p. 553-563.
189. Kobayashi, J., et al., *A method for selective ablation of neurons in C. elegans using the phototoxic fluorescent protein, KillerRed.* Neurosci Lett, 2013. **548**: p. 261-4.
190. Landis, G.N. and J. Tower, *Superoxide dismutase evolution and life span regulation.* Mechanisms of Ageing and Development, 2005. **126**(3): p. 365-379.
191. Takemoto, K., et al., *SuperNova, a monomeric photosensitizing fluorescent protein for chromophore-assisted light inactivation.* Scientific Reports, 2013. **3**(1).
192. Shu, X., et al., *A Genetically Encoded Tag for Correlated Light and Electron Microscopy of Intact Cells, Tissues, and Organisms.* PLoS Biology, 2011. **9**(4): p. e1001041.
193. Pimenta, F.M., et al., *Oxygen-Dependent Photochemistry and Photophysics of "MiniSOG," a Protein-Encased Flavin.* Photochemistry and Photobiology, 2013. **89**(5): p. 1116-1126.
194. Borden, W.T., et al., *Dioxygen: What Makes This Triplet Diradical Kinetically Persistent?* Journal of the American Chemical Society, 2017. **139**(26): p. 9010-9018.
195. Griffiths, D.J. and D.F. Schroeter, *Introduction to quantum mechanics*. 2018: Cambridge University Press.
196. Wilkinson, F. and J.G. Brummer, *Rate constants for the decay and reactions of the lowest electronically excited singlet state of molecular oxygen in solution.* Journal of Physical and Chemical Reference Data, 1981. **10**(4): p. 809-999.
197. Westberg, M., et al., *Rational Design of an Efficient, Genetically Encodable, Protein-Encased Singlet Oxygen Photosensitizer.* Journal of the American

Chemical Society, 2015. **137**(4): p. 1632-1642.

198. Wilkinson, F., W.P. Helman, and A.B. Ross, *Rate Constants for the Decay and Reactions of the Lowest Electronically Excited Singlet State of Molecular Oxygen in Solution. An Expanded and Revised Compilation.* Journal of Physical and Chemical Reference Data, 1995. **24**(2): p. 663-677.
199. Davies, M.J., *Singlet oxygen-mediated damage to proteins and its consequences.* Biochemical and Biophysical Research Communications, 2003. **305**(3): p. 761-770.
200. Pattison, D.I., A.S. Rahmanto, and M.J. Davies, *Photo-oxidation of proteins.* Photochem. Photobiol. Sci., 2012. **11**(1): p. 38-53.
201. Edwards, S.L., et al., *A Novel Molecular Solution for Ultraviolet Light Detection in Caenorhabditis elegans.* PLoS Biology, 2008. **6**(8): p. e198.
202. Ward, A., et al., *Light-sensitive neurons and channels mediate phototaxis in C. elegans.* Nat Neurosci, 2008. **11**(8): p. 916-22.
203. Xu, S. and A.D. Chisholm, *Highly efficient optogenetic cell ablation in C. elegans using membrane-targeted miniSOG.* Scientific Reports, 2016. **6**(1).
204. Yang, W.S. and B.R. Stockwell, *Ferroptosis: Death by Lipid Peroxidation.* Trends in Cell Biology, 2016. **26**(3): p. 165-176.
205. Buytaert, E., M. Dewaele, and P. Agostinis, *Molecular effectors of multiple cell death pathways initiated by photodynamic therapy.* Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 2007. **1776**(1): p. 86-107.
206. Bargmann, C.I. and A. L., *Laser killing of cells in Caenorhabditis elegans.* Methods Cell Biol., 1995. **48**: p. 225-50.
207. Pelliccioli, A.P. and J. Wirz, *Photoremovable protecting groups: reaction mechanisms and applications.* Photochemical & Photobiological Sciences, 2002. **1**(7): p. 441-458.
208. Mayer, G. and A. Heckel, *Biologically active molecules with a "light switch".* Angew Chem Int Ed Engl, 2006. **45**(30).
209. Nagel, G., et al., *Channelrhodopsin-2, a directly light-gated cation-selective membrane channel.* Proceedings of the National Academy of Sciences, 2003. **100**(24): p. 13940-13945.

210. Boyden, E.S., et al., *Millisecond-timescale, genetically targeted optical control of neural activity.* Nature Neuroscience, 2005. **8**(9): p. 1263-1268.

211. Li, X., et al., *Fast noninvasive activation and inhibition of neural and network activity by vertebrate rhodopsin and green algae channelrhodopsin.* Proceedings of the National Academy of Sciences, 2005. **102**(49): p. 17816-17821.

212. Jarillo, J.A., et al., *Phototropin-related NPL1 controls chloroplast relocation induced by blue light.* Nature, 2001. **410**(6831): p. 952-4.

213. Kagawa, T., et al., *Arabidopsis NPL1: A Phototropin Homolog Controlling the Chloroplast High-Light Avoidance Response.* Science, 2001. **291**(5511): p. 2138-41.

214. Baier, J., et al., *Singlet Oxygen Generation by UVA Light Exposure of Endogenous Photosensitizers.* Biophysical Journal, 2006. **91**(4): p. 1452-1459.

215. Strickland, D., K. Moffat, and T.R. Sosnick, *Light-activated DNA binding in a designed allosteric protein.* Proceedings of the National Academy of Sciences, 2008. **105**(31): p. 10709-10714.

216. Swartz, T.E., et al., *Vibration Spectroscopy Reveals Light-Induced Chromophore and Protein Structural Changes in the LOV2 Domain of the Plant Blue-Light Receptor Phototropin 1.* Biochemistry, 2002. **41**(23): p. 7183-9.

217. Wu, Y.I., et al., *A genetically encoded photoactivatable Rac controls the motility of living cells.* Nature, 2009. **461**(7260): p. 104-108.

218. Baker, A.S. and A. Deiters, *Optical Control of Protein Function through Unnatural Amino Acid Mutagenesis and Other Optogenetic Approaches.* ACS Chemical Biology, 2014. **9**(7): p. 1398-1407.

219. Wilcox, M., et al., *Synthesis of photolabile precursors of amino acid neurotransmitters.* The Journal of Organic Chemistry, 1990. **55**(5): p. 1585-9.

220. Mendel, D., J.A. Ellman, and P.G. Schultz, *Construction of a light-activated protein by unnatural amino acid mutagenesis.* J Am Chem Soc, 1991. **113**(7): p. 2758-60.

221. Billington, A.P., et al., *Synthesis and photochemistry of photolabile N-glycine derivatives and effects of one on the glycine receptor.* Biochemistry, 1992. **31**(24): p. 5500-7.

222. Meldrum, B.S., *Glutamate as a neurotransmitter in the brain: review of physiology and pathology.* J Nutr, 2000. **130**(4S Suppl): p. 1007S-15S.

223. Legendre, P., *The glycinergic inhibitory synapse.* Cell Mol Life Sci, 2001. **58**(5-6): p. 760-93.

224. Perry, L.J. and R. Wetzel, *Unpaired cysteine-54 interferes with the ability of an engineered disulfide to stabilize T4 lysozyme.* Biochemistry, 1986. **25**(3): p. 733-9.

225. Zubay, G., *In vitro synthesis of protein in microbial systems.* Annu Rev Genet, 1973. **7**(267-87).

226. Philipson, K.D., et al., *Incorporation of caged cysteine and caged tyrosine into a transmembrane segment of the nicotinic ACh receptor.* Am J Physiol Cell Physiol, 2001. **281**(1): p. 195-206.

227. Deiters, A., et al., *A Genetically Encoded Photocaged Tyrosine.* Angewandte Chemie International Edition, 2006. **45**(17): p. 2728-2731.

228. Brown, W., et al., *Cell-Lineage Tracing in Zebrafish Embryos with an Expanded Genetic Code.* ChemBioChem, 2018. **19**(12): p. 1244-1249.

229. Engels, J. and E.J. Schlaeger, *Synthesis, structure, and reactivity of adenosine cyclic 3', 5'-phosphate benzyl triesters.* J Med Chem, 1977. **20**(7): p. 907-11.

230. Engels, J. and R. Reidys, *Synthesis and application of the photolabile guanosine 3′, 5′-phosphoric-o-nitrobenzylester.* Experientia, 1978. **34**(1): p. 14-5.

231. Kaplan, J.H., B. Forbush, and J.F. Hoffman, *Rapid photolytic release of adenosine 5′-triphosphate from a protected analogue: Utilization by the Na:K pump of human red blood cell ghosts.* Biochemistry, 1978. **17**(10): p. 1929-35.

232. Bochet, C.G., *Photolabile protecting groups and linkers.* J. Chem. Soc., 2002. **Perkin Transactions 1**(2): p. 125-42.

233. Wootton, J.F. and D.R. Trentham, *'Caged' compounds to probe the dynamics of cellular processes- synthesis and properties of some novel photosensitive P-2-nitrobenzyl esters of nucleotides*, in *Photochemical probes in biochemistry*, P.E. Nielsen, Editor. 1989, Springer-Verlag.

234. Pan, P. and H. Bayley, *Caged cysteine and thiophosphoryl peptides.* FEBS letters, 1997. **405**(1): p. 81-5.

235. Adams, S.R. and R.Y. Tsien, *Controlling Cell Chemistry with Caged Compounds.* Annu Rev Physiol., 1993. **55**(1): p. 755-84.

236. Kuluncsics, Z., et al., *Wavelength dependence of ultraviolet-induced DNA damage distribution: involvement of direct or indirect mechanisms and possible artefacts.* J Photochem Photobiol B, 1999. **49**(1): p. 71-80.

237. Jean-Roch Meunier, Alain Sarasin, and L. Marrot., *Photogenotoxicity of Mammalian Cells: A Review of the Different Assays for In Vitro Testing.* Photochemistry and Photobiology, 2002. **75**(5): p. 437-47.

238. Dahle, J., et al., *Bystander Effects may Modulate Ultraviolet A and B Radiation-Induced Delayed Mutagenesis.* Radiation Research, 2005. **163**(3): p. 289-295.

239. Nerbonne, J.M., et al., *New photoactivatable cyclic nucleotides produce intracellular jumps in cyclic AMP and cyclic GMP concentrations.* Nature, 1984. **310**(5972): p. 74-6.

240. Ren, W., A. Ji, and H.W. Ai, *Light activation of protein splicing with a photocaged fast intein.* J Am Chem Soc, 2015. **137**(6): p. 2155-8.

241. Nguyen, D.P., et al., *Genetic encoding of photocaged cysteine allows photoactivation of TEV protease in live mammalian cells.* J Am Chem Soc, 2014. **136**(6): p. 2240-3.

242. Gautier, A., et al., *Genetically Encoded Photocontrol of Protein Localization in Mammalian Cells.* J Am Chem Soc, 2010: p. 4086-8.

243. Redemann, S., et al., *Codon adaptation-based control of protein expression in C. elegans.* Nat Methods, 2011. **8**(3): p. 250-4.

244. Veljkovic, E., et al., *Functional characterization of Caenorhabditis elegans heteromeric amino acid transporters.* J Biol Chem, 2004. **279**(9): p. 7655-62.

245. Chang, H., et al., *Light-induced protein translocation by genetically encoded unnatural amino acid in Caenorhabditis elegans.* Protein & Cell, 2013. **4**(12): p. 883-886.

246. Zheng, S.Q., et al., *Drug absorption efficiency in Caenorhbditis elegans delivered by different methods.* PLoS One, 2013. **8**(2): p. e56877.

247. Davis, L. and S. Greiss, *Genetic Encoding of Unnatural Amino Acids in C. elegans*, in *Noncanonical Amino Acids* E.A. Lemke, Editor. 2018, Humana Press: New York. p. 389-408.

248. Yochem, J., T. Gu, and M. Han, *A New Marker for Mosaic Analysis in Caenorhabditis elegans Indicates a Fusion Between hyp6 and hyp7, Two Major Components of the Hypodermis.* Genetics, 1998. **149**(3): p. 1323-34.

249. Greiss, S., et al., *Transcriptional profiling in C. elegans suggests DNA damage dependent apoptosis as an ancient function of the p53 family.* BMC Genomics, 2008. **9**(1).

250. Spencer, W.C., et al., *A spatial and temporal map of C. elegans gene expression.* Genome Research, 2010. **21**(2): p. 325-341.

251. Kimble, J., et al., *Suppression of an amber mutation by microinjection of suppressor tRNA in C. elegans.* Nature, 1982. **299**(5882): p. 456-8.

252. Lyssenko, N.N., W. Hanna-Rose, and R.A. Schlegel, *Cognate putative nuclear localization signal effects strong nuclear localization of a GFP reporter and facilitates gene expression studies in Caenorhabditis elegans.* BioTechniques, 2007. **43**(5): p. 596-600.

253. Yu, S., et al., *Guanylyl Cyclase Expression in Specific Sensory Neurons: A New Family of Chemosensory Receptors.* Proc Natl Acad Sci U S A, 1997. **94**(7): p. 3384-7.

254. Gomez, M., et al., *Ca2+ signaling via the neuronal calcium sensor-1 regulates associative learning and memory in C. elegans.* Neuron, 2001. **30**(1): p. 241-8.

255. Nass, R., et al., *Neurotoxin-induced degeneration of dopamine neurons in Caenorhabditis elegans.* Proceedings of the National Academy of Sciences, 2002. **99**(5): p. 3264-3269.

256. Rogers, C., et al., *Inhibition of Caenorhabditis elegans social feeding by FMRFamide-related peptide activation of NPR-1.* Nature Neuroscience, 2003. **6**(11): p. 1178-1185.

257. Tsalik, E.L., et al., *LIM homeobox gene-dependent expression of biogenic amine receptors in restricted regions of the C. elegans nervous system.* Developmental Biology, 2003. **263**(1): p. 81-102.

258. Wenick, A.S. and O. Hobert, *Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans.* Dev Cell, 2004. **6**(6): p. 757-70.

259. Altun, Z.F., et al., *High resolution map of Caenorhabditis elegans gap junction proteins.* Dev Dyn, 2009. **238**(8): p. 1936-50.

260. Sagi, D., et al., *Tissue- and Time-Specific Expression of Otherwise Identical tRNA Genes.* PLOS Genetics, 2016. **12**(8): p. e1006264.

261. !!! INVALID CITATION !!! {Yu, 1997 #2473;Chen, 2005 #2372}.

262. Clokey, G.V. and L.A. Jacobson, *The autofluorescent "lipofuscin granules" in the intestinal cells of Caenorhabditis elegans are secondary lysosomes.* Mech Ageing Dev, 1986. **35**(1): p. 79-94.

263. Hermann, G.J., et al., *Genetic analysis of lysosomal trafficking in Caenorhabditis elegans.* Mol Biol Cell, 2005. **16**(7): p. 3273-88.

264. *WormAtlas*, Z.F. Altun, et al., Editors. 2002-2021: http://www.wormatlas.org.

265. Scolnick, E., et al., *Release Factors Differing in Specificity for Terminator Codons.* Proc Natl Acad Sci U S A, 1968. **61**(2): p. 768-74.

266. Kisselev, L.L.B., Richard H, *Translational termination comes of age.* Trends in Biochemical Sciences, 2000. **25**(11): p. 561-6.

267. Krause, M. and D. Hirsh, *The actin genes in Caenorhabditis elegans*, in *Cell and Molecular Biology of the Cytoskeleton*. 1986, Springer: Boston, MA. p. 151-78.

268. Li, M.-H., *Effects of nonionic and ionic surfactants on survival, oxidative stress, and cholinesterase activity of planarian.* Chemosphere, 2008. **70**(10): p. 1796-1803.

269. De la Parra-Guerra, A. and J. Olivero-Verbel, *Toxicity of nonylphenol and nonylphenol ethoxylate on Caenorhabditis elegans.* Ecotoxicology and Environmental Safety, 2020. **187**: p. 109709.
270. Andersen, G.R., et al., *Structural Basis for Nucleotide Exchange and Competition with tRNA in the Yeast Elongation Factor Complex eEF1A:eEF1Bα.* Mol Cell, 2000. **6**(5): p. 1261-6.
271. Albers, S., et al., *Repurposing tRNAs for nonsense suppression.* Nature Communications, 2021. **12**(1).
272. Culbertson, M.R. and P.F. Leeds, *Looking at mRNA decay pathways through the window of molecular evolution.* Current Opinion in Genetics & Development, 2003. **13**(2): p. 207-214.
273. Muir, V.S., A.P. Gasch, and P. Anderson, *The Substrates of Nonsense-Mediated mRNA Decay in Caenorhabditis elegans.* G3 (Bethesda), 2018. **8**(1): p. 195-205.
274. Sakaki, K., et al., *RNA surveillance is required for endoplasmic reticulum homeostasis.* Proc Natl Acad Sci U S A, 2012. **109**(21): p. 8079-84.
275. Sakaki, K. and R.J. Kaufman, *Interaction between quality control systems for ER protein folding and RNA biogenesis.* Worm, 2014. **2**(2): p. e23005.
276. Longman, D., et al., *Identification of a localized nonsense-mediated decay pathway at the endoplasmic reticulum.* Genes & Development, 2020. **34**(15-16): p. 1075-1088.
277. Douglas, P.M. and A. Dillin, *Protein homeostasis and aging in neurodegeneration.* Journal of Cell Biology, 2010. **190**(5): p. 719-729.
278. Son, H.G., et al., *RNA surveillance via nonsense-mediated mRNA decay is crucial for longevity in daf-2/insulin/IGF-1 mutant C. elegans.* Nature Communications, 2017. **8**(1).
279. Hsin, H. and C. Kenyon, *Signals from the reproductive system regulate the lifespan of C. elegans.* Nature, 1999. **399**(6734): p. 362-6.
280. Huang, L. and M.F. Wilkinson, *Regulation of nonsense-mediated mRNA decay.* Wiley Interdisciplinary Reviews: RNA, 2012. **3**(6): p. 807-828.
281. Nikić, I., et al., *Debugging Eukaryotic Genetic Code Expansion for*

Site-Specific Click-PAINT Super-Resolution Microscopy. Angewandte Chemie International Edition, 2016. **55**(52): p. 16172-16176.

282. Wen, W., et al., *Identification of a signal for rapid export of proteins from the nucleus.* Cell, 1995. **82**(3): p. 463-73.
283. Güttler, T., et al., *NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1.* Nature Structural & Molecular Biology, 2010. **17**(11): p. 1367-1376.
284. Ryu, Y. and P.G. Schultz, *Efficient incorporation of unnatural amino acids into proteins in Escherichia coli.* Nature Methods, 2006. **3**(4): p. 263-265.
285. Liu, C.C., et al., *Efficient expression of tyrosine-sulfated proteins in E. coli using an expanded genetic code.* Nature Protocols, 2009. **4**(12): p. 1784-1789.
286. Sakamoto, K., et al., *Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells.* Nucleic Acids Res, 2002. **30**(21): p. 4692-99.
287. Katoh, T., Y. Iwane, and H. Suga, *tRNA engineering for manipulating genetic code.* RNA Biology, 2017. **15**(4-5): p. 453-460.
288. Hoffman, K.S., A. Crnkovic, and D. Soll, *Versatility of Synthetic tRNAs in Genetic Code Expansion.* Genes (Basel), 2018. **9**(11).
289. Guo, J., et al., *Evolution of Amber Suppressor tRNAs for Efficient Bacterial Production of Proteins Containing Nonnatural Amino Acids.* Angewandte Chemie International Edition, 2009. **48**(48): p. 9148-51.
290. Serfling, R., et al., *Designer tRNAs for efficient incorporation of non-canonical amino acids by the pyrrolysine system in mammalian cells.* Nucleic Acids Res, 2017. **1**.
291. Juhling, F., et al., *tRNAdb 2009: compilation of tRNA sequences and tRNA genes.* Nucleic Acids Research, 2009. **37**(Database): p. D159-D162.
292. Ambrogelly, A., et al., *Pyrrolysine is not hardwired for cotranslational insertion at UAG codons.* Proceedings of the National Academy of Sciences, 2007. **104**(9): p. 3141-3146.
293. Charette, G.M. and M. W, *Pseudouridine in RNA: What, Where, How, and*

Why. IUBMB life, 2000. **49**(5): p. 341-51.

294. Davis, D.R. and C.D. Poulter, *1H-15N NMR studies of Escherichia coli tRNA(Phe) from hisT mutants: a structural role for pseudouridine.* Biochemistry, 1991. **30**(17): p. 4223-31.

295. Shakespeare, W., *The tragicall historie of Hamlet, Prince of Denmarke*. 1869: N. Trübner and Company.

296. Xu, X., Y. Lai, and Z.-C. Hua, *Apoptosis and apoptotic body: disease message and therapeutic target potentials.* Bioscience Reports, 2019. **39**(1).

297. Ghose, P. and S. Shaham, *Cell death in animal development.* Development, 2020. **147**(14).

298. Glücksmann, A., *Cell Deaths in Normal Vertebrate Ontogeny.* Biological Reviews, 1951. **26**(1): p. 59-86.

299. Lockshin, R.A. and C.M. Williams, *Programmed cell death-II. Endocrine potentiation of the breakdown of the intersegmental muscles of silkmoths.* Journal of Insect Physiology, 1964. **10**(4): p. 643-9.

300. Kerr, J F., A.H. Wyllie, and A.R. Currie, *Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics.* British journal of cancer, 1972. **26**(4): p. 239-57.

301. Berghe, T.V., et al., *Regulated necrosis: the expanding network of non-apoptotic cell death pathways.* Nature Reviews Molecular Cell Biology, 2014. **15**(2): p. 135-147.

302. Fuchs, Y. and H. Steller, *Live to die another way: modes of programmed cell death and the signals emanating from dying cells.* Nat Rev Mol Cell Biol, 2015. **16**(6): p. 329-44.

303. D'Arcy, M.S., *Cell death: a review of the major forms of apoptosis, necrosis and autophagy.* Cell Biology International, 2019. **43**(6): p. 582-592.

304. Sulston, J.E. and H.R. Horvitz, *Post-embryonic Cell Lineages of the Nematode, Caenorhabditis Elegans.* Dev Biol, 1977. **56**

(1): p. 110-56.

305. Giorgi, F. and P. Deri, *Cell Death in Ovarian Chambers of Drosophila Melanogaster.* J Embryol Exp Morphol, 1976. **35**(3): p. 521-33.

306. Song, W., X. Lu, and Q. Feng, *Tumor necrosis factor-α induces apoptosis via inducible nitric oxide synthase in neonatal mouse cardiomyocytes.* Cardiovascular Research, 1970. **4**(4): p. 595-602.
307. Ellis, H.M. and H.R. Horvitz, *Genetic control of programmed cell death in the nematode C. elegans.* Cell, 1986. **44**(6): p. 817-29.
308. Green, D.R. and P. Fitzgerald, *Just So Stories about the Evolution of Apoptosis.* Curr Biol, 2016. **26**(13): p. R620-R627.
309. Lettre, G. and M.O. Hengartner, *Developmental apoptosis in C. elegans: a complex CEDnario.* Nat Rev Mol Cell Biol, 2006. **7**(2): p. 97-108.
310. Wyllie, A.H., *Glucocorticoid-induced thymocyte apoptosis is associated with endogenous endonuclease activation.* Nature, 1980. **284**: p. 555-6.
311. Wyllie, A.H., J.F.R. Kerr, and A.R. Currie, *Cell Death: The Significance of Apoptosis.* 1980. **68**: p. 251-306.
312. Robertson, A.M.G. and J.N. Thomson, *Morphology of programmed cell death in the ventral nerve cord of Caenorhabditis elegans larvae.* Journal of Embryology and Experimental Morphology, 1982. **67**: p. 89-100.
313. Wang, X. and C. Yang, *Programmed cell death and clearance of cell corpses in Caenorhabditis elegans.* Cell Mol Life Sci, 2016. **73**(11-12): p. 2221-36.
314. Gartner, A., et al., *A conserved checkpoint pathway mediates DNA damage-induced apoptosis and cell cycle arrest in C. elegans.* Mol Cell, 2000. **5**(3): p. 435-43.
315. Aballay, A. and F.M. Ausubel, *Programmed cell death mediated by ced-3 and ced-4 protects Caenorhabditis elegans from Salmonella typhimurium-mediated killing.* Proceedings of the National Academy of Sciences, 2001. **98**(5): p. 2735-2739.
316. Yuan, J., et al., *The C. elegans Cell Death Gene Ced-3 Encodes a Protein Similar to Mammalian Interleukin-1 Beta-Converting Enzyme.* Cell, 1993. **75**: p. 641-52.
317. Fernandes-Alnemri, T., G. Litwack, and E.S. Alnemri, *CPP32, a novel human apoptotic protein with homology to Caenorhabditis elegans cell death protein Ced-3 and mammalian interleukin-1-beta-converting enzyme.* The Journal of

biological chemistry, 1994. **269**(49): p. 30761-4.

318. Xue, D., S. Shaham, and H.R. Horvitz, *The Caenorhabditis elegans cell-death protein CED-3 is a cysteine protease with substrate specificities similar to those of the human CPP32 protease.* Genes Dev 1996. **10**: p. 1073-83.
319. Huang, W., et al., *Mechanistic insights into CED-4-mediated activation of CED-3.* Genes Dev, 2013. **27**(18): p. 2039-48.
320. Shi, Y., *Mechanisms of Caspase Activation and Inhibition during Apoptosis.* Molecular Cell, 2002. **9**(3): p. 459-470.
321. Shi, Y., *Caspase activation: revisiting the induced proximity model.* Cell, 2004. **117**(7): p. 855-8.
322. Qi, S., et al., *Crystal Structure of the Caenorhabditis elegans Apoptosome Reveals an Octameric Assembly of CED-4.* Cell, 2010. **141**(3): p. 446-457.
323. Shaham, S., *Identification of multiple Caenorhabditis elegans caspases and their potential roles in proteolytic cascades.* Journal of Biological Chemistry, 1998. **273**(52): p. 35109-17.
324. Geng, X., et al., *Inhibition of CED-3 zymogen activation and apoptosis in Caenorhabditis elegans by caspase homolog CSP-3.* Nat Struct Mol Biol, 2008. **15**(10): p. 1094-101.
325. Geng, X., et al., *Caenorhabditis elegans caspase homolog CSP-2 inhibits CED-3 autoactivation and apoptosis in germ cells.* Cell Death & Differentiation, 2009. **16**(10): p. 1385-1394.
326. Chen, X., et al., *Regulation of CED-3 caspase localization and activation by C. elegans nuclear-membrane protein NPP-14.* Nat Struct Mol Biol, 2016. **23**(11): p. 958-964.
327. Hengartner, M.O. and H.R. Horvitz, *C. elegans cell survival gene ced-9 encodes a functional homolog of the mammalian proto-oncogene bcl-2.* Cell, 1994. **76**(4): p. 665-76.
328. Shaham, S. and H.R. Horvitz, *Developing Caenorhabditis elegans neurons may contain both cell-death protective and killer activities.* Genes & Development, 1996. **10**(5): p. 578-91.

329. Spector, M.S., et al., *Interaction between the C. elegans cell-death regulators CED-9 and CED-4.* Nature, 1997. **385**: p. 653-6.

330. Wu, D., et al., *Interaction and regulation of the Caenorhabditis elegans death protease CED-3 by CED-4 and CED-9.* J. Biol. Chem., 1997. **272**: p. 21449-54.

331. Yan, N., et al., *Structure of the CED-4-CED-9 complex provides insights into programmed cell death in Caenorhabditis elegans.* Nature, 2005. **437**(7060): p. 831-7.

332. Yang, X., H.Y. Chang, and D. Baltimore, *Essential role of CED-4 oligomerization in CED-3 activation and apoptosis.* Science, 1998. **281**: p. 1355-7.

333. Parrish, J., et al., *Demonstration of the in vivo interaction of key cell death regulators by structure-based design of second-site suppressors.* Proc Natl Acad Sci U S A, 2000. **97**(22): p. 11916-21.

334. Yan, N., et al., *Structural, biochemical, and functional analyses of CED-9 recognition by the proapoptotic proteins EGL-1 and CED-4.* Mol Cell, 2004. **15**(6): p. 999-1006.

335. Del Peso, L., V.M. González, and G. Núñez, *Caenorhabditis elegans EGL-1 disrupts the interaction of CED-9 with CED-4 and promotes CED-3 activation.* J. Biol. Chem, 1998. **273**(50): p. 33495-500.

336. Chen, F., et al., *Translocation of C. elegans CED-4 to Nuclear Membranes during Programmed Cell Death.* Science, 2000. **287**(5457): p. 1485-9.

337. Tzur, Y.B., et al., *Matefin/SUN-1 is a nuclear envelope receptor for CED-4 during Caenorhabditis elegans apoptosis.* Proc Natl Acad Sci U S A, 2006. **103**(36): p. 13397-402.

338. Yan, N., Y. Xu, and Y. Shi, *2:1 Stoichiometry of the CED4–CED9 Complex and the Tetrameric CED-4: Insights into the Regulation of CED-3 Activation.* Cell Cycle, 2006. **5**(1): p. 31-34.

339. Nicholson, D.W., *ICE/CED3-like Proteases as Therapeutic Targets for the Control of Inappropriate Apoptosis.* Nat Biotechnol, 1996. **14**(3): p. 297.

340. Xue, D. and H.R. Horvitz, *Caenorhabditis elegans CED-9 protein is a*

bifunctional cell-death inhibitor. Nature, 1997. **390**(6657): p. 305-8.

341. Minn, I.L., et al., *SUN-1 and ZYG-12, mediators of centrosome-nucleus attachment, are a functional SUN/KASH pair in Caenorhabditis elegans.* Mol Biol Cell, 2009. **20**(21).

342. Xue, D. and H.R. Horvitz, *Inhibition of the Caenorhabditis elegans cell-death protease CED-3 by a CED-3 cleavage site in baculovirus p35 protein.* Nature, 1995. **377**(6546): p. 248-51.

343. Shaham, S.H., H.Robert, *An Alternatively Spliced C. elegans ced-4 RNA Encodes a Novel Cell Death Inhibitor.* Cell, 1996. **86**(2): p. 201-8.

344. Wang, C. and R.J. Youle, *The role of mitochondria in apoptosis.* Annu Rev Genet, 2009. **43**: p. 95-118.

345. Conradt, B. and H.R. Horvitz, *The C. elegans Protein EGL-1 Is Required for Programmed Cell Death and Interacts with the Bcl-2–like Protein CED-9.* Cell, 1998. **93**(4): p. 519-29.

346. Zou, H., et al., *Apaf-1, a human protein homologous to C-elegans CED-4, participates in cytochrome c-dependent activation of caspase-3.* Cell, 1997. **90**(3): p. 405-13.

347. Cohen, G.M., *Caspases: the executioners of apoptosis.* Biochem. J., 1997. **326**: p. 1–16.

348. Kluck, R.M., et al., *The release of cytochrome c from mitochondria: A primary site for Bcl-2 regulation of apoptosis.* Science, 1997. **275**(5303): p. 1132-6.

349. Yang, J., *Prevention of apoptosis by Bcl-2: release of cytochrome c from mitochondria blocked.* Science, 1997. **275**(5303): p. 1129-32.

350. Zong, W.X., et al., *BH3-only proteins that bind pro-survival Bcl-2 family members fail to induce apoptosis in the absence of Bax and Bak.* Genes and Development, 2001. **15**(12): p. 1481-6.

351. Tan, C., et al., *Auto-activation of the Apoptosis Protein Bax Increases Mitochondrial Membrane Permeability and Is Inhibited by Bcl-2.* Journal of Biological Chemistry, 2006. **281**(21): p. 14764-14775.

352. Cory, S. and J.M. Adams, *The Bcl2 family: regulators of the cellular life-or-death switch.* Nature Reviews Cancer, 2002. **2**(9): p. 647-656.

353. Li, P., et al., *Cytochrome c and dATP-dependent formation of Apaf-1/Caspase-9 complex initiates an apoptotic protease cascade.* Cell, 1997. **91**(4): p. 479-89.

354. Kim, H.E., et al., *Formation of apoptosome is initiated by cytochrome c-induced dATP hydrolysis and subsequent nucleotide exchange on Apaf-1.* Proc Natl Acad Sci U S A, 2005. **102**(49): p. 17545-50.

355. Larisch S, et al., *A novel mitochondrial septin-like protein, ARTS, mediates apoptosis dependent on its P-loop motif.* Nat Cell Biol, 2000. **2**: p. 915-21.

356. Gottfried, Y., et al., *The mitochondrial ARTS protein promotes apoptosis through targeting XIAP.* The EMBO Journal, 2004. **23**(7): p. 1627-35.

357. Takahashi, R., et al., *A Single BIR Domain of XIAP Sufficient for Inhibiting Caspases.* Journal of Biological Chemistry, 1998. **273**(14): p. 7787-90.

358. Deveraux, Q.L., et al., *X-linked IAP is a direct inhibitor of cell-death proteases.* Nature, 1997. **388**: p. 300-4.

359. Slee, E.A., et al., *Ordering the cytochrome c-initiated caspase cascade: hierarchical activation of caspases-2, -3, -6, -7, -8, and -10 in a caspase-9-dependent manner.* J Cell Biol., 1999. **144**(2): p. 281-92.

360. Tewari, M., et al., *Yama/CPP32β, a mammalian homolog of CED-3, is a CrmA-inhibitable protease that cleaves the death substrate poly(ADP-ribose) polymerase.* Cell, 1995. **81**(5): p. 801-9.

361. Duan, H., et al., *ICE-LAP6, a Novel Member of the ICE/Ced-3 Gene Family, Is Activated by the Cytotoxic T Cell Protease Granzyme B.* Journal of Biological Chemistry, 1996. **271**(28): p. 16720-16724.

362. Srinivasula, S.M., et al., *The Ced-3/Interleukin 1β Converting Enzyme-like Homolog Mch6 and the Lamin-cleaving Enzyme Mch2α Are Substrates for the Apoptotic Mediator CPP32.* Journal of Biological Chemistry, 1996. **271**(43): p. 27099-27106.

363. Thornberry, N.A., et al., *A combinatorial approach defines specificities of members of the caspase family and granzyme B: Functional relationships established for key mediators of apoptosis.* The Journal of biological chemistry, 1997. **272**(29): p. 17907-11.

364. Walters, J., et al., *A constitutively active and uninhibitable caspase-3 zymogen efficiently induces apoptosis.* Biochemical Journal, 2009. **424**(3): p. 335-345.

365. Morgan, C.W., et al., *Turning ON Caspases with Genetics and Small Molecules*, in *Regulated Cell Death Part A: Apoptotic Mechanisms*, A. Ashkenazi, J. Yuan, and J.A. Wells, Editors. 2014. p. 179-213.

366. Putt, K.S., et al., *Small-molecule activation of procaspase-3 to caspase-3 as a personalized anticancer strategy.* Nature Chemical Biology, 2006. **2**(10): p. 543-550.

367. MacCorkle, R.A., K.W. Freeman, and D.M. Spencer, *Synthetic Activation of Caspases: Artificial Death Switches.* Proc Natl Acad Sci U S A, 1998. **95**(7): p. 3655-60.

368. Pop, C., et al., *Mutations in the Procaspase-3 Dimer Interface Affect the Activity of the Zymogen.* Biochemistry, 2003. **42**(42): p. 12311-20.

369. Chu, Y., et al., *Novel caspase-suicide proteins for tamoxifen-inducible apoptosis.* Genesis, 2008. **46**(10): p. 530-536.

370. Gray, D.C., S. Mahrus, and J.A. Wells, *Activation of Specific Apoptotic Caspases with an Engineered Small-Molecule-Activated Protease.* Cell, 2010. **142**(4): p. 637-646.

371. Smart, A.D., et al., *Engineering a light-activated caspase-3 for precise ablation of neurons in vivo.* Proc Natl Acad Sci U S A, 2017. **114**(39): p. E8174-E8183.

372. Wang, J., et al., *Time-resolved protein activation by proximal decaging in living systems.* Nature, 2019. **569**(7757): p. 509-513.

373. Srinivasula, S.M., et al., *Generation of constitutively active recombinant caspases-3 and -6 by rearrangement of their subunits.* J Biol Chem 1998. **273**(17): p. 10107-11.

374. Wang, S., et al., *The Drosophila caspase inhibitor DIAP1 is essential for cell survival and is negatively regulated by HID.* Cell, 1999. **98**(4): p. 453-63.

375. Schipper, J.L., et al., *A bifunctional allosteric site in the dimer interface of procaspase-3.* Biophysical Chemistry, 2011. **159**(1): p. 100-109.

376. Hardy, J.A., et al., *Discovery of an allosteric site in the caspases.* Proc Natl Acad Sci U S A, 2004. **101**(34): p. 12461-6.

377. Feeney, B., et al., *Role of Loop Bundle Hydrogen Bonds in the Maturation and Activity of (Pro)caspase-3.* Biochemistry, 2006. **45**(44): p. 13249-63.

378. Ghosh, I., A.D. Hamilton, and L. Regan, *Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein.* J Am Chem Soc, 2000. **122**: p. 5658-9.

379. Ganesan, R., et al., *Extended Substrate Recognition in Caspase-3 Revealed by High Resolution X-ray Structure Analysis.* Journal of Molecular Biology, 2006. **359**(5): p. 1378-1388.

380. Brady, K.D., et al., *A catalytic mechanism for caspase-1 and for bimodal inhibition of caspase-1 by activated aspartic ketones.* Bioorganic & Medicinal Chemistry Letters, 1999. **7**(4): p. 621-31.

381. Stennicke, H.R. and G.S. Salvesen, *Catalytic properties of the caspases.* Cell Death Differ, 1999. **6**(11): p. 1054-9.

382. Chéreau, D., et al., *Structural and functional analysis of caspase active sites.* Biochemistry, 2003. **42**(14): p. 4151-60.

383. Sulpizi, M., et al., *Reaction mechanism of caspases: insights from QM/MM Car-Parrinello simulations.* Proteins, 2003. **52**(2): p. 212-24.

384. Colussi, P.A., et al., *Conversion of procaspase-3 to an autoactivating caspase by fusion to the caspase-2 prodomain.* J Biol Chem, 1998. **273**(41): p. 26566-70.

385. Mittl, P.R., et al., *Structure of recombinant human CPP32 in complex with the tetrapeptide acetyl-Asp-Val-Ala-Asp fluoromethyl ketone.* J. Biol. Chem. , 1997. **272**: p. 6539-47.

386. Powers, J.C., et al., *Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases.* Chemical reviews, 2002. **102**(12): p. 4639-750.

387. Wang, Z., et al., *Kinetic and structural characterization of caspase-3 and caspase-8 inhibition by a novel class of irreversible inhibitors.* Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 2010. **1804**(9): p. 1817-1831.

388. Forman, J., M. Dietrich, and W. Todd Monroe, *Photobiological and thermal effects of photoactivating UVA light doses on cell cultures.* Photochemical & Photobiological Sciences, 2007. **6**(6): p. 649.

389. Fire, A., *Fire Lab C. Elegans Vector Kit 1995*. 1995, Stanford Universit School of Medicine.

390. Gallegos, J.E. and A.B. Rose, *The enduring mystery of intron-mediated enhancement.* Plant Science, 2015. **237**: p. 8-15.

391. Crane, M.M., et al., *In vivo measurements reveal a single 5′-intron is sufficient to increase protein expression level in Caenorhabditis elegans.* Scientific Reports, 2019. **9**(1).

392. Donnelly, M.L.L., et al., *Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'.* Journal of General Virology, 2001. **82**(5): p. 1013-25.

393. Atkins, J.F., et al., *A case for "StopGo": Reprogramming translation to augment codon meaning of GGN by promoting unconventional termination (Stop) after addition of glycine and then allowing continued translation (Go).* RNA, 2007. **13**(6): p. 803-810.

394. Krause, M. and D. Hirsh, *A trans-spliced leader sequence on actin mRNA in C. elegans.* Cell, 1987. **49**(6): p. 753-61.

395. Spieth, J., et al., *Operons in C. elegans: Polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions.* Cell, 1993. **73**(3): p. 521-32.

396. Ryan, M.D., A.M. King, and G.P. Thomas, *Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence.* Journal of General Virology, 1991. **72**: p. 2727-32.

397. Ahier, A. and S. Jarriault, *Simultaneous expression of multiple proteins under a single promoter in Caenorhabditis elegans via a versatile 2A-based toolkit.* Genetics, 2014. **196**(3): p. 605-13.

398. Chang, J.S., et al., *Oxygen sensation and social feeding mediated by a C. elegans guanylate cyclase homologue.* Nature, 2004. **430**(6997): p.

317-22.

399. Shapiro, S.S. and M.B. Wilk, *An Analysis of Variance Test for Normality (Complete Samples).* Biometrika, 1965. **52**(3/4): p. 591-611.

400. Watanabe, M., et al., *Regulation of intracellular dynamics of Smad4 by its leucine-rich nuclear export signal.* EMBO Rep., 2000. **1**(2): p. 176-82.

401. Watanabe, M., et al., *Regulation of intracellular dynamics of Smad4 by its leucine-rich nuclear export signal.* EMBO Rep. , 2000. **1**(2): p. 176-82.

402. Kasagi, N., et al., *Apoptotic cell death in human gastric carcinoma: analysis by terminal deoxynucleotidyl transferase-mediated dUTP-biotin nick end labeling.* Jpn J Cancer Res., 1994. **85**(9): p. 939-45.

403. Gavrieli, Y., Y. Sherman, and S.A. Ben-Sasson, *Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation.* J Cell Biol, 1992. **119**(3): p. 493-501.

404. Zhang, J., et al., *Visualization of caspase-3-like activity in cells using a genetically encoded fluorescent biosensor activated by protein cleavage.* Nat Commun, 2013. **4**: p. 2157.

405. To, T.L., et al., *Rationally designed fluorogenic protease reporter visualizes spatiotemporal dynamics of apoptosis in vivo.* Proc Natl Acad Sci U S A, 2015. **112**(11): p. 3338-43.

406. De Crescenzo, G., et al., *Real-Time Monitoring of the Interactions of Two-Stranded de Novo Designed Coiled-Coils: Effect of Chain Length on the Kinetic and Thermodynamic Constants of Binding.* Biochemistry, 2003. **42**(6): p. 1754-63.

407. To, T.L., et al., *Rational Design of a GFP-Based Fluorogenic Caspase Reporter for Imaging Apoptosis In Vivo.* Cell Chem Biol, 2016. **23**(7): p. 875-82.

408. Iwai, H., et al., *Highly efficient proteintrans-splicing by a naturally split DnaE intein fromNostoc punctiforme.* FEBS Letters, 2006. **580**(7): p. 1853-1858.

409. Zettler, J., V. Schütz, and H.D. Mootz, *The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the proteintrans-splicing reaction.*

FEBS Letters, 2009. **583**(5): p. 909-914.

410. Duggan, A., C. Ma, and M. Chalfie, *Regulation of touch receptor differentiation by the Caenorhabditis elegans mec-3 and unc-86 genes.* Development, 1998. **125**: p. 4107-19.

411. Savage, C., et al., *mec-7 is a beta-tubulin gene required for the production of 15-protofilament microtubules in Caenorhabditis elegans.* Genes Dev, 1989. **3**(6): p. 870-81.

412. Tavernarakis, N. and M. Driscoll, *Molecular modeling of mechanotransduction in the nematode Caenorhabditis elegans.* Annu Rev Physiol., 1997. **59**: p. 659-89.

413. Fukushige, T., et al., *MEC-12, an α-tubulin required for touch sensitivity in C. elegans.* J Cell Sci, 1999. **112**(3): p. 395-403.

414. Mitani, S., et al., *Combinatorial Control of Touch Receptor Neuron Expression in Caenorhabditis Elegans.* Development, 1993. **119**(3): p. 773-83.

415. Chalfie, M., et al., *The neural circuit for touch sensitivity in Caenorhabditis elegans.* J Neurosci, 1985. **5**(4): p. 956-64.

416. Evans, T.C., *Transformation and microinjection (April 6)*, in *WormBook*. 2006, The C. elegans Research Community: WormBook, doi/10.1895/wormbook.1.108.1, http://www.wormbook.org.

417. Lai, C.C., et al., *Sequence and Transmembrane Topology of MEC-4, an Ion Channel Subunit Required for Mechanotransduction in Caenorhabditis elegans.* J Cell Biol, 1996. **133**(5): p. 1071-81.

418. Hosono, R., *Sterilization and growth inhibition of Caenorhabditis elegans by 5-fluorodeoxyuridine.* Exp Gerontol, 1978. **13**(5): p. 369-74.

419. Okkema, P.G., et al., *Sequence Requirements for Myosin Gene Expression and Regulation in Caenorhabditis elegans.* Genetics, 1993. **135**(2): p. 385-404.

420. Lindenboim, L., et al., *The nuclear envelope: target and mediator of the apoptotic process.* Cell Death Discovery, 2020. **6**(1).

421. Chalfie, M.S., John, *Developmental genetics of the mechanosensory*

neurons of Caenorhabditis elegans. Dev Biol, 1981. **82**(2): p. 358-70.

422. Zhang, S., et al., *MEC-2 Is Recruited to the Putative Mechanosensory Complex in C. elegans Touch Receptor Neurons through Its Stomatin-like Domain.* Current Biology, 2004. **14**(21): p. 1888-1896.

423. Hart, A., *Behavior (July 3)*, in *WormBook*. 2006, The C. elegans Research Community: doi/10.1895/wormbook.1.87.1, http://www.wormbook.org.

424. Barnes, T.M. and S. Hekimi, *The Caenorhabditis elegans avermectin resistance and anesthetic response gene unc-9 encodes a member of a protein family implicated in electrical coupling of excitable cells.* J Neurochem, 1997. **69**(6): p. 2251-60.

425. Bouhours, M., et al., *A co-operative regulation of neuronal excitability by UNC-7 innexin and NCA/NALCN leak channel.* Mol Brain, 2011. **4**: p. 16.

426. Mathews, E.A., et al., *Differential expression and function of synaptotagmin 1 isoforms in Caenorhabditis elegans.* Mol Cell Neurosci, 2007. **34**(4): p. 642-52.

427. Wong, S.S.C., et al., *Split-intein mediated re-assembly of genetically encoded Ca2+ indicators.* Cell Calcium, 2012. **51**(1): p. 57-64.

428. Shah, N.H. and T.W. Muir, *Inteins: Nature's Gift to Protein Chemists.* Chem Sci, 2014. **5**(1): p. 446-461.

429. Topilina, N.I. and K.V. Mills, *Recent advances in in vivo applications of intein-mediated protein splicing.* Mobile DNA, 2014. **5**(1): p. 5-18.

430. Ramirez, M., et al., *Engineering split intein DnaE from Nostoc punctiforme for rapid protein purification.* Protein Engineering Design and Selection, 2012. **26**(3): p. 215-223.

431. Zhang, B., et al., *A Dual-Intein Autoprocessing Domain that Directs Synchronized Protein Co-Expression in Both Prokaryotes and Eukaryotes.* Scientific Reports, 2015. **5**(1).

432. Kozak, M., *Pushing the limits of the scanning mechanism for initiation of translation.* Gene, 2002. **299**(1-2): p. 1-34.

433. Andrews, S.J. and J.A. Rothnagel, *Emerging evidence for functional peptides encoded by short open reading frames.* Nature Reviews Genetics,

2014. **15**(3): p. 193-204.

434. Kozak, M., *Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes.* Mol Cell Biol, 1987. **7**(10): p. 3438-45.

435. Abastado, J.-P., et al., *Suppression of ribosomal reinitiation at upstream open reading frames in amino acid-starved cells forms the basis for GCN4 translational control.* Mol Cell Biol, 1991. **11**(1): p. 486-96.

436. Garcia-Barrio, M.T., et al., *GCD10, a translational repressor of GCN4, is the RNA-binding subunit of eukaryotic translation initiation factor-3.* Genes Dev, 1995. **9**(14): p. 1781-96.

437. Hinnebusch, A.G., *Translational Regulation of Yeast GCN4.* Journal of Biological Chemistry, 1997. **272**(35): p. 21661-21664.

438. Diba, F., C.S. Watson, and B. Gametchu, *5′ UTR sequences of the glucocorticoid receptor 1A transcript encode a peptide associated with translational regulation of the glucocorticoid receptor.* J Cell Biochem., 2001. **81**(1): p. 149-61.

439. Luke, G.A., et al., *Occurrence, function and evolutionary origins of '2A-like' sequences in virus genomes.* Journal of General Virology, 2008. **89**(4): p. 1036-1042.

440. Doronina, V.A., et al., *Site-Specific Release of Nascent Chains from Ribosomes at a Sense Codon.* Molecular and Cellular Biology, 2008. **28**(13): p. 4227-4239.

441. Rychlík, I., et al., *Substrate specificity of ribosomal peptidyl transferase. The effect of the nature of the amino acid side chain on the acceptor activity of 2'(3')-O-aminoacyladenosines.* Eur J Biochem., 1970. **16**(1): p. 136-42.

442. Cesaratto, F., et al., *BiP/GRP78 Mediates ERAD Targeting of Proteins Produced by Membrane-Bound Ribosomes Stalled at the STOP-Codon.* Journal of Molecular Biology, 2019. **431**(2): p. 123-141.

443. Compton, Jonathan L., Amy N. Hellman, and V. Venugopalan, *Hydrodynamic Determinants of Cell Necrosis and Molecular Delivery Produced by Pulsed Laser Microbeam Irradiation of Adherent Cells.* Biophysical Journal, 2013. **105**(9): p. 2221-2231.

444. Kimble, J., *Alterations in cell lineage following laser ablation of cells in the somatic gonad of Caenorhabditis elegans.* Developmental biology, 1981. **87**(2): p. 286-300.

445. White, J.Q., et al., *The sensory circuitry for sexual attraction in C. elegans males.* Curr Biol, 2007. **17**(21): p. 1847-57.

446. Orrenius, S., V. Gogvadze, and B. Zhivotovsky, *Mitochondrial Oxidative Stress: Implications for Cell Death.* Annual Review of Pharmacology and Toxicology, 2007. **47**(1): p. 143-183.

447. Reed, R. and T. Maniatis, *An extensive network of coupling among gene expression machines.* Nature, 2002. **416**(6880): p. 499-506.

448. Nott, A., H. Le Hir, and M.J. Moore, *Splicing enhances translation in mammalian cells: An additional function of the exon junction complex.* Genes & development, 2004. **18**(2): p. 210-22.

449. Dusenbery, D.B., R.E. Sheridan, and R.L. Russell, *Chemotaxis-defective mutants of the nematode Caenorhabditis elegans.* Genetics, 1975. **80**(2): p. 297-309.

450. Stiernagle, T., *Maintenance of C. elegans (February 11)*, in *WormBook*. 2006, The C. elegans Research Community: WormBook, doi/10.1895/wormbook.1.101.1, http://www.wormbook.org.

451. Etchberger, J.F. and O. Hobert, *Vector-free DNA constructs improve transgene expression in C. elegans.* Nat Methods, 2008. **5**(1): p. 3.

452. Ronda, C.R., T. Ju¨stel, and H. Nikol, *Rare earth phosphors: fundamentals and applications.* Journal of Alloys and Compounds 1998: p. 669–76.

453. Sheldon, M.R., M.J. Fillyaw, and W.D. Thompson, *The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs.* Physiother Res Int, 1996. **1**(4): p. 221-8.

454. Hobert, O., R.J. Johnston, Jr., and S. Chang, *Left-right asymmetry in the nervous system: the Caenorhabditis elegans model.* Nat Rev Neurosci, 2002. **3**(8): p. 629-40.

455. Hobert, O., *Development of left/right asymmetry in the Caenorhabditis elegans nervous system: From zygote to postmitotic neuron.* genesis, 2014.

52(6): p. 528-543.

456. Pinan-Lucarre, B., et al., *The Core Apoptotic Executioner Proteins CED-3 and CED-4 Promote Initiation of Neuronal Regeneration in Caenorhabditis elegans.* PLoS Biology, 2012. **10**(5): p. e1001331.

457. Ohsawa, S., et al., *Caspase-mediated changes in histone H1 in early apoptosis: prolonged caspase activation in developing olfactory sensory neurons.* Cell Death & Differentiation, 2008. **15**(9): p. 1429-1439.

458. Ohsawa, S., et al., *Maturation of the olfactory sensory neurons by Apaf-1/caspase-9-mediated caspase activity.* Proc Natl Acad Sci U S A, 2010. **107**(30).

459. Yuan, J., A. Najafov, and B.F. Py, *Roles of Caspases in Necrotic Cell Death.* Cell, 2016. **167**(7): p. 1693-1704.

460. Shi, J., et al., *Inflammatory caspases are innate immune receptors for intracellular LPS.* Nature, 2014. **514**(7521): p. 187-192.

461. Dahl, G., et al., *Cell/cell channel formation involves disulfide exchange.* European journal of biochemistry/FEBS, 1991. **197**: p. 141-144.

462. Dahl, G., et al., *Mutational analysis of gap junction formation.* Biophysical journal, 1992. **62**: p. 172-180.

463. Foote, C.I., et al., *The pattern of disulfide linkages in the extracellular loop regions of connexin 32 suggests a model for the docking interface of gap junctions.* J Cell Biol, 1998. **140**: p. 1187-1197.

464. Hall, D.H., *Gap junctions in C. elegans: Their roles in behavior and development.* Dev Neurobiol, 2016.

465. Hicke, L., *Protein regulation by monoubiquitin.* Nat Rev Mol Cell Biol, 2001. **2**(3): p. 195-201.

466. Haglund, K., et al., *Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation.* Nature Cell Biology, 2003. **5**(5): p. 461-466.

467. Hochstrasser, M., *Lingering Mysteries of Ubiquitin-Chain Assembly.* Cell, 2006. **124**(1): p. 27-34.

468. Portbury, A.L., et al., *Back to your heart: Ubiquitin proteasome system-regulated signal transduction.* Journal of Molecular and Cellular

Cardiology, 2012. **52**(3): p. 526-537.

469. Papaevgeniou, N. and N. Chondrogianni, *The ubiquitin proteasome system in Caenorhabditis elegans and its regulation.* Redox Biology, 2014. **2**: p. 333-347.

470. Reyes-Turcu, F.E., K.H. Ventii, and K.D. Wilkinson, *Regulation and Cellular Roles of Ubiquitin-Specific Deubiquitinating Enzymes.* Annual Review of Biochemistry, 2009. **78**(1): p. 363-397.

471. Reyes-Turcu, F.E. and K.D. Wilkinson, *Polyubiquitin binding and disassembly by deubiquitinating enzymes.* Chem Rev, 2009. **109**(4): p. 1495-508.

472. Passmore, L. and D. Barford, *Getting into position: the catalytic mechanisms of protein ubiquitylation.* Biochem J., 2004. **379**: p. 513-25.

473. Mullins, M., M. Kulkarni, and H.E. Smith, *E1 Ubiquitin-Activating Enzyme UBA-1 Plays Multiple Roles throughout C. elegans Development.* PLoS Genetics, 2008. **4**(7): p. e1000131.

474. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.

475. Frøkjær-Jensen, C., et al., *Random and targeted transgene insertion in Caenorhabditis elegans using a modified Mos1 transposon.* Nature Methods, 2014. **11**(5): p. 529-534.

476. Dickinson, D.J., et al., *Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette.* Genetics, 2015. **200**(4): p. 1035-1049.

477. Paix, A., et al., *High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes.* Genetics, 2015. **201**(1): p. 47-54.

478. Dickinson, D.J. and B. Goldstein, *CRISPR-Based Methods for Caenorhabditis elegans Genome Engineering.* Genetics, 2016. **202**(3): p. 885-901.

479. Schlake, T. and J. Bode, *Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci.* Biochemistry 1994. **33**: p. 12746-51.

480. Turan, S., et al., *Recombinase-mediated cassette exchange (RMCE) — A rapidly-expanding toolbox for targeted genomic modifications.* Gene, 2013. **515**(1): p. 1-27.

481. Nonet, M.L., *Efficient Transgenesis in Caenorhabditis elegans Using Flp Recombinase-Mediated Cassette Exchange.* Genetics, 2020. **215**(4): p. 903-921.

482. Gossen, M. and H. Bujard, *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.* Proc Natl Acad Sci U S A, 1992: p. 5547-51

483. Brand, A.H. and N. Perrimon, *Targeted gene expression as a means of altering cell fates and generating dominant phenotypes.* Development, 1993. **118**: p. 401-15.

484. Wang, H., et al., *cGAL, a temperature-robust GAL4–UAS system for Caenorhabditis elegans.* Nature Methods, 2017. **14**(2): p. 145-148.

485. Mao, S., et al., *A Tet/Q Hybrid System for Robust and Versatile Control of Transgene Expression in C. elegans.* iScience, 2019. **11**: p. 224-237.

486. Fashena, S.J., I.G. Serebriiskii, and E.A. Golemis, *LexA- based two-hybrid systems.* Methods Enzymol, 2000. **328**: p. 14-26.

487. Potter, C.J., et al., *The Q System: A Repressible Binary System for Transgene Expression, Lineage Tracing, and Mosaic Analysis.* Cell, 2010. **141**(3): p. 536-548.

488. Wei, X., et al., *Controlling gene expression with the Q repressible binary expression system in Caenorhabditis elegans.* Nature Methods, 2012. **9**(4): p. 391-395.

489. Monsalve, G.C., K.R. Yamamoto, and J.D. Ward, *A New Tool for Inducible Gene Expression in Caenorhabditis elegans.* Genetics, 2019. **211**(2): p. 419-430.

490. Aljohani, M.D., et al., *Engineering rules that minimize germline silencing of transgenes in simple extrachromosomal arrays in C. elegans.* Nature Communications, 2020. **11**(1).

491. Fire, A., S.W. Harrison, and D. Dixon, *A modular set of lacZ fusion vectors for*

studying gene expression in Caenorhabditis elegans. Gene, 1990. **93**(2): p. 189-98.

492. Knight, S.W. and B.L. Bass, *The role of RNA editing by ADARs in RNAi.* Mol Cell, 2002. **10**(4): p. 809-17.

493. Bagijn, M.P., et al., *Function, targets, and evolution of Caenorhabditis elegans piRNAs.* Science, 2012. **337**(6094): p. 574-78.

494. Wu, W.-S., et al., *pirScan: a webserver to predict piRNA targeting sites and to avoid transgene silencing in C. elegans.* Nucleic Acids Research, 2018. **46**(W1): p. W43-W48.

495. Fire, A., R. Alcazar, and F. Tan, *Unusual DNA Structures Associated With Germline Genetic Activity in Caenorhabditis elegans.* Genetics, 2006. **173**(3): p. 1259-1273.

496. Frøkjær-Jensen, C., et al., *An Abundant Class of Non-coding DNA Can Prevent Stochastic Gene Silencing in the C. elegans Germline.* Cell, 2016. **166**(2): p. 343-357.

497. Rog, O., S. Köhler, and A.F. Dernburg, *The synaptonemal complex has liquid crystalline properties and spatially regulates meiotic recombination factors.* eLife, 2017. **6**.

498. Fielmich, L.-E., et al., *Optogenetic dissection of mitotic spindle positioning in vivo.* eLife, 2018. **7**.

499. Zhang, D., et al., *The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes.* Science, 2018. **359**(6375): p. 587-92.

500. Artiles, K.L., A.Z. Fire, and C. Frøkjær-Jensen, *Assessment and Maintenance of Unigametic Germline Inheritance for C. elegans.* Developmental Cell, 2019. **48**(6): p. 827-839.e9.

501. Reinke, V. and A.D. Cutter, *Germline Expression Influences Operon Organization in the Caenorhabditis elegans Genome.* Genetics, 2009. **181**(4): p. 1219-1228.