TECHNISCHE UNIVERSITÄT MÜNCHEN

Department Chemie

Labor für Synthetische Biochemie

Expanding the chemical biology toolbox: Site-specific incorporation of unnatural amino acids and bioorthogonal protein labeling to study structure and function of proteins

Susanne Veronika Mayer

Vollständiger Abdruck der, von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzende/-r: Prof. Dr. Matthias Feige

Prüfende/-r der Dissertation:
1. Prof. Dr. Kathrin Lang
2. Dr. Stephan Hacker, Junior Group Leader

Die Dissertation wurde am 16.07.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 30.07.2019 angenommen.

ABSTRACT

The fundamental understanding of biological processes at a molecular level constitutes the basis for the development of new drugs and therapies against diseases. Endowing biomolecules such as proteins with physical probes and chemical reporters provides an opportunity to elucidate these molecular mechanisms, by providing information on structure, dynamics, function and localization of labeled proteins. An especially exciting area consists in the development of new tools for the introduction of chemical reporters into proteins in their native environment - the living cell.

Combining biological and chemical methods has proven to be a powerful approach in developing such new tools. One example is genetic code expansion (GCE), which represents an elegant and potent approach to site-specifically and co-translationally install reporters or probes into a protein of interest. This is achieved by using orthogonal aminoacyl-tRNA synthetase (aaRS)/tRNA pairs that direct the incorporation of an unnatural amino acid (uAA) into a protein in response to an amber stop codon introduced in a gene of interest. An emerging area with enormous potential consists in the site-specific incorporation of uAAs that contain functional groups that allow subsequent labeling with biophysical probes at a single site in a protein of interest. These uAAs, due to their small size, only minimally perturb structure and function of the target protein and decorate the protein with a bioorthogonal handle for further chemical or biochemical modification. Upon exogenous addition of a reaction partner the so installed handle can be decorated chemoselectively in a modular manner. The Diels-Alder cycloaddition with inverse-electron demand (iEDDAC) between electron-poor tetrazines (dienes) and electron-rich strained alkenes or alkynes (dienophiles) has emerged as a very powerful representative of bioorthogonal labeling reactions. It occurs very rapidly under physiological conditions and generates stable, covalent linkages with the formation of gaseous nitrogen as the only, nontoxic byproduct.

In this thesis, GCE in combination with iEDDAC was used as a basis for the development of new tools to study protein structure and dynamics and to induce site-specific protein labeling in living cells. We contribute to the broad spectrum of bioorthogonal protein modification reactions by designing a new, very rapid photo-inducible iEDDAC labeling reaction with spatio-temporal control. Hereby, a photo-labile cyclopropenone-caged dibenzoannulated bicyclononyne probe (photo-DMBO) specifically and chemoselectively reacts only upon decarbonylation, triggered at 365 nm, with novel tetrazine-based uAAs TetK and mTetK. Incorporation of these uAAs into proteins with a novel orthogonal aaRS/tRNA pair, allowed us to determine very fast on-protein rate-constants of ~50 $M^{-1}s^{-1}$ between photo-DMBO and mTetK-bearing sfGFP, which is significantly faster than other photo-induced labeling reactions. Protein labeling with fluorophore conjugates of photo-DMBO at low µM concentrations was performed on purified tetrazine-bearing proteins, within the *E. coli* proteome or on the surface of living *E. coli* in a photo-induced manner.

Furthermore, a new approach based on GCE and iEDDAC using a small, rigid tetrazine amino acid (TetF) to site-specifically introduce nitroxide spin labels bearing cyclopropene (Cp) moieties into proteins was developed during this thesis. In collaboration with the group of Michael Sattler (Department of Chemistry, TUM), we characterized the resulting spin-labeled protein by pulsed electron electron double resonance (PELDOR) and paramagnetic resonance enhancement (PRE) experiments, as well as molecular dynamics (MD) simulations to benchmark our site-specific spin labeling approach against existing cysteine-based site-directed spin labeling (SDSL) techniques. By introducing one or two spin labels into the double-stranded RNA (dsRNA)-binding protein Loquacious-PD (Loqs-PD), which recognizes and processes endogenous siRNA in *D. melanogaster*, we were able to determine distinct intra-protein distances by PELDOR upon binding to double-stranded RNA, indicating that this binding event induces an ordered orientation of the two otherwise independently acting dsRNA-binding domains of Loqs-PD.

In collaboration with Henning Mootz (Institute for Biochemistry, Universität Münster), we have developed an enzyme-mediated approach for the deprotection of a masked natural amino acid to control and modulate enzymatic activity. The bulky phenylacetyl (Pac) group was chosen as protecting group for lysine (PacK) to locally perturb the structure of a target protein and thereby abolish its activity. Together with the Mootz group we have discovered two different enzymes, penicillin G acylase (PGA) and a sirtuin (SrtN) from *Bacillus subtilis* that are both able to facilitate the cleavage of the Pac group from model proteins as well as more complex multi-domain proteins, where PacK was installed via GCE. This is the first report of an artificial enzyme-triggered decaging approach. The mild conditions of our enzyme-mediated approach render it a valuable alternative to chemical or optical triggers for the modulation of protein function.

In conclusion, we have added several novel and powerful approaches, based on a combination of GCE with bioorthogonal labeling and decaging reactions, to the chemical biology toolkit, which in the future will contribute to the elucidation and modulation of biological processes in *in vivo* cellular settings.

ZUSAMMENFASSUNG

Das grundlegende Verständnis von biologischen Prozessen auf molekularer Ebene bildet die Basis für die Entwicklung neuer Wirkstoffe und Therapien zur Bekämpfung von Krankheiten. Die Ausstattung von Biomolekülen wie Proteinen mit biophysikalischen und chemischen Sonden liefert Informationen über Struktur, Dynamik, Funktion und Ort des markierten Biomoleküls. Dies ermöglicht die Aufklärung dieser molekularen Mechanismen. Ein besonders interessantes Gebiet stellt dabei die Entwicklung neuer Methoden für die Einführung von chemischen Sonden in Proteine in ihrer natürlichen Umgebung - der lebenden Zelle - dar.

Die Kombination biologischer und chemischer Methoden hat sich als wirkungsvolle Vorgehensweise für die Entwicklung solch neuer Methoden erwiesen. Ein Beispiel dafür ist die Erweiterung des genetischen Codes (GCE), die einen eleganten und effektiven Ansatz darstellt, um Sonden während der Proteinbiosynthese positionsspezifisch in ein Zielprotein einzubringen. Orthogonale Aminoacyl-RNA Synthetase (aaRS)/tRNA Paare dirigieren dabei den Einbau von unnatürlichen Aminosäuren (uAA) als Antwort auf ein genetisch installiertes Ambercodon in das Zielprotein. Der positionsspezifische Einbau von uAAs, die Funktionalitäten tragen, welche daraufhin an einer bestimmten Stelle im Protein mit biophysikalischen Sonden versehen werden können, beinhaltet dabei großes Potential. Wegen ihrer geringen Größe stören diese uAAs die Struktur und Funktion des Zielproteins nur geringfügig und statten dieses mit einer bioorthogonalen Gruppe aus, die durch Zugabe eines entsprechenden Reaktionspartners biochemisch oder chemoselektiv über das Baukastenprinzip weiter funktionalisiert werden kann. Die Diels-Alder Cycloaddition mit inversem Elektronenbedarf (iEDDAC) zwischen elektronenarmen Tetrazinen (Diene) und elektronenreichen, ringgespannten Alkenen oder Alkinen (Dienophile) hat sich dabei als äußerst wirkungsvoller Vertreter dieser Reaktionen zur bioorthogonalen Markierung herausgestellt. Sie verläuft sehr schnell unter physiologischen Bedingungen und erzeugt stabile, kovalente Verbindungen und als einziges, ungiftiges Nebenprodukt wird gasförmiger Stickstoff freigesetzt.

In dieser Arbeit wurden auf der Basis von GCE und iEDDAC neue Werkzeuge für die Untersuchung von Proteinstruktur und -dynamik, sowie zur Aktivierung positionsspezifischer Markierung von Proteinen in lebenden Zellen entwickelt. Durch Konstruktion einer neuen, sehr schnellen, lichtinduzierbaren iEDDAC Markierungsreaktion mit räumlicher und zeitlicher Kontrolle haben wir zum breiten Spektrum bioorthogonaler Reaktionen für die Dekoration von Proteinen beigetragen. Dabei reagiert eine photolabile Cyclopropenon-geschützte, dibenzoannulierte Bicyclononin-Sonde (photo-DMBO) erst nach Decarbonylierung bei 365 nm spezifisch und chemoselektiv mit den neuen, auf Tetrazinen basierenden uAAs TetK und mTetK. Der Einbau dieser uAAs in Proteine mit einem neuen orthogonalen aaRS/tRNA Paar ermöglichte uns die Bestimmung einer sehr schnellen Geschwindigkeitskonstante am Protein von ~50 $M^{-1}s^{-1}$ zwischen photo-DMBO und mTetK-modifiziertem sfGFP, was deutlich schneller ist als andere lichtinduzierbare Modifikationsreaktionen. Die lichtaktivierte Markierung von Proteinen mit einem photo-

DMBO Fluorophor-Konjugat wurde bei niedrigen µM Konzentrationen an aufgereinigten Proteinen, sowie an Proteinen im *E. coli* Proteom und an der Oberfläche von lebenden *E. coli* erfolgreich durchgeführt.

Weiterhin wurde ein neuer Ansatz basierend auf der GCE und iEDDAC entwickelt, der die kleine, starre Tetrazin-Aminosäure (TetF) nutzt, um Nitroxid-Spinmarker, die mit einer Cyclopropengruppe versehen sind, positionsspezifisch in Proteine einzubringen. In Zusammenarbeit mit der Arbeitsgruppe von Michael Sattler (Department für Chemie, TUM) wurde die resultierende Spinmarkierung über gepulste Elektron-Elektron Doppelresonanz (PELDOR) und paramagnetische Relaxationsverstärkungsexperimente (paramagentic relaxation enhancement, PRE), sowie Simulation der Molekulardynamik (MD) charakterisiert, um unseren positionsspezifischen Spinmarkierungsansatz (site-directed spin labeling SDSL) mit bestehenden Cystein-basierten Methoden zu vergleichen. Durch die Einführung von einem oder zwei Spinmarkern in Loquacious-PD (Loqs-PD), ein Protein, das doppelsträngige RNA (dsRNA) bindet und dessen Rolle die Erkennung und Prozessierung von endogener siRNA in *D. melanogaster* ist, konnten wir genaue Distanzen innerhalb des Proteins nach Bindung doppelsträngiger RNA mittels PELDOR bestimmen. Diese weisen darauf hin, dass dieses Bindungsevent eine geordnete Orientierung der beiden sonst unabhängig agierenden Domänen, welche die dsRNA binden, hervorruft.

In Zusammenarbeit mit der Arbeitsgruppe von Henning Mootz (Institut für Biochemie, Universität Münster) haben wir einen Ansatz für die Enzym-vermittelte Entschützung von maskierten natürlichen Aminosäuren zur Kontrolle und Regulierung enzymatischer Aktivität entwickelt. Die sperrige Phenylacetylgruppe (Pac) wurde als Schutzgruppe für Lysine (PacK) gewählt, um lokal die Struktur des Zielproteins zu stören und damit gezielt dessen Aktivität zu beeinträchtigen. Zusammen mit der Mootz Gruppe haben wir zwei unterschiedliche Enzyme, Penicillin G acylase (PGA) und ein Sirtuin (SrtN) aus *Bacillus subtilis*, gefunden, die beide in der Lage sind, die Spaltung der Pac Gruppe von Modellproteinen, sowie komplexeren Multi-Domänen-Proteinen, in welchen PacK positionsspezifisch über GCE installiert wurde, zu bewerkstelligen. Dies ist der erste Bericht eines künstlichen, Enzym-veranlassten Ansatzes zur Demaskierung. Die milden Konditionen unseres Enzym-vermittelten Vorgehens machen den Ansatz zu einer wertvollen Alternative gegenüber chemischen oder optischen Stimuli zur Regulation von Proteinfunktionen.

Zusammenfassend haben wir mehrere neue und wirkungsvolle Werkzeuge zur Markierung oder Demaskierung von Biomolekülen zum Repertoire der Chemischen Biologie hinzugefügt, die auf einer Kombination von GCE und bioorthogonalen Reaktionen basieren. In der Zukunft werden diese dazu beitragen, biologische Prozesse in lebendem, zellulärem Milieu aufzuklären und zu regulieren.

ACKNOWLEDGEMENTS - DANKSAGUNG

An dieser Stelle möchte ich mich ganz herzlich bei allen Leuten bedanken, die mich in den letzten viereinhalb Jahren bei der Durchführung meiner Dissertation unterstützt und begleitet haben - ohne euch wäre das alles nicht möglich gewesen!

Prof. Kathrin Lang danke ich für die Möglichkeit, meine Dissertation als Teil ihrer Gruppe durchzuführen und auch aktiv am Aufbau des Labors und der Gruppe mitzuarbeiten. Als ich auf die Stellenanzeige aufmerksam geworden bin, wusste ich sofort, dass dieses Thema an der Schnittstelle zwischen organischer Chemie und Biochemie genau das Richtige für mich ist. Vielen Dank dafür, dass du immer ein offenes Ohr für mich hattest. Ich werde die guten Gespräche über wissenschaftliche Probleme, aber auch private Themen vermissen. Außerdem habe ich mich dank dir und den Tetrazinen mit der Farbe Pink versöhnt.

Ein großes Dankeschön gilt meiner Arbeitsgruppe, der "KlangBang", bestehend aus Marko Cigler, Marie-Kristin von Wrisberg, Tuan-Anh Nguyen, Max Fottner, Kristina Krauskopf, Marie-Lena Jokisch, Toni Murnauer und Vera Wanka, sowie den ehemaligen Mitgliedern Mathias Enderle, Thorsten Müller, Dominik Essig und Tanja Reichhart. Ich werde euch und unsere gemeinsamen Aktivitäten in Zukunft sehr vermissen. Dank der guten Atmosphäre im Labor und der gegenseitigen Unterstützung gab es keinen einzigen Tag in den letzten viereinhalb Jahren, an dem ich nicht gerne in die Arbeit gekommen wäre und das obwohl experimentell nicht immer alles glatt gelaufen ist. Besonders bedanken möchte ich mich bei Marko, meinem Mitstreiter erster Stunde, für die wertvollen Diskussionen über organische Chemie als es nur uns zwei Chemiker gab; bei Marie für die gemeinsamen Mittagspausen; dafür, dass du hast mir Klonieren beigebracht hast und du immer da bist, wenn man dich braucht; bei Kristina, für die positiven Aufmunterungen in stressigen Situation; dafür, dass ich mich immer bei dir melden kann und ich freue mich in dir einen neuen Konzertpartner gefunden zu haben und bei Toni, für deine offene Art und besonders als Co-Partner des photo-iEDDAC Projektes, das ohne dich jetzt sicher nicht so dastehen würde. Meinem Kooperationspartner im Rahmen des SDSL Projektes Stefan Gaussmann danke ich für das geduldige, wiederholte Erklären von verschiedenen NMR und EPR Grundlagen für unser Projekt. Ein Dankeschön gilt auch all meinen Forschungsstudenten, unter anderem Patrick Allihn, Julia Freivogel und Maria Mühlhofer, die mir geholfen haben die Projekte voranzubringen. Marko, Marie, Kristina, Toni, Stefan, Marie-Lena und Vera - vielen Dank auch für das kurzfristige Korrekturlesen von Teilen dieser Arbeit, sie würde sonst sicher (mehr) Fehler enthalten!

Ich möchte mich auch bei unseren Kollaborationspartnern für den wissenschaftlichen Austausch und die gute Zusammenarbeit bedanken. Innerhalb des SDSL Projektes waren das Stefan Gaussmann, Prof. Michael Sattler (TUM), Dr. Ana Gamiz-Hernandez, Prof. Ville Kaila (TUM), Dr. Burkhard Endeward und Prof. Thomas Prisner (Goethe Universität Frankfurt), und innerhalb des PacK Projektes Dr. Marie Reille-Seroussi und Prof. Henning Mootz (Universität Münster).

Abschließend möchte ich mich bei meinen Eltern, meinem Bruder und seiner Familie, sowie unserer Großfamilie bedanken, ihr seid immer für mich da und gebt mir das Vertrauen und die Kraft auch in schwierigen Situationen gut durchs Leben zu kommen.

LIST OF PUBLICATIONS

During this thesis parts of this work have been published in the following papers, reviews and book chapters or presented as posters:

Journal Publications (peer-reviewed)

M. Reille-Seroussi, **S.V. Mayer**, W. Dörner, K. Lang and H.D.D. Mootz. (2019). Expanding the Genetic Code with a Lysine Derivative Bearing an Enzymatically Removable Phenylacetyl Group *ChemComm*, 55, 4793-4796. DOI: 10.1039/C9CC00475K

S. Mayer, K. Lang. (2017). Tetrazines in Inverse-Electron-Demand Diels–Alder Cycloadditions and Their Use in Biology. *Synthesis*, 49, 830-848. DOI: 10.1055/s-0036-1588682.

Submitted Journal Publications (peer-reviewed)

S.V. Mayer,[#] A. Murnauer,[#] M.-K. von Wrisberg, M.-L. Jokisch and K. Lang. Photo-induced and rapid labeling of tetrazine-bearing proteins via cyclopropenone-caged bicyclononynes. **2019**. Manuscript submitted for publication.

[#] these authors contributed equally

Book chapters

S. Mayer, K. Lang (2017). Ein erweiterter genetischer Code - neue Proteine aus dem Labor. In T. K. Lindhorst, H.-J. Quadbeck-Seeger & GDCh (Eds.), *Unendliche Weiten: Kreuz und quer durchs Chemie-Universum* (pp. 52-56). Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.

Reports

S. Mayer and K. Lang (2016). Chemie in lebenden Systemen. Nachr. Chem., 64, 301-305. DOI: 10.1002/nadc.20164047516.

Poster presentations (peer-reviewed)

S. V. Mayer, A. Murnauer, M.-K. von Wrisberg, M.-L. Jokisch, K. Lang. (2019) Caged bicyclononynes allow photo-induced labeling of tetrazine-bearing proteins. 8th CPS Meeting, 16. -19.06.2019, Berlin, Germany. Recognized with a poster award from Nature Reviews.

S. V. Mayer, S. Gaussmann, M. Sattler, K. Lang (2018). Site-directed spin labeling via tetrazine ligation to probe protein structure. GCE conference 2018, 09. -12.08.2018, Corvallis, Oregon, USA.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviations of natural amino acids

A, Ala	alanine
C, Cys	cysteine
D, Asp	aspartic acid
E, Glu	glutamic acid
F, Phe	phenylalanine
G, Gly	glycine
H, His	histidine
I, Ile	isoleucine
K, Lys	lysine
L, Leu	leucine
M, Met	methionine
N, Asn	asparagine
P, Pro	proline
R, Arg	arginine
Q, Gln	glutamine
S, Ser	serine
T, Thr	threonine
V, Val	valine
W, Trp	tryptophane
Y, Tyr	tyrosine

General abbreviations

aaRS	aminoacyl-tRNA synthetase
AgOAc	silver acetate
AcOH	acetic acid
AR	antibacterial resistance
BARAC	biarylazacyclooctynone
bb	backbone
BCN	bicyclo[6.1.0]nonyne
BCN-OH	bicyclo[6.1.0]non-4-yn-9-ylmethanol
BF_3*Et_2O	boron trifluoride diethyl etherate
Boc	*tert*-butoxycarbonyl
Boc_2O	di-*tert*-butyl dicarbonate
Cam	chloramphenicol
CAT	chloramphenicol acetyltransferase
$CDCl_3$	deuterated chloroform
Cp	cyclopropene
CpK	N^6-(((2-methylcycloprop-2-en-1-yl)methoxy)carbonyl)-L-lysine
Cp-PROXYL	(*N*-((2-methylcycloprop-2-en-1-yl)methyl))-2,2,5,5-tetramethyl-1-pyrrolidinyloxyl-3-carboxamide
Cp-TEMPO	(2-methylcycloprop-2-en-1-yl)methyl (*N*-(2,2,6,6-tetramethylpiperidin-

	1-oxyl))carbamate
CpSLA	*N*-((2-methylcycloprop-2-en-1-yl)methyl)cyclopentanecarboxamide (spin label analog)
CSP	chemical shift perturbation
DA	Diels-Alder cycloaddition
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DIBAC or ADIBO	aza-dibenzocyclooctyne
DIBO	dibenzocyclooctyne
DIBOD	dibenzo[*a*,*e*]cyclooctadiyne
DIPEA	diisopropylethylamine
DMAP	dimethylaminopyridine
DMBO	di(methoxybenzo)bicyclo[6.1.0]nonyne
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DMSO-d_6	deuterated dimethylsulfoxide
DNA	desoxyribonucleic acid
DPEPhos	bis[(2-diphenylphosphino)phenyl] ether
DPPA	diphenylphosphoryl azide
E. coli	*Escherichia coli*
EDC*HCl	*N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide hydrochloride
EPR	electron paramagnetic resonance
et al.	and others
eq	equivalent
ESI	electrospray ionization
Et_2O	diethylether
EtOH	ethanol
EtOAc	ethyl acetate
FA	formic acid
Fmoc	fluorenylmethyloxycarbonyl
FRET	Förster resonance energy transfer
GCE	genetic code expansion
H_2O	water
HATU	1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HCl	hydrochloric acid
HFIP	hexafluoroisopropanol
HiFi	NEB HiFi DNA Assembly
HmTetK	N^6-(((3-(1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine
HOAt	1-hydroxy-7-azabenzotriazol
HOBt	1-hydroxybenzotriazol
HPLC	high performance liquid chromatography
Hz	Hertz
iEDDAC	inverse-electron demand Diels-Alder cycloaddition
Kan	kanamycin

KI	potassium iodide
KmTetK	N^2-(L-lysyl)-N^6-(((3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)-carbonyl)-L-lysine
KpTetK	N^2-(L-lysyl)-N^6-(((4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)-carbonyl)-L-lysine
L	liter
LB	lysogeny broth
LC-MS	liquid chromatography mass spectrometry
LiOH	lithium hydroxide
$LiPyrrBH_3$	lithium pyrrolidinoborohydride
M	molar
M. alvus (Ma)	*Methanomethylophilus alvus*
M. barkeri (Mb)	*Methanosarcina barkeri*
MC-DIBOD	mono-cyclopropenone-caged dibenzocyclooctadiynes
Me	methyl
MeOH	methanol
min	minutes
miRNA	micro RNA
MHz	megaHertz
M. jannaschii (Mj)	*Methanocaldococcus jannaschii*
mL	milliliter
mM	millimolar
mmol	millimol
M. mazei (Mm)	*Methanosarcina mazei*
MS	mass spectrometry
MQ	Milli-Q
Myo	myoglobin
mTetK	N^6-(((3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine
mTet2K	N^6-(2-(3-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)acetyl)-L-lysine
$NaHCO_3$	sodium hydrogencarbonate
NaOH	sodium hydroxide
$NaNO_2$	sodium nitrate
NEt_3	triethylamine
$Ni(acac)_2$	nickel(II) acetylacetonate
$Ni(OTf)_2$	nickel(II) trifluoromethanesulfonate
NEM	*N*-ethylmorpholine
nm	nanometer
nM	nanomolar
nmol	nanomol
NMR	nuclear magnetic resonance
Nor	norbornene
Nor-OH	5-norbornene-2-ol
ODIBO	Oxa-dibenzocyclooctyne
OtBu	*tert*-butyl
$Pd_2(dba)_3$	tris(dibenzylideneacetone)dipalladium(0)

$Pd(OAc)_4$	palladium(IV) tetraacetate
PELDOR	pulsed electron electron double resonance
photo-DIBO	cyclopropenone-masked dibenzocyclooctyne
photo-DIBOD	bis-cyclopropenone-masked dibenzo[*a*,*e*]cyclooctadiyne
photo-DMBO	cyclopropenone-caged di(methoxybenzo)bicyclo[6.1.0]nonyne
photo-ODIBO	cyclopropenone-masked oxa-dibenzocyclooctyne
ppm	parts per million
PRE	paramagnetic resonance enhancement
PROXYL	2,2,5,5-tetramethyl-1-pyrrolidinyloxy
pTetK	N^6-(((4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-*L*-lysine
Pyl	pyrrolysine
PylRS	pyrrolysl tRNA synthetase
QC	quikchange mutagenesis
RC	restriction cloning
RISC	RNA induced silencing complex
RNA	ribonucleic acid
rpm	revolutions per minute
rt	room temperature
SCO	strained cyclooctyne
SDSL	site-directed spin labeling
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
sfGFP	superfolder Green Fluorescent Protein
siRNA	short interfering RNA
SL	spin label
Sm	spectinomycin
SOB	super optimal broth
SOC	super optimal broth with catabolite repression
SPhos	2-dicyclohexylphosphino-2′,6′-dimethoxybiphenyl
TAMRA	tetramethylrhodamine
TBAF	tetrabutylammonium fluoride
TCO	*trans*-cyclooctene
TEMPO	2,2,6,6-tetramethylpiperidin-1-oxyl
Tcn	tetracycline
Tet	1,2,4,5-tetrazine
TetA	*ß*-(6-methyl-1,2,4,5-tetrazin-3-yl)-L-alanine
TetF	*ß*-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alanine
TetK	N^6-((2-(6-methyl-1,2,4,5-tetrazin-3-yl)ethoxy)-carbonyl)-L-lysine
Tet-PROXYL	(*N*-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)-2,2,5,5-tetramethyl-1-pyrrolidinyloxyl–3-carboxamide
Tet-TEMPO	*N*-(2,2,6,6-tetramethylpiperidin-1-oxyl)-4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzamide
TFA	trifluoroacetic acid
THF	tetrahydrofuran
tRNA	transfer RNA

TyrRS (Mj)	tyrosyl tRNA synthetase from *M. jannaschii*
uAA	unnatural amino acid
Ub	ubiquitin
µM	micromolar
µmol	micromol
UV	ultraviolet
WB	western blot
$Zn(OTf)_2$	zinc(II) trifluoromethanesulfonate
$ZnCl_2$	zinc(II) chloride

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

I. General Introduction

The information in the following chapter I.4 has been published as:

S. Mayer and K. Lang. Tetrazines in Inverse-Electron-Demand Diels–Alder Cycloadditions and Their Use in Biology. *Synthesis* **2017**, *49*, 830-848.

I. General Introduction

I.1 Decoration of proteins

Proteins fulfill a plethora of tasks in cellular biological processes. They build up the cytoskeleton and are integrated in a great number of processes from cell signaling and transport of molecules to replication of DNA, protein biosynthesis and cell death. Therefore, it is of great interest to have tools to track and visualize their location, movement and pathways in the cellular context. First approaches to attach biophysical probes like fluorophores in a covalent manner to the proteins of interest (POIs) were based on lysine- and cysteine-reactive partners (for example NHS-esters, maleiimides or iodoacetamides) and therefore limited to *in vitro* applications.[1] Immunolabeling techniques, using a primary antibody to generate selectivity for a certain protein and a secondary antibody-fluorophore conjugate and can only be applied on permeabilized cells or extracellular proteins.[2] The discovery of fluorescent proteins (FP) and their application as genetically encoded protein tags in living organisms greatly influenced imaging of biological processes.[2-3] Their first representative is green fluorescent protein (GFP), discovered in 1962 in the marine organism *Aequorea victoria*, which was found to display fluorescence (excitation: 395 and 475 nm; emission: 508 nm) also when expressed in different organisms.[3b] Formation of the GFP chromophore (Figure I.1b) in the central α-helix at the heart of the 11-β-sheet barrel structure of the 240 amino acid long protein happens posttranslationally and host-independently by autocatalytic maturation of the tripeptide Ser65-Tyr66-Gly67 due to internal cyclization and oxidation to generate 4-(*p*-hydroxybenzylidene)-imidazolidin-5-one.[4]

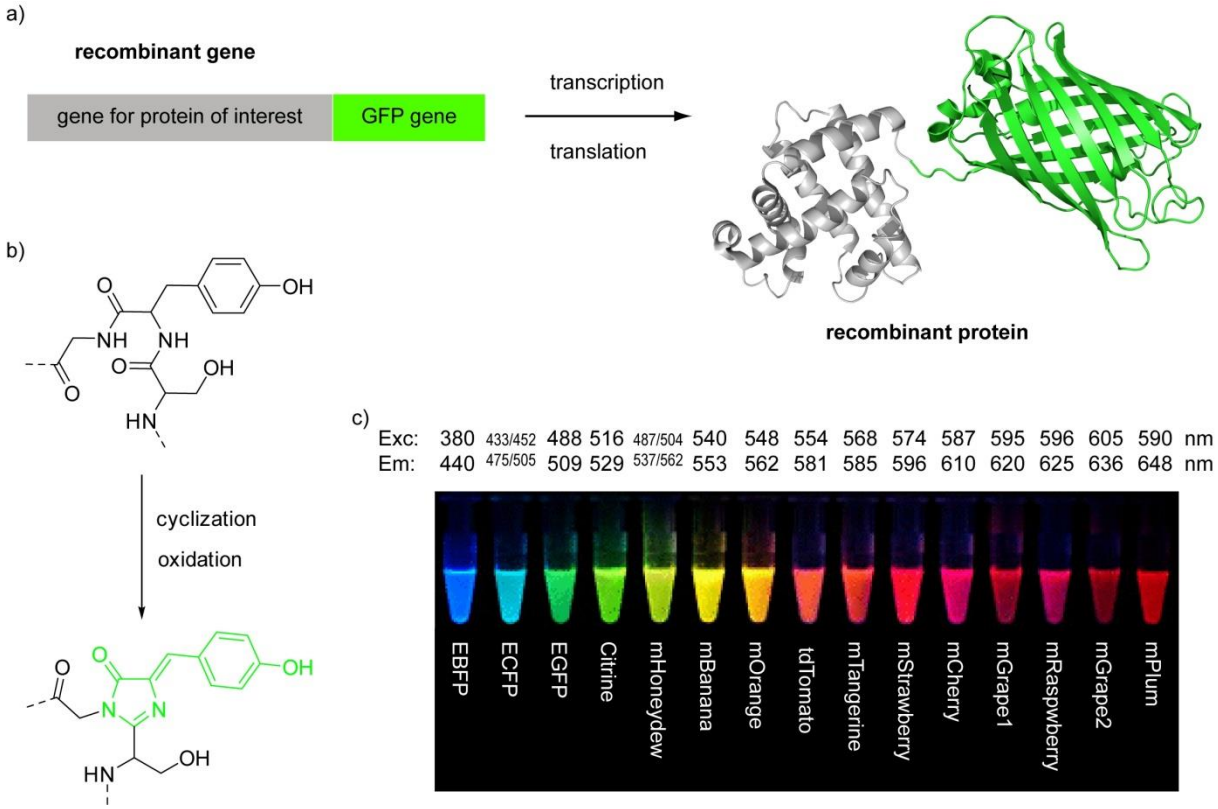


Figure I.1: Fluorescent proteins (FPs).
a) Genetic fusion of FP to POI; GFP structure in green. b) Chromophore formation of GFP. c) Overview of different FPs varying in excitation and emission wavelengths (E = enhanced versions of GFP, m = monomeric proteins and tdTomato = head-to-tail dimer).[5] Adapted by permission from Nathan C. Shaner.

Spectroscopic behavior like quantum yield, photostability and small changes in color are due to interactions with the local environment in the protein and the surrounding H-bonding network.[6] FPs are usually genetically fused to the POI (Figure I.1a) which allows the localization of proteins and analyze its dynamics in the cell with spatial and temporal resolution by visualization via fluorescence.[6b] In the last 25 years a variety of different FPs have been developed with varying photostability, pH dependence, as well as different excitation and emission spectra, so today they span most of the visible light palette (Figure I.1c).[7] Photo-activatable or photo-shiftable FPs, whose fluorescence changes, either turned on from a dark to a bright state or by shifting the wavelength when irradiated with a specific wavelength, have found use in super-resolution microscopy.[8] FPs have also been engineered to study protein dynamics via Förster resonance energy transfer (FRET) pairs and protein-protein interactions by fluorescence complementation of split FPs.[3a] The drawback of fluorescent proteins however is their size of about 30 kDa, which, when fused to the POIs, may influence their structure or affect their subcellular localization. Due to their size they are usually attached via *N*-or *C*-terminal fusion to the POI, while only in a few reported cases they were introduced within the sequence.[9] There it might also cause disturbance to the folding and dynamics of the proteins in addition to perturb its function. Additionally, small molecule fluorophores exhibit often higher brightness and better photostability.

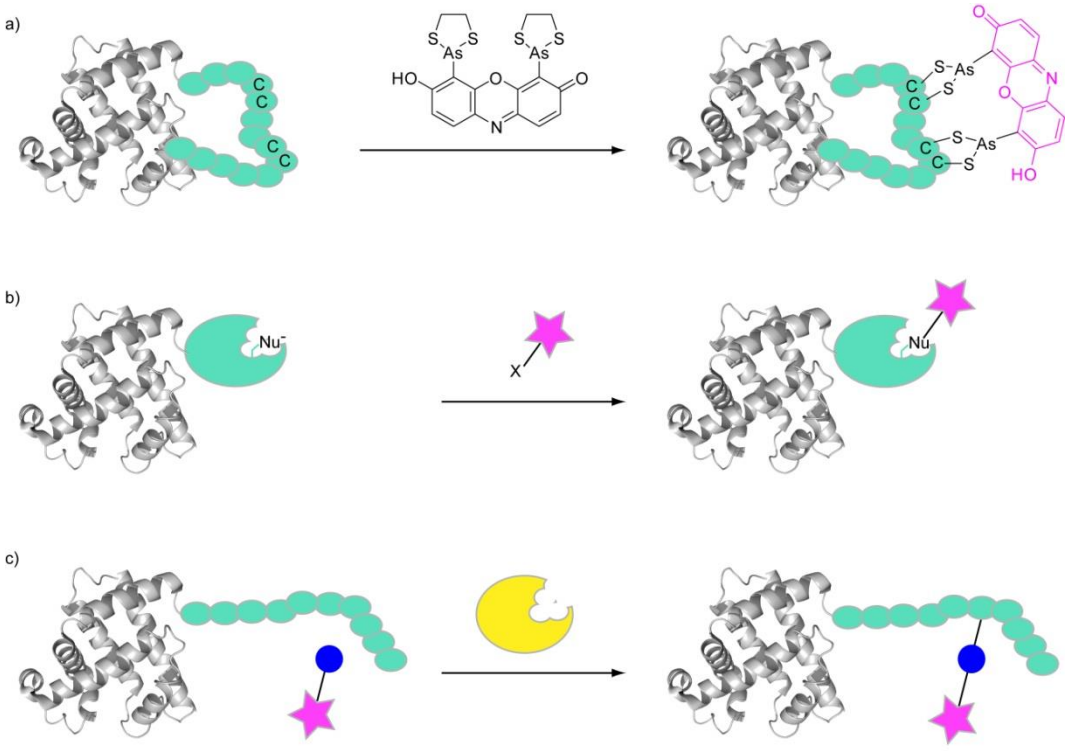


Figure I.2: Various methods for the decoration of proteins.
a) Self-labeling tags. b) Self-labeling enzymes c) Enzyme (yellow shape)-mediated labeling. Green egg shapes denote natural amino acids; pink stars represent biophysical probes (e.g. fluorophores). Nu is an abbreviation for a nucleophilic amino acid and X stands for a chemical leaving group.

To overcome these problems, alternative labeling strategies for proteins were introduced. One of those approaches genetically encodes small peptide sequences (self-labeling tags, Figure I.2a) containing four cysteines (CCXXCC, X indicating any natural amino acid besides cysteine, but XX is preferably PG) into the protein of interest, which can then be recognized

by bisarsenic modified organic dyes like fluorescein (FlAsH) or resorufin (ReAsH).[10] The highly fluorogenic dyes are administered as 1,2-ethanedithiol (EDT) derivatives to disable nonspecific interactions. Upon binding to the tetracysteine tag the fluorescence is intensified by 1000-fold. Millimolar concentration of EDT outcompete the labeling reaction and render it reversible.[10] Due to its small size, the tag can easily be incorporated into different structural motifs within the sequence, for example loops. It has been used in a number of *in vivo* applications like fluorescence[11] and electron microscopy imaging[12], the investigation of protein interactions and folding[13] as well as pulse labeling experiments.[12] However, the arsenic dyes also have limitations as they are toxic to living cells and often display a high fluorescence background due to nonspecific labeling of cysteine-rich sequences in the proteome.[14] To evade cytotoxicity, a bisboronic acid derivative of rhodamine (RhoBo), initially developed as monosaccharide sensor[15], was found to recognize and bind to a tetraserine motif (SSPGSS), but for specific labeling the sequence has to be further optimized as more than 100 proteins of the human proteome contain this sequence and were labeled with RhoBo in HeLa cells.[16]

Alternative methods to label proteins which do not suffer from the off-target labeling of the self-labeling tags use enzymes facilitating the covalent modification of the POI. One class uses so called self-labeling protein tags, which specifically recognize and attach small molecule probes to the POI they are genetically fused to (Figure I.2b). The most common tags are mutants of O^6-alkylguanine-DNA alkyltransferase (AGT), an enzyme performing DNA repair, called SNAP- and CLIP-tag, recognizing O^6-linked benzylguanine or benzylcytosine moieties respectively and transfer probes like biotin or fluorophores to the catalytic cysteine moiety in the tag.[17] The orthogonal tags have been applied for *in vivo* imaging and have been crucial in development of superresolution imaging technologies.[18] To eliminate the washing steps which remove excess of dye after labeling, a quenched O^6-benzylguanine derivate was developed, which becomes fluorescent only after covalent attachment and release of the quencher.[19] Another enzyme called haloalkane dehalogenase, which catalyzes the displacement of a alkyl halide for a water-derived hydroxy group in a bimolecular nucleophilic substitution, was genetically altered to covalently trap synthetically modified substrates for single molecule analysis[20] and to visualize location and transport of proteins in living cells.[21] Furthermore, non-covalent interactions between a receptor and a ligand have been explored as labeling approach. One example is the high affinity interaction between *Escherichia coli* dihydrofolate reductase (*Ec*DHFR) and trimethoprim, a folic acid analogue, which enables the application as tag in mammalian cells due to the 1000-fold lower affinity for endogenous mammalian DHFR.[22] A drawback of these methods based on self-labeling enzyme tags is their size of 18-20 kDa, almost the size of fluorescent proteins, which might disrupt the fold, function and localization of the studied proteins.[23]

To overcome aforementioned limitations, other approaches have been developed that use genetic in-frame fusion of small peptide sequences to the POI, which are recognized by co-expressed enzymes that mediate the reaction with a labeled substrate (Figure I.2c).[24] Phosphopantetheinyltransferases (PPTases) endogenously facilitate the transfer of CoA to an acyl carrier protein (ACP) in the NRPS pathway. They also accept CoA derivates and have therefore been applied for cell-surface labeling of ACP-fusion proteins.[25] To circumvent

perturbation of the protein by the 8-10 kDa ACP recognition sequence, a minimal motif of 8 amino acids was identified and exploited for *in vitro* and *in vivo* fluorescent labeling on the cell surface.[26] PPTase-mediated labeling is confined to the cell surface due to off-target crossreactivity with endogenous CoA. Another peptide-based labeling approach exploits the tight binding between biotin and its natural interaction partner streptavidin (K_D of 10^{-15} M^{-1}), which renders it interesting for visualization of proteins.[27] The biotin ligase BirA from *E. coli* endogenously biotinylates a specific lysine side chain of the biotin carboxyl carrier protein (BCCP), a subunit of acetyl-CoA carboxylase. The recognition sequence of the natural substrate was engineered to a minimal motif of a 15-mer, which can be attached to proteins of interest for BirA mediated cell-surface labeling.[28] Visualization of the labeling is performed with a biophysical probe attached to streptavidin, which requires washing steps to remove excess of biotin. The application is also restricted to cell-surface labeling and imaging as there are endogenously biotinylated proteins in *E. coli* and mammalian cells.[29] Another enzyme from *E. coli*, the lipoic acid ligase, which naturally catalyzes the reaction between lipoic acid and a lysine residue in an acceptor protein involved in oxidative metabolism, was engineered for protein labeling. Structure-guided mutagenesis of LplA identified mutants accepting azidoalkyl[30], fluorophore[31], trans-cyclooctene (TCO)[32], tetrazine[33] and norbornene[34] derivates of lipoic acid for enzyme-mediated modification of a 13 amino acid[35] recognition peptide sequence for cell-surface and intracellular labeling.[30, 36] The shortest peptide motif LPXTG is recognized by sortase (SrtA), a transpeptidase from *Staphylocoocus aureus*. It cleaves between the amino acids T and G and facilitates the attachment to a polyglycine peptide via a new peptide bond.[37] This approach was used to modify proteins for cell surface labeling.[38]

All these methods are based on the genetic introduction of either a peptide tag varying from 5 to 80 amino acids or a protein in the range of 20-30 kDa to the protein of interest for direct, self- or enzyme-mediated labeling, which might perturb the function or fold of the POIs. The self-labeling tags can also be introduced within the sequence, while the enzyme-base approaches are confined to being fused either on the *N*- or *C*-terminus of the POI due to the size of these modifications. The application of FPs is nevertheless very facile as they are mostly very stable proteins, which also do not show phototoxicity. Enzyme-mediated labeling takes advantage of the excellent specificity of enzyme-substrate interactions and the small size of their recognition tag; however some of these enzymes have endogenous substrates, which limit their application to the cell surface.

A powerful method emerged over the last twenty years, which allows the site-specific decoration of proteins via incorporation of unnatural amino acids (uAAs) by genetic code expansion. These uAAs bear handles with unnatural moieties which can be modified via chemoselective bioorthogonal labeling reactions.[39]

I.2 Genetic Code Expansion

The genetic information of all organisms is stored in DNA. This biopolymer has a phosphatedesoxyribose backbone bearing four different nucleobases adenine (A), thymine (T), guanine (G) and cytosine (C) on the 1' position of the sugar moiety, linked via a glycosidic bond. By Watson-Crick base pairing of A-T and G-C through two and three hydrogen bonds, respectively, DNA exists as double-stranded helix, stabilized by hydrophobic effects and van-der-Waals forces.[40] The nucleobase sequence and its translation into the amino acid sequence of a protein is referred to as the genetic code. It is represented by basetriplets consisting of three nucleobases - codons. Each basetriplet codes for one amino acid. This relation was first decoded in the 1960s in the groups of M. Nirenberg, H.G. Khorana and R. W. Holley, who were therefore awarded the Nobel Prize in 1968.[41] With 4 different nucleobases, there are 64 (4^3) possibilities to arrange codons; three of them are stop codons (amber = UAG, ochre = UAA and opal = UGA) while 61 encode for the 20 natural, proteinogenic amino acids. This means that the genetic code is degenerate and there is more than one codon for most amino acids, so called synonyms. The number of synonyms correlates with the occurrence of an amino acid in the proteome and these synonyms differ mostly in the last nucleobase. This degeneration ensures that random mutations in the genetic code do not have a harmful effect, which would in the worst case lead to truncation of the protein.[40] Protein biosynthesis does not directly occur on the DNA, but the genetic information of a gene is first transcribed into a single stranded messenger RNA (mRNA) by RNA polymerases. In contrast to DNA, RNA contains a free hydroxyl moiety on the 2' position of the ribose, resulting in reduction of stability. RNA exists mostly single stranded and base pairs adenine with the base uracil (U) instead of thymine.[40, 42]

First nucleobase	Second nucleobase: U		Second nucleobase: C		Second nucleobase: A		Second nucleobase: G		Third nucleobase
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	**U**
	UUC		UCC		UAC		UGC		**C**
	UUA	Leu	UCA		**UAA**	**Stop**	**UGA**	**Stop**	**A**
	UUG		UCG		**UAG**	**Stop**	UGG	Trp	**G**
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	**U**
	CUC		CCC		CAC		CGC		**C**
	CUA		CCA		CAA	Gln	CGA		**A**
	CUG		CCG		CAG		CGG		**G**
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	**U**
	AUC		ACC		AAC		AGC		**C**
	AUA		ACA		AAA	Lys	AGA	Arg	**A**
	AUG	**Met**	ACG		AAG		AGG		**G**
G	GUU	Val	GCU	Ala	GGU	Asp	GGU	Gly	**U**
	GUC		GCC		GGC		GGC		**C**
	GUA		GCA		GGA	Glu	GGA		**A**
	GUG		GCG		GGG		GGG		**G**

Figure I.3: Translation of the RNA codon into natural amino acids

Translation is the process of transformation of the genetic information (Figure I.3) on the mRNA into an amino acid sequence. In prokaryotes, which do not have a nucleus, translation can directly take place on the newly forming mRNA, while in eukaryotes transcription occurs in the nucleus and the mRNA has to be transported to the cytoplasm first to be translated into proteins. Transfer RNAs (tRNA), equipped with an amino acid by the corresponding aminoacyl-tRNA synthetase (aaRS), transport the amino acid to the ribosome, which decodes the information stored in the mRNA (Figure I.4). tRNAs adopt a cloverleaf secondary structure, which is folded into a 3D structure that resembles an inverted L. One arm of this reversed L structure contains the anticodon loop complementary to a codon on the mRNA, which differs for every tRNA. The other arm comprises the acceptor step at the conserved 3' end, where the tRNA synthetase charges the amino acid onto the tRNA via an ester bond. The aaRS specifically recognizes the amino acids via interactions of their side chains with its active center to avoid mislabeling of the tRNAs. The amino acid is activated by formation of a mixed anhydride with ATP under release of pyrophosphate. A nucleophilic attack of the 2' or 3' hydroxyl moiety of A76 of the tRNA on the activated carboxyl functionality leads to formation of an ester bond. The tRNA synthetase recognizes the corresponding tRNA by primary interactions with the acceptor stem and in most cases the anticodon loop. Further interactions with other structural parts depend on the tRNA synthetase/tRNA pair.[42]

The charged tRNAs are transported to ribosomes with the aid of initiation factors (IF) or elongation factors (EF). A bacterial ribosome is assembled by two subunits (30S and 50S), which freely associate for translation and. The ribosomal subunits contain the 16S-RNA (30S) and the 23S and 5S-RNA (50S), as well as a number of proteins each, which are important for structure and fine tuning of the subunits. Ribosomal RNA (rRNA) is responsible for catalytic activity and interaction with mRNA and tRNA, which is why ribosomes are also called ribozymes. Translation is a sequence of three steps: initiation, elongation and termination The mRNA contains a purine-rich (A and G) sequence, the so-called Shine-Dalgarno-Sequence upstream of the start codon AUG. Inside a sequence AUG encodes the natural amino acid methionine, while as start codon it interacts with the initiator tRNA, which bears formylated methionine. Every protein sequence starts therefore with a methionine, which can be cleaved off posttranslationally. The 16S-RNA recognizes the Shine-Dalgarno sequence by complementation, which leads to formation of 70S-initiation complex from the 30S and 50S subunit, the mRNA and the initiator tRNA, directed by the three initiation factors (IF1-3). The 50S subunit contains three binding sites for tRNAs, the A (aminoacyl), the P (peptidyl) and the E (exit) site. Elongation factors (EF) direct the elongation of the amino acid sequence by transport of the charged tRNA to the A side (EF-Tu). Catalysis of the peptide bond formation occurs at the peptidyltransferase center by placement and arrangement of the amino acid attached to the tRNA in the A site towards the activated carboxyl moiety of methionine attached to the initiator tRNA occupying the P site. The amino group attacks the activated ester on the carbonyl carbon and builds a peptide bond via a tetrahedral transition state. Simultaneously, the bond to the tRNA in the P site is cleaved leading to formation of a dipeptide attached to the tRNA in the A site. Translocation of the tRNAs to E and P site is facilitated by EF-G, which in turn is activated by GTP. In the E site the off-loaded tRNA can dissociate from the ribosome and is once more available for equipment with an amino acid by a tRNA synthetase. Elongation of the amino acid chain takes place until a stop codon in the

mRNA is reached, which induces binding of one of the two release factors RF1 or 2, as there are no tRNAs with anticodons corresponding to stop codons. This leads to hydrolytic cleavage the amino acid sequence from the tRNA and release of the protein, followed by dissociation of the 70S complex. The ribosome is a very efficient ribozyme that catalyzes the formation of about 20 peptide bonds per second with only a error rate of 10^{-4}.[42]

In bacteria, there are 30-40 different tRNAs to decode 64 different codons, with 3 of them being stop codons that do not usually encode for amino acids as well as there are no tRNAs with the corresponding anticodons. This means that some tRNAs recognize more than one codon, which can be explained by the so called wobble-hypothesis. After transport of a charged tRNA to the A site, dissociation of EF-Tu occurs only after verification of the correct base pairing of codon and anticodon by three conserved nucleobases in the 16S RNA. Here, interactions with the first two bases of a triplet are evaluated, but not the third base, which leaves some tolerance in the base pairing. Amino acids, whose synonyms differ in the third base can therefore be decoded by the same tRNA.[42-43] The genetic code is highly conserved among all organisms of life and protein biosynthesis occurs almost identical by using the 20 natural amino acids. Only a few organisms are known to incorporate the amino acids selenocysteine and pyrrolysine in response to the stop codons UGA or UAG respectively.[44] These variations of the genetic code can be used to expand it by the incorporation of new non-canonical amino acids, bearing functionalities that do not naturally exist in organisms.[45]

First methods to introduce unnatural amino acids (uAAs) into proteins were based on a chemical approach. Peptide fragments containing the wanted uAAs in their sequence were synthesized on solid phase, which were then connected via native chemical ligation to generate the protein. This requires an *N*-terminal cysteine residue that attacks a *C*-terminal peptidethioester and forms a new thioester, which can be transformed into a native peptide bond by S->N acyl shift.[46]

The easiest procedure to incorporate uAAs in live cells uses the promiscuity of the natural tRNA synthetases for residue-specific labeling. Some natural aaRS also accept derivates of their natural substrates and charge them to their corresponding tRNA. Both, the natural and unnatural substrate compete with each other, which leads to statistical labeling with the uAA of the whole proteome. This is usually performed in auxotrophic strains, in which the synthesis of one or more natural amino acids is knocked out. The labeling can be influenced via the addition of defined concentrations of amino acids. With this approach selenocysteine and the uAAs azidohomoalanine, homopropargylglycine and homoallylglycine have been successfully incorporated by the wild-type methionyl-tRNA-synthetase.[47] Other reassignments used modified leucyl- and phenylalanyl-tRNA-synthetases to decode leucine and phenylalanine sense codons.[48]

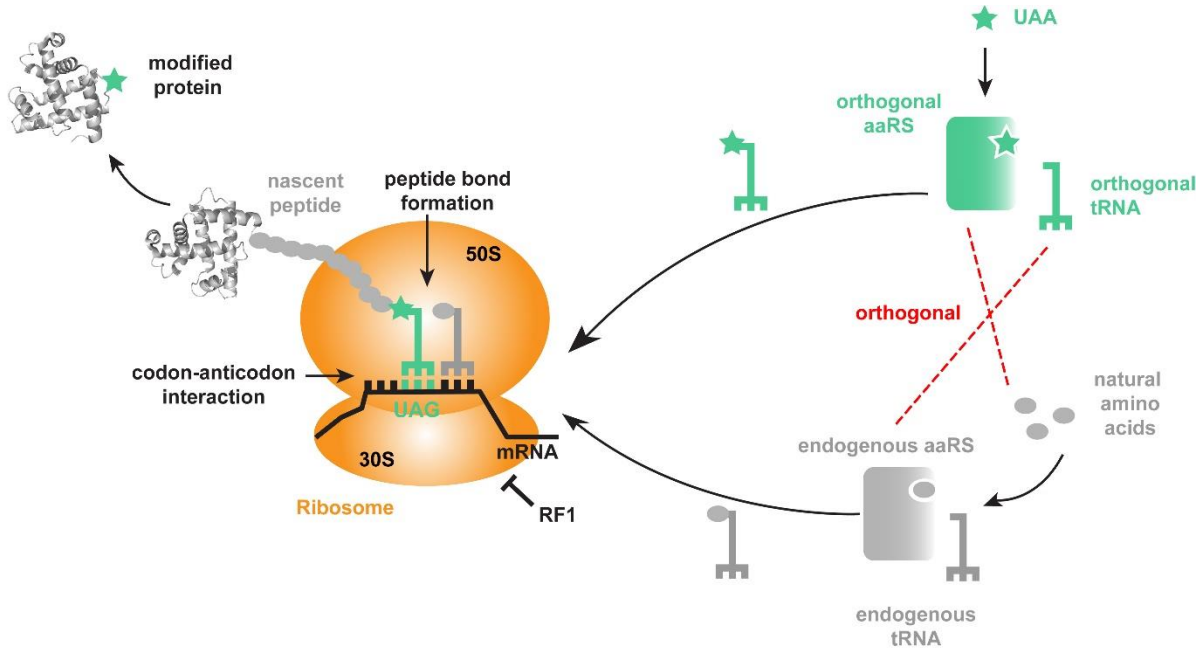


Figure I.4: Co-translational incorporation of uAAs into proteins in response to an amber codon.
Genetic information transcribed into mRNA is translated into an amino acid (grey oval shape) sequence by the ribosome by decoding base-triplets with the corresponding anti-codon on the tRNA. The endogenous or orthogonal tRNA is equipped with a natural amino acid or uAA by an endogenous or orthogonal tRNA synthetase, respectively. The resulting protein is site-specifically decorated with an uAA (green star) for further modifications.

A first approach for the site-specific introduction of a uAA into proteins uses an *in vitro* translation system, which combines all components necessary for translation in a cell-free setting. This allows the addition of artificial components like uAAs as well as modified mRNA and tRNAs. The amber stop codon was used as a blank codon, which was inserted at a specific site into the protein. Then a tRNA from yeast was modified in the anticodon loop (mutated to CUA) to efficiently suppress the UAG codon and then chemically charged on the 3' end with the uAA. In this translation system the stop codon does not lead to truncation of the protein, but is read through by the modified amber suppressor tRNA, which is why the method is also known as amber suppression.[49] The limiting factor is the chemical loading of the tRNA, which is tedious and only gives low yields, which in turn influences the amount of protein that can be produced.[50]

A breakthrough for amber suppression was the development of an approach that is able to use the cell's natural translation machinery to assemble the modified protein *in vivo*. Here, an orthogonal aminoacyl-tRNA-synthetase specifically recognizes the desired unnatural amino acid and charges it to the corresponding tRNA. The orthogonal tRNA then decodes a blank codon introduced at a specific site in the gene of the protein of interest, resulting in the incorporation of the uAA. Orthogonality means there is no crossreactivity between the orthogonal pair and the endogenous tRNA synthetases and tRNAs, concerning recognition and loading of the tRNAs to avoid misacylation.[39a] As there are no free codons in the genetic code due to its degeneration, only the three stop codons UAG, UAA and UGA are possible candidates to use as blank codons, as they are not decoded by any of the endogenous tRNAs. They usually trigger termination of protein biosynthesis by recognition of the release factors RF1 (UAA and UAG) and RF2 (UAA and UGA).[42] As the amber codon UAG is the least used in *E. coli* (~ 9 %) and yeast (~ 23 %), it is preferably utilized for incorporation of uAAs

in vivo.[51] The orthogonal aaRS/tRNA pair and the modified POI are encoded on two different vectors using molecular biotechnological methods and the uAA is fed to the cell. Cells harboring these vectors are capable of expressing all components to co-translationally produce the altered protein from an mRNA transcript with an internal amber codon in presence of uAA. The desired uAA has to be stable under physiological conditions and cannot be toxic to the cells. Competition with RF1 leads to the formation of truncated protein as a side reaction. Therefore it is desirable to have a tRNA synthetase that effectively suppresses the amber codon.[39a]

In 2001 Schultz and coworkers reported the first aaRS/tRNA pair from methanogenic archaea *Methanocaldococcus jannaschii* (*Mj, M. jannaschii*) that is orthogonal in *E. coli:* the *Mj* tyrosyl-tRNA synthetase (TyrRS). The *Mj* TyrRS was genetically engineered to incorporate the chemically modified *O*-methyl-L-tyrosine in response to the amber codon into the protein dihydrofolate reductase.[52] Since then further orthogonal pairs have been discovered and used for the incorporation of uAAs in prokaryotic and eukaryotic cells. The tyrosyl- and leucyl-aaRS/tRNA pairs from *E. coli* are only orthogonal in eukaryotes and they had to be engineered to abolish their activity towards natural amino acids. The most applied orthogonal aaRS/$tRNA_{CUA}$ pair is the pyrrolysyl-tRNA synthetase (PylRS) pair from *Methanosarcina mazei* (*Mm, M. mazei*) or *barkeri* (*Mb, M. barkeri*), which is orthogonal in bacteria, mammalian cells and animals.[45] Its natural substrate, the 22nd natural amino acid pyrrolysine was discovered in 2002 in the catalytic center of essential proteins in *M. barkeri*. There, the PylRS and its cognate $tRNA_{CUA}$ decode the UAG codon to incorporate pyrrolysine into the proteome.[44a, 53] Recently, a new PylRS/tRNA pair from *Methanomethylophilus alvus* (*Ma, M. alvus*) was engineered to be orthogonal to the *Mm*PylRS/tRNA pair.[54] The wild-type PylRS is very promiscuous and accepts structurally different amino acids to charge them to its corresponding tRNA.[39a, 51]

To develop new mutant aaRS a directed evolution approach is used (Figure I.5). Therefore, a library of synthetase mutants is generated by choosing 3 to 6 positions for saturation mutagenesis in the active site where amino acid recognition takes place. This results in an aaRS library with $10^5 – 10^{10}$ members, which are still able to be taken up by competent *E. coli* cells, where directed evolution is usually performed.[55] Alternating rounds of positive and negative selection steps reduce the number of mutants until, hopefully, a aaRS specific for the desired uAA is found. For the positive selection an amber codon is introduced into an essential gene and the desired uAA is added to the growth medium. Clones can only survive if the amber codon is successfully suppressed. As in this selection steps also false-positive hits which incorporate natural amino acids into the essential protein survive, a subsequent negative selection step is necessary to remove them. Here, the UAG codon is introduced into a toxic gene and the desired uAA is not added, which leads to death of cells, when a natural amino acid is incorporated in response to the amber codon. Surviving members of several rounds of positive and negative selection are subjected to a fluorescence-based read-out assay comparing sfGFP fluorescence in cultures grown with and without uAA. Positive hits have a modified active center, which specifically recognizes the uAA and does not show any crossreactivity with one of the natural ones.[56]

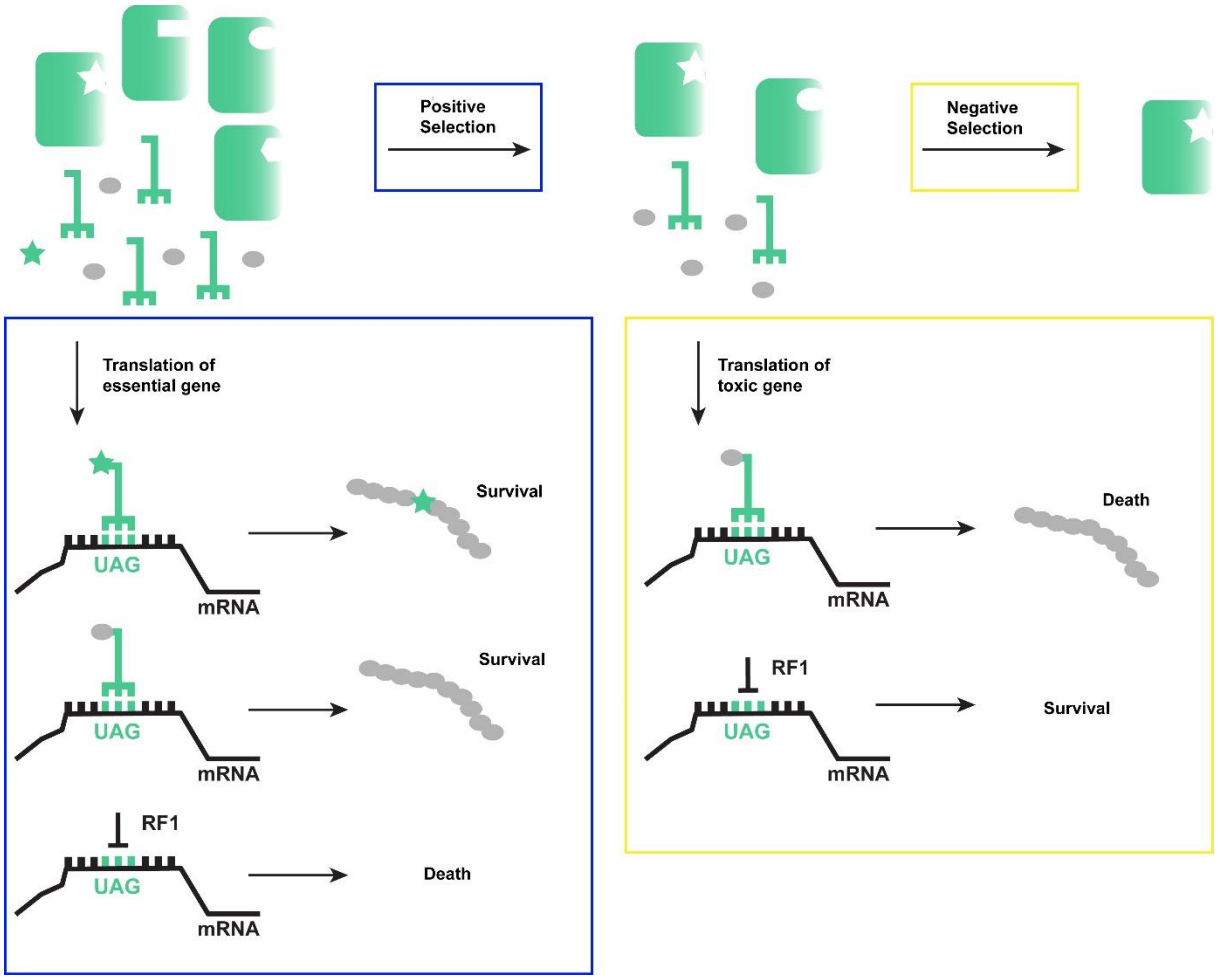


Figure I.5: Directed evolution via sequential positive and negative selection steps to find a new tRNA synthetase (green rectangle with star modification) with specific aminoacylation activity towards a desired unnatural amino acid.

Altering the recognition site by directed evolution or approaches based on rational, structure guided mutagenesis has created new mutant aaRS/tRNA pairs that incorporate a great number of new uAAs into different bacterial and eukaryotic proteins.[57] More than 250 uAAs bearing different functionalities for various applications have been introduced into proteins *in vivo* and there are reports of new uAAs being incorporated on a monthly basis.[39a, 57] The uAAs can be biophysical probes itself (Figure I.6b) or they bear bioorthogonal handles that allow the subsequent decoration via bioorthogonal reactions. The class that is mostly used are bioorthogonal ligation reactions, where new covalent bonds are formed between the uAA and the biophysical probe (Figure I.6a), for example for the site-specific positioning of fluorophores or spin labels. These modified protein can be studied and analyzed via fluorescence imaging[58] or different nuclear magnetic resonance (NMR)[59] and electron paramagnetic resonance (EPR)[60] methods for example to obtain information about structure and dynamics of a protein or its localization and pathways in a cell.[58a]

A further class of amino acids used for the covalent trapping of protein interactions with other proteins, nucleic acids or ligands are so called crosslinker amino acids. Photocrosslinkers (Figure I.6c) are activated by light impulses and react unspecifically via radical reactions with any molecule in the surrounding area. They are able to trap transient interactions, which has been used to study receptor-ligand binding in living cells.[61] Chemical crosslinkers (Figure

I.6d) react with nucleophilic amino acids in S_N2 reactions and are tuned in their electrophilicity that the reaction takes place only when enhanced by the proximity of the reaction partners.[62] With this approach low affinity complexes have been successfully stabilized for crystallization.[63]

An emerging technology in the last 10 years was the development of decaging reactions, where covalent bonds are broken instead of formed.[64] Here, natural amino acid like lysine, cysteine, serine or tyrosine are chemically masked, incorporated site-specifically into proteins and then decaged by light or chemical reactions (Figure I.6e, f). This allows studying enzyme function, when the caged amino acid is introduced in the active center of the enzyme and activity is abolished by the mask. Triggering the demasking reaction leads to restoration of their activity.[65] Recently, a enzyme-facilitated deprotection of genetically incorporated lysine bearing a phenylacetyl protection group has been reported.[66]

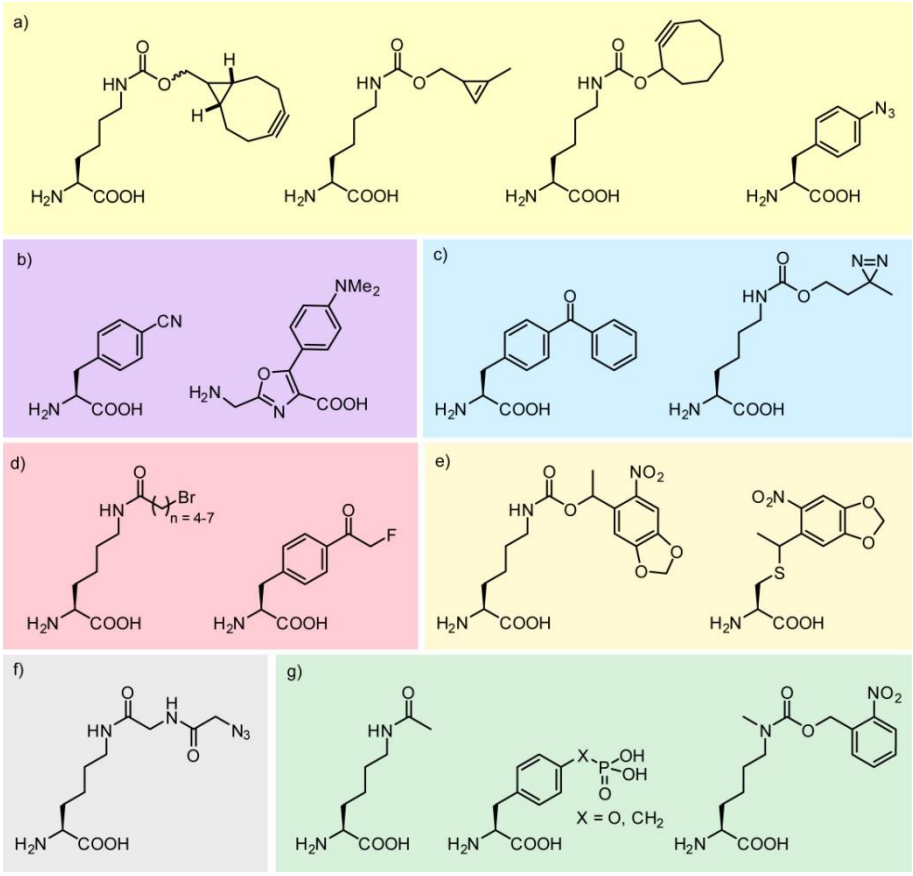


Figure I.6: Selection of unnatural amino acids that can be incorporated via genetic code expansion *in vivo*.
a) uAAs for bioorthogonal ligation reactions b) uAAs for spectroscopic applications c) photocrosslinker uAAs d) chemical crosslinker uAAs e) photo-caged uAAs f) uAA for sortase mediated ligation g) uAAs bearing PTMs.

Post-translational modifications (PTMs) of proteins such as phosphorylation, acetylation, methylation, lipidation, glycosylation and ubiquitination play an important role in many biological processes, including enzyme regulation, cell signaling, protein localization and degradation.[67] They are introduced mostly in eukaryotes after protein biosynthesis by the cell and are difficult to access with biotechnological methods, as many of the enzymes that are responsible for the modifications are still unknown. The site-specific incorporation of amino acids bearing PTMs (Figure I.6g) opens therefore possibilities to study their influence in a

biological context. aaRS have been engineered together with an improved biosynthesis of phosphoserine and phosphothreonine to use the amber suppression technology to incorporate these phosphorylated amino acids into proteins.[68] Also for the incorporation of phosphorylated tyrosine and its derivates, mutant aaRS have been developed.[69] Non-hydrolyzable analogs allow identifying enzymes that interact with PTMs. Several lysine PTMs like acetylation, methylation and crotonylation are found in histones, where they form a distinct code for epigenetic regulation. Here, genetic code expansion has been able to introduce those PTMs site-specifically as a tool to research their impact.[70] Ubiquitination occurs between the C-terminus of ubiquitin and a lysine residue in a target protein, which is ubiquitinated post-translationally via a native isopeptide bond. Monoubiquitination occurs as well as the formation of polyubiquitin chains, as ubiquitin itself has seven lysine residues which can be modified. The modification pattern determines the fate of the protein, its localization or possible degradation by the proteasome.[71] Methods to build different ubiquitination patterns by bioorthogonal ligations all resulted in non-native peptide bonds or had to be deprotected using harsh methods like desulfurization not tolerated by a great number of proteins.[72] A very recently reported approach applies the amber suppression methodology to introduce a small handle into the protein of interest for sortase-mediated ubiquitylation of the target protein *in vitro* and in live cells (Figure I.6f).[73]

As termination by release factor RF1 competes with amber suppression, only 20-30 % of full-length uAA bearing protein is produced compared to the wild-type protein.[74] This further decreases if the incorporation of the uAA is not effective or if two or more uAAs are to be incorporated into a POI. Several approaches have been developed to address this problem. One class of them focuses on the creation of RF1 knockout strains with a reinterpretation of the endogenous amber stop codons in essential genes.[75] Despite efforts to rescue the well being of these engineered bacteria, all deletion strains display a decrease in growth rate. Other attempts focused on the development of orthogonal ribosomes, which have been observed to enhance amber suppression.[74] Orthogonality here was achieved by directed evolution of the Shine-Dalgarno sequence on the mRNA and its counterpart in the 16S rRNA to force interaction between an orthogonal ribosome and an orthogonal message on the mRNA.[76] The orthogonal biosynthetic machinery is not involved with endogenous protein synthesis and is therefore more tolerant towards genetic engineering.[77] Directed evolution of the highly conserved rRNA resulted in the orthogonal ribosome ribo-X, which shows enhanced amber suppression yields also with more than one UAG codon due to a decreased interaction with RF1.[74]

To study protein structure and dynamics by FRET or the interplay of several different PTMs in one protein, the introduction of more than one type of uAA for modification with the FRET pair or various PTMs, is of great interest. Since the genetic code is degenerate and contains no blank codon, this requires the genetic recoding of a sense or another stop codon besides the commonly used amber codon for suppression of a single uAA. The recoding of sense codons is more complex, since it needs to be performed genome wide and necessitates abolishing the corresponding endogenous tRNAs decoding this sense codon.[78] Therefore, first approaches for the incorporation of two distinct uAAs into proteins were based on a combination of amber and ochre stop[79]codon, as well as pairing of amber codon with a quadruplet codon.[80]

However, the production of double-modified protein was low yielding. A quadruplet codon consists hereby of four instead of three nucleobases and expansion of the genetic code with these quadruplet codons would result in 256 (4^4) possibilities and new blank codons to incorporate uAAs into proteins. Early methods to decode quadruplet codons *in vitro*[81] and *in vivo*[82] used chemically aminoacylated tRNAs with an extended anticodon loop and suffered from low yields. To separate the quadruplet decoding from the endogenous machinery, orthogonal ribosomes were further evolved in 16S rRNA where tRNA recognition takes place to result ribo-Q1.[83] ribo-Q1 is able to facilitate incorporation of uAAs in response to amber codons, quadruplet codons or a combination of both in an orthogonal mRNA for the formation of an intramolecular bioorthogonal crosslink or dual labeling via two orthogonal bioorthogonal reactions.[84] Only the 16S subunit has been genetically engineered to be orthogonal to the endogenous one and associates with one of the endogenous 23S subunits to form the ribosomal complex. Directing specific 16S and 23S subunits towards each other for the formation of the ribosomal complex would enable evolution of the 23S subunit towards new functions. First approaches tethered both subunits together resulting in stapled ribosomes, which however suffered from a lower activity.[85] Tethering does not only lead to the desired intramolecular interaction of the linked subunits, but association with endogenous free subunits or other tethered subunits were observed. To abolish the formation of these unwanted ribosomal complexes, further engineering of the linker region was necessary. The resulting tethered ribosome also showed enhanced activity.[86] These developments may open the possibility to use genetic code expansion to decorate proteins with further unnatural monomers, like *D*- or *β*-amino acids as well as dipeptides or dipeptidomimetics.[87] A very recent publication displayed data for the incorporation of an oxazole dipeptide analog into GFP *in cellulo* with the wild-type PylRS system in response to an amber codon using a modified ribosome.[88] The bottleneck here will be finding or engineering orthogonal aminoacyl-tRNA-synthetases with activity towards these unnatural substrates.

I.3 Bioorthogonal Reactions

To understand biological processes, approaches for the decoration of biomolecules with probes that do not perturb structure and function to study these modified biomolecules in their native surrounding have been developed in the last twenty years. These methods are usually based on a two-step labeling approach. One reaction partner bearing a bioorthogonal functionality is introduced into a biomolecule via chemical, genetic or metabolic engineering. The reaction partner is added exogenously and reacts selectively and specifically with the incorporated handle. This modular approach enables the introduction of various probes into a biomolecule via one bioorthogonal reaction. Several requirements need to be fulfilled to achieve bioorthogonality: (i) the bioorthogonal reaction partners must chemoselectively react with each other without displaying crossreactivity to naturally occurring molecules; (ii) reactions have to take place under physiological conditions, meaning aqueous buffers, temperatures around 37 °C, ambient pressure and the reducing cytosolic environment; (iii) reactants cannot be toxic to the living system and have to be stable (thermodynamically, kinetically and metabolically) in this environment; (iv) the reactions have to yield stable products with covalent linkages (in case of a ligation reaction); (v) fast kinetics are favorable, as they allow to follow the fast biological processes, label low abundant biomolecules or enable the uses of only little excess of labeling agent.[39, 89] Meeting these criteria is quite a challenge, but there are a number of chemoselective reactions that have been reported over the last years that fulfill these requirements. [39b, 89-90] Some of them are less chemoselective than others in regard to naturally occurring functionalities and have been used for the decoration of biomolecules *in vitro*, while others have also been applied in *in vivo* settings like living cells or animals.

Ketones and aldehydes were among the first functionalities employed for bioconjugation reactions. Under acidic conditions (pH 4-6), their carbonyl moiety is protonated and forms a reversible Schiff base after attack of amines. Strong α-effect nucleophiles like hydrazine, hydroxyl- or alkoxyamines shift the equilibrium from the free carbonyl group to the ligation products, which are hydrazones or oximes (Figure I.7a).[91] Due to the rather slow kinetic rates[92] of 10^{-4} - 10^{-3} $M^{-1}s^{-1}$ [93] a high excess of the labeling reagent (in the mM range) is required to accomplish sufficient labeling, which might lead to a high background signal and even be toxic to the cells. As carbonyl groups are present in living cells, side reactions with the α-effect nucleophile labeling reagents might occur there. Due to these aspects these ligations are suitable for applications *in vitro* or for cells surface labeling.[94] Aniline catalysis (Figure I.7b) enables the performance of the ligation reactions at pH 7 and enhances the reaction rate of the hydrazone ligation to 170 $M^{-1}s^{-1}$ albeit a 100-1000 fold high excess of the catalyst is needed.[95] The modified versions have been used to label biomolecules on the cell-surface and in the cell.[96]

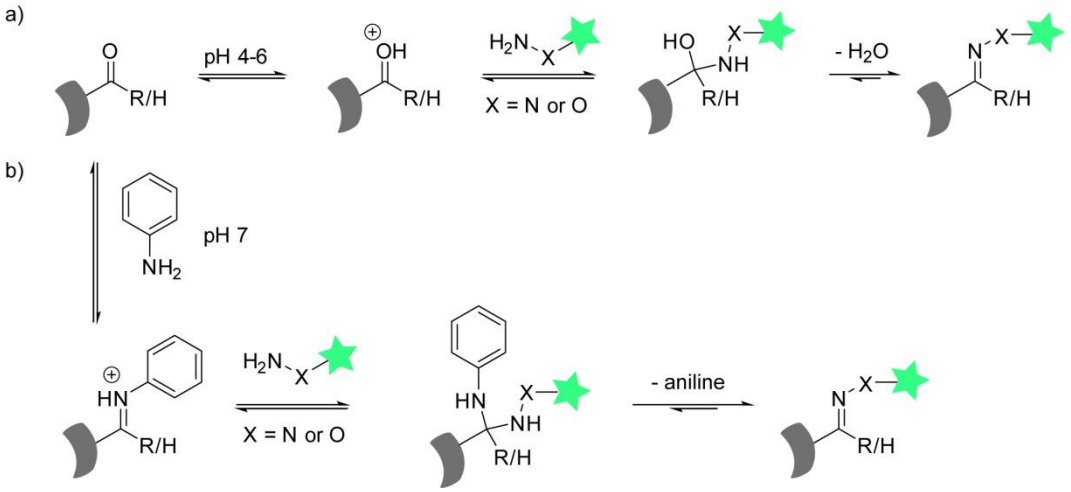


Figure I.7: Labeling of aldehydes and ketones with hydrazine (hydrazone ligation) and hydroxyl- or alkoxyamines (oxime ligation). a) acid catalysis b) aniline catalysis. Grey shape denotes a biomolecule and the green star represents a biophysical probe (e.g. fluorophore).

In contrast to ketones and aldehydes, azides are mostly not present in biological systems and therefore azides are truly bioorthogonal functionalities that react in a number of different bioorthogonal chemoselective reactions.[39, 94a, 97] In one of these reactions, known as the Staudinger reduction (Figure I.8a), azides are reduced to primary amines upon treatment with phosphines. The phosphorus atom attacks on the terminal electrophilic nitrogen atom of the azide under formation of an aza-ylide after release of N_2 from the phosphazide intermediate. The aza-ylide hydrolyzes to the primary amine and phosphine oxide, which renders the Staudinger reduction a decaging reaction instead of a bioconjugation. Azides have been used to mask primary amines and the decaging reaction has been triggered using phosphines in living cells for chemical control of protein function,[98] antigen cross presentation[99] or to provide a free glycine moiety for sortylation.[100] To apply the Staudinger reaction as a bioconjugation reaction the phosphine reactant needed to be modified to electrophilically trap the aza-ylide before destructive hydrolysis. Therefore, an ester was introduced in *ortho* position to the phosphorus atom which directs the formation of a five-membered ring via intramolecular cyclization, followed by hydrolysis to a stable amide bond. This Staudinger ligation reaction (Figure I.8b) displays rather slow kinetics of $10^{-3}\,M^{-1}s^{-1}$, with the nucleophilic attack of the phosphine on the azide being the rate-limiting step of the reaction. Phosphines with increased nucleophilicity that would show faster kinetics are in turn more prone to oxidation by air, making them unreactive.[101] A further modification of the phosphine reagents lead to the development of a "traceless" Staudinger ligation (Figure I.8c), where the phosphine oxide is not part of the ligation product anymore. For this the ester was arranged as phenolic ester, which releases the phosphine oxide upon formation of the native amide bond.[102] A similar approach uses phosphinethiol to link two peptide fragments - one containing a *C*-terminal thioester, the other one an *N*-terminal azide moiety as an alternative method to native chemical ligation.[103] These different types of Staudinger ligations have been used to successfully label various biomolecules *in vitro* and in living cells and animals.[102-104]

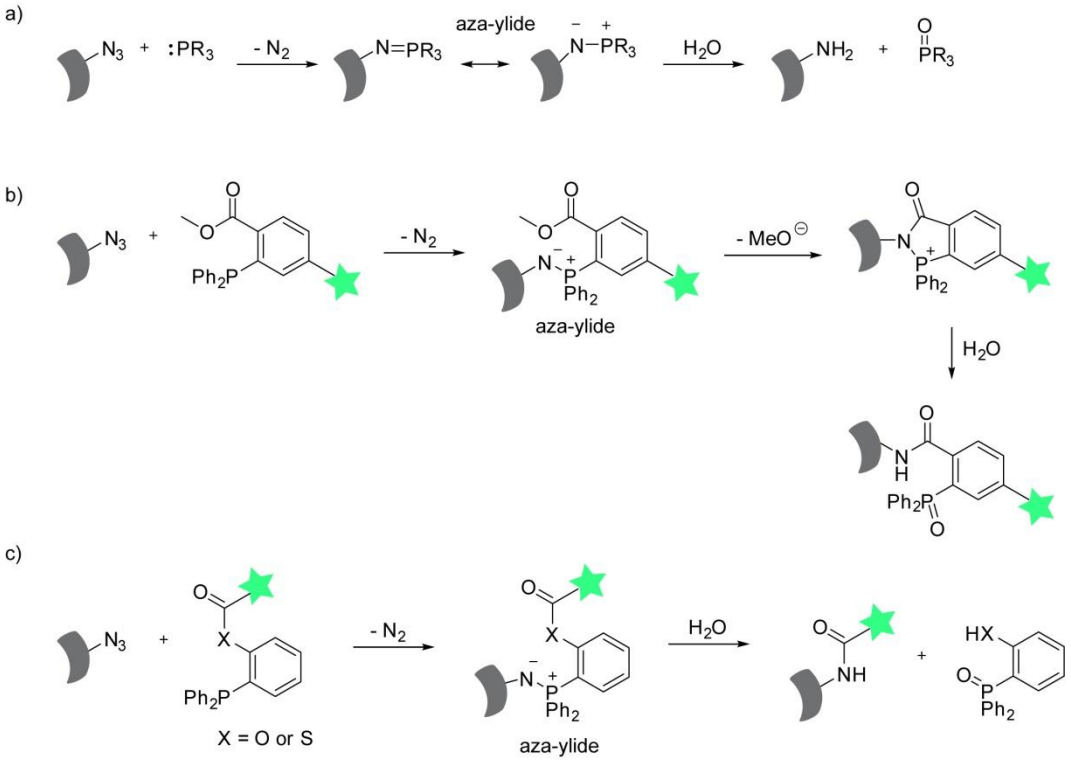


Figure I.8: a) Staudinger reduction from azide to amine b) Staudinger ligation to form a stable amine c) "traceless" Staudinger ligation. Grey shape denotes a biomolecule and the green star represents a biophysical probe (e.g. fluorophore).

Azides can also be described as 1,3-dipoles and can undergo [3+2] cycloaddition reactions with dipolarophiles such as activated alkynes to form stable triazoles. Huisgen first described the reaction in the 1960s. Without activation of the alkyne the reaction requires high temperatures and/or pressure, which renders it unsuitable for biological applications.[105] Around 20 years ago the groups of Sharpless and Meldal independently reported the activation of terminal alkynes with copper(I) salts with a rate enhancement of 6-7 orders of magnitude compared to the uncatalyzed reaction.[106] Cu(I) catalysis also leads to the predominant formation of 1,4-disubstituted 1,2,3-triazoles.[107] This copper catalyzed azide alkyne cycloaddition (CuAAC) is nowadays also known as "click" reaction (Figure I.9a).[107] The reaction enhancement is due to the formation of a Cu(I) acetylide after coordination of the alkyne to the metal. The azide then also coordinates to Cu(I) via the nitrogen next to the carbon atom and forms a six-membered Cu(III) metallacycle intermediate that proceeds to a triazolyl-copper derivative, followed by protonolysis to the triazole product (Figure I.9b).[108] Compared to the Staudinger ligation the reaction kinetics of CuAAC reactions are about 25 fold faster,[109] which enabled their successful application to label azide- or alkyne-modified nucleic acid[110], lipids[111] and virus particles[112], as well as to functionalize proteins *in vitro* and *in vivo*.[109b, 113] Drawback of the reaction is the cellular toxicity of Cu(I), as it induces the formation of reactive oxygen species (ROS) from O_2,[114] which influence cell viability of *E. coli* and mammalian cells.[113b, 115] Decreasing Cu(I) concentration which would go hand in hand with a lowered toxicity leads to significant decrease of the reaction rate.[116] To overcome these problems water soluble ligands[117] that coordinate the catalyst (Figure I.9c) or azides with a copper-chelating moiety[118] have been used to reduce the necessary Cu(I) concentration into low micromolar range.[116b]

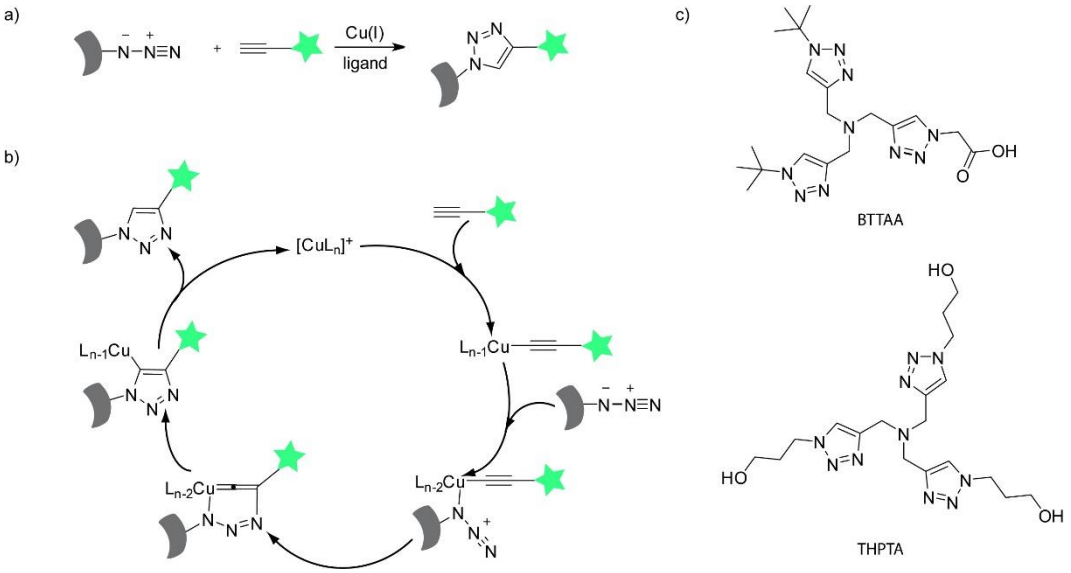


Figure I.9: a) Copper(I) catalyzed azide alkyne cycloaddition (CuAAC) b) reaction mechanism of CuAAC c) water-soluble ligands 2-(4-((bis((1-(*tert*-butyl)-1*H*-1,2,3-triazol-4-yl)methyl)amino)methyl)-1*H*-1,2,3-triazol-1-yl)acetic acid (BTTAA) and tris((1-hydroxy-propyl-1*H*-1,2,3-triazol-4-yl)methyl)amine (THPTA). Grey shape denotes a biomolecule and the green star represents a biophysical probe (e.g. fluorophore).

Another approach to eliminate the cytotoxicity of the Cu(I) catalyst utilizes the introduction of ring strain into the alkyne to activate the reaction, which is also known as strain-promoted azide alkyne cycloaddition (SPAAC, Figure I.10a).[119] This reaction was first described between phenyl azide and cyclooctyne in the 1960s by Wittig and Krebs.[120] Cyclooctyne is the smallest stable cyclic compound harboring an alkyne with a ring strain of about 20 kcal/mol.[121] The first probes for SPAAC are based on this scaffold. The ether-substituted cyclooctynes (OCT, Figure I.10a) reacted only slowly with benzyl azides in aqueous acetonitrile with reaction rates of 0.0012 $M^{-1}s^{-1}$,[109a] which is slower than the Staudinger ligation with 0.0025 $M^{-1}s^{-1}$.[122] However, this reaction could be applied for cell-surface labeling of azide-modified glycans avoiding the negative effects of the Cu(I) catalyst.[119] A further improvement concerning the reaction rate was the introduction of electron-withdrawing groups like fluorines (3-fluorocyclooctyne (MOFO) and 3,3-difluorocyclooctyne (DIFO) derivatives, Figure I.10b)[109a, 123] into the scaffold to lower the activation energy of the concerted [3+2] cycloaddition with azides compared to cyclooctyne, which resulted in a 63-fold rate enhancement.[121, 123] These modified alkynes have been used to label azide-bearing glycans on the surface of live cells, in mice and zebrafish embryos.[124] The change in electronic properties however make those compounds more susceptible towards attack by cysteines and therefore less bioorthogonal.[125] Other approaches modulate the ring strain of the cyclooctyne scaffold by fusion of two aryl rings (Figure I.10c), which increases the ring strain of the compound, resulting in dibenzocyclooctyne (DIBO) derivatives, which have similar rate constants as DIFO.[126] The implementation of another sp^2-like center into DIBO further enhances the reaction rates to values around 0.31-1 $M^{-1}s^{-1}$ as observed with aza-dibenzocyclooctyne (DIBAC or ADIBO, Figure I.10c) and an amide bond-containing derivative (BARAC).[126a, 127] Instead of introducing two aryl rings, another approach fused a pyrrol moiety to the cyclooctyne scaffold to yield pyrrolocyclooctyne (PYRROC, Figure I.10c), which displays the highest reaction rates of around 500 $M^{-1}s^{-1}$ in aqueous buffer with azides so far.[128]

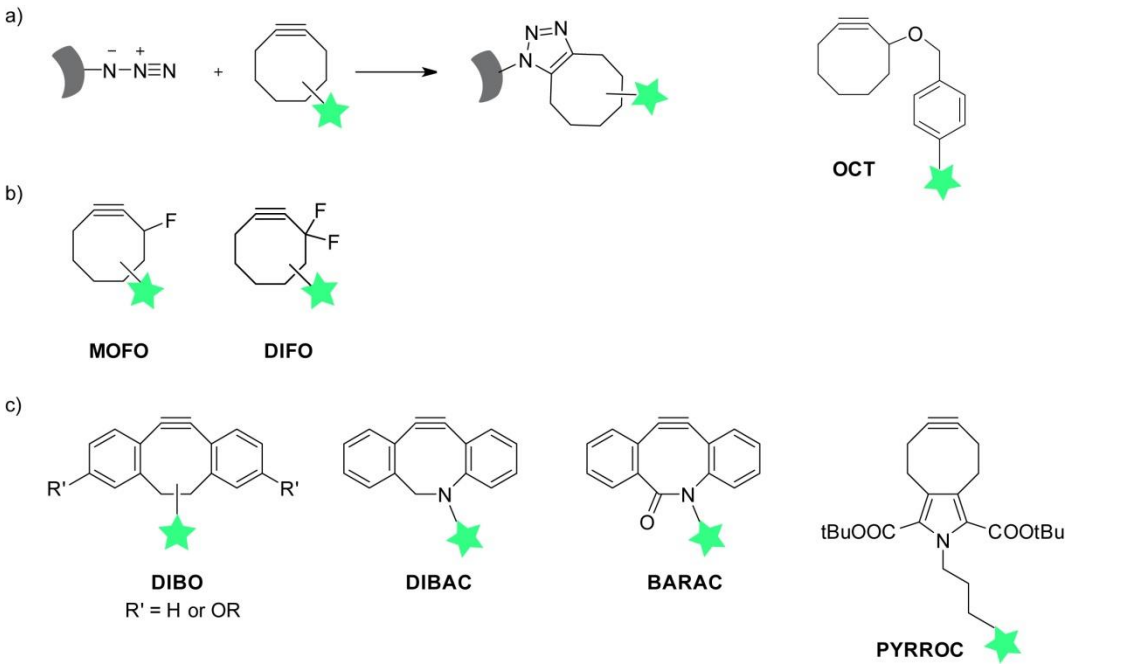


Figure I.10: a) Strain-promoted azide alkyne cycloaddition (SPAAC). b,c) Tuning of SPAAC reactivity by modulation of electronic properties (b) or ring strain (c). Grey shape denotes a biomolecule and the green star represents a biophysical probe (e.g. fluorophore).

To expand the scope of strain-promoted [3+2] cycloadditions concerning kinetics and biocompatibility after modulation of the strained alkynes different 1,3-dipoles like nitrones[129], nitrile oxides[130], nitrile imines[131], nitrile ylides[132], sydnones[133] and diazo compounds[134] have been explored and applied for *in vitro* bioconjugation of proteins.

In recent years also organometallic reactions like the ruthenium-catalyzed olefin metathesis and the palladium (Pd)-facilitated crossmetathesis have been employed as bioorthogonal labeling reactions for *in vitro* or cell surface applications.[135] Sonogashira coupling was mediated by palladium nanoparticles in living *E. coli* cells for protein modification.[136] Palladium has also been used as a trigger for the decaging of *N*-propargyloxycarbonyl on cell surface glycans.[137]

Rediscovered almost a decade ago, the inverse electron demand Diels-Alder cycloaddition (iEDDAC) between 1,2,4,5-tetrazines and strained alkenes or alkynes has become one of the most applied and versatile bioorthogonal ligation reactions, as it displays very fast reaction kinetics and shows good biocompatibility in live biological settings.[90a, 138] The following chapter will explain the reactivity, synthesis of the reaction partners and the biological applications of iEDDA cycloadditions in more detail.

I.4 Tetrazine Ligation – inverse electron demand Diels-Alder cycloaddition (iEDDAC)

I.4.1 Reactivity of the iEDDAC

In the late 1920 Otto Diels and Kurt Alder planted the basis for the so called Diels-Alder (DA) reaction by first describing the mechanism of a cycloaddition between cyclopentadiene and quinone.[139] The Diels-Alder reaction generates a (hetero)cyclohexene system in a pericyclic[4+2]-cycloaddition reaction between a diene (a 4π-electron system) and a dienophile (a 2π-electron system). [140] The reaction proceeds stereospecifically by suprafacial interaction of the diene and the dienophile. The electronic properties of diene and dienophile, influenced by their substituents, categorize the reaction into a DA cycloaddition with normal electron demand and inverse-electron demand Diels-Alder cycloadditions (iEDDAC). In a normal DA reaction, an electron-deficient dienophile reacts with an electron-rich diene, while the iEDDAC proceeds reversibly between an electron-poor diene and a dienophile bearing electron-donating groups (EDG). The energy difference between the HOMO (highest occupied molecular orbital) and the LUMO (lowest unoccupied molecular orbital) of the two reaction partners correlates with the reaction rate of the cycloaddition. The smaller the difference between them, the faster the reaction will occur. Therefore, the reaction rates can be tuned by introduction of electron-withdrawing groups (EWG) and substituents which donate electrons on the diene or dienophile (Figure I.11).

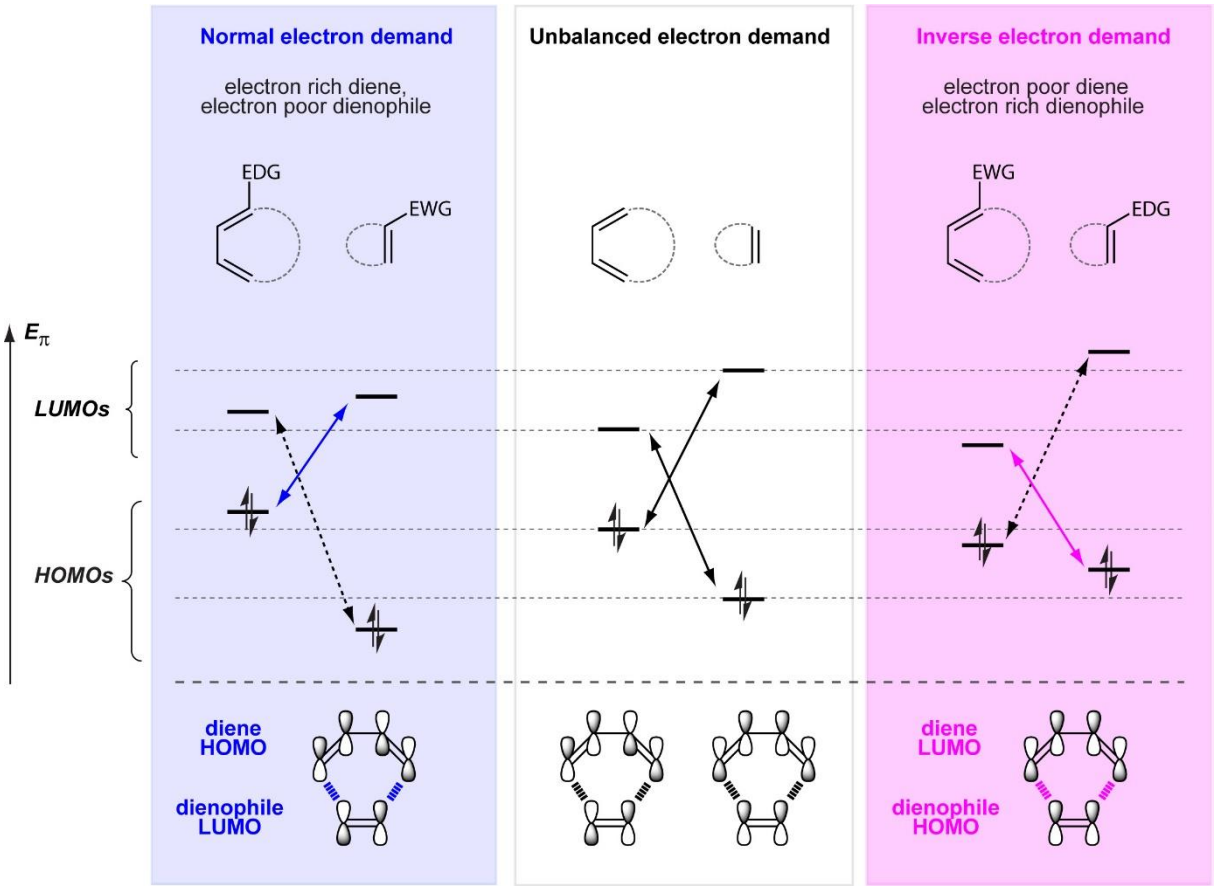
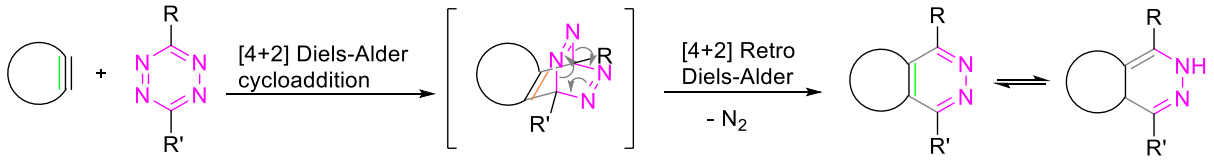


Figure I.11: Participating MOs and the energy difference between them determine if a Diels-Alder cycloaddition has a normal or an inverse electron demand. EWGs (electron-withdrawing groups) lower MOs, while EDGs (electron donating groups) raise them, which greatly influences the kinetics of DA or iEDDAC reactions.

Another rate enhancement was observed in polar solvents like methanol and water, which renders the application interesting in biological and physiological systems.[141] Bioorthogonal reactions require stable reactants without interactions with endogenous molecules, which would be challenging for the electron-deficient dienophiles used in a normal DA, as they are likely to react with endogenous nucleophiles *in vivo*.[135a] iEDDAC cycloadditions between 1,2,4,5-tetrazines and strained alkenes or alkynes in turn have become a valuable tool for the decoration of proteins via bioorthogonal chemistry.



Scheme I.1: Mechanism of an inverse electron demand Diels Alder cycloaddition (iEDDAC) reaction between strained alkenes/alkynes (green bond). Formation of a bicyclic intermediate by a [4+2] Diels-Alder cycloaddition, followed by extrusion of nitrogen in a [4+2] retro Diels-Alder step.

The first report of a Diels-Alder reaction with inverse electron demand (iEDDAC) was in 1949 by Bachmann and Deno, who described a number of electrophilic dienes, which reacted with nucleophilic alkenes.[142] In the late 1950, Lindsey and Carboni first mentioned 1,2,4,5-tetrazines as suitable electron-deficient dienes for the synthesis of pyridazines, which is why the iEDDAC is also known as Carboni-Lindsey reaction.[143] Since then a number of synthetic strategies in organic synthesis have relied on it, especially for the assembly of complex ring structures of natural products.[144] The mechanism of an iEDDAC (Scheme I.1) between an electron-poor azadiene and an electron-rich dienophile can be divided into two parts: first a [4+2] Diels-Alder cycloaddition takes place and results in a bicyclic intermediate, followed by a retro-Diels-Alder step, where nitrogen is released and the corresponding dihydopyridazine conjugation product (pyridazine product if the dienophile is an alkyne) is formed. The reaction, in which the cycloaddition represents the rate limiting step, displays second-order kinetics. Like the normal Diels-Alder reaction, the iEDDAC between tetrazines and strained alkenes proceeds with a high stereo-, regio- and chemoselectivity.[144-145]

The first experiments displaying the rapidness of iEDDACs date back to 1990, where Thalhammer *et al.* showed that second order rate constants between 3,6-bismethoxy- (**9**) or 3,6-bistrifluoromethyl-1,2,4,5-tetrazine (**8**) and diverse strained alkenes and alkynes in dioxane can reach rate constants up to 10^4 $M^{-1}s^{-1}$.[146] Unfortunately, these tetrazines are too reactive to be used in a biological context. Therefore a number of differently substituted tetrazines have been synthesized in the last few years and studied under physiological conditions. UV/VIS spectroscopy, fluorescence spectroscopy and/or ^{1}H NMR spectroscopy were the main methods used to determine second order rate constants of the iEDDAC reaction with different dienophiles under pseudo first order conditions.[58a, 147] Substituents on the reaction partner influence electronic properties of the reaction by modulation of the HOMO-LUMO gap. EWGs on the tetrazine and electron-donating groups on the dienophile as well as ring strain minimize the gap between the interacting MOs and hence favorably influence the reactivity and rate constant of the reaction (Figure I.12). [145, 148]

Karver *et al.* confirmed this correlation comparing the reactivity of a number of differently 3,6-substituted tetrazines (**2-4, 6**) under physiological conditions towards *trans*-cyclooctene (TCO, **17**), the smallest stable cycloalkene with a trans-configured double bond.[147a, 147b] Substituents in positions 3 and 6 were varied to alkyl groups, protons or heteroaryl moieties. Here, dialkyltetrazines (**2**) display the slowest rate constant of 10^2 $M^{-1}s^{-1}$ due to the weak donating characteristics of the alkyl moieties.[147b] Tetrazines, carrying an alkyl and an aryl (**3**) react with a slightly faster rate constant in the range of 10^2 - 10^3 $M^{-1}s^{-1}$.[147b, 147c] The introduction of heteroatoms into aryl rings such as pyridinyl or pyrimidyl substituents (5-**7**) further accelerate the reaction rates up to another magnitude.[147c] **7** is therefore on par with monosubstituted tetrazines (**4**), which were the fastest reacting tetrazines in that study, owing to a decrease in steric hindrance.[147b] Due to high cell permeability they have often been used for intracellular labeling.[58a, 147b, 149] Donor-substitution of tetrazines (**1**), for example with amines, leads to a higher LUMO, therefore they display the lowest kinetics (10^{-3} $M^{-1}s^{-1}$ with TCO) towards dienophiles.[147b, 150] Their reactivity can be recovered by 400-700-fold to 0.5-9.5 $M^{-1}s^{-1}$ by acylation of the amino moiety.[150] An increase in reactivity does however often correlate with a decrease in stability. Monosubstituted tetrazines and other reactive tetrazines are prone to degradation and show reactivity towards cysteines, which is a problem for application in *in vivo* settings.[147b, 151]

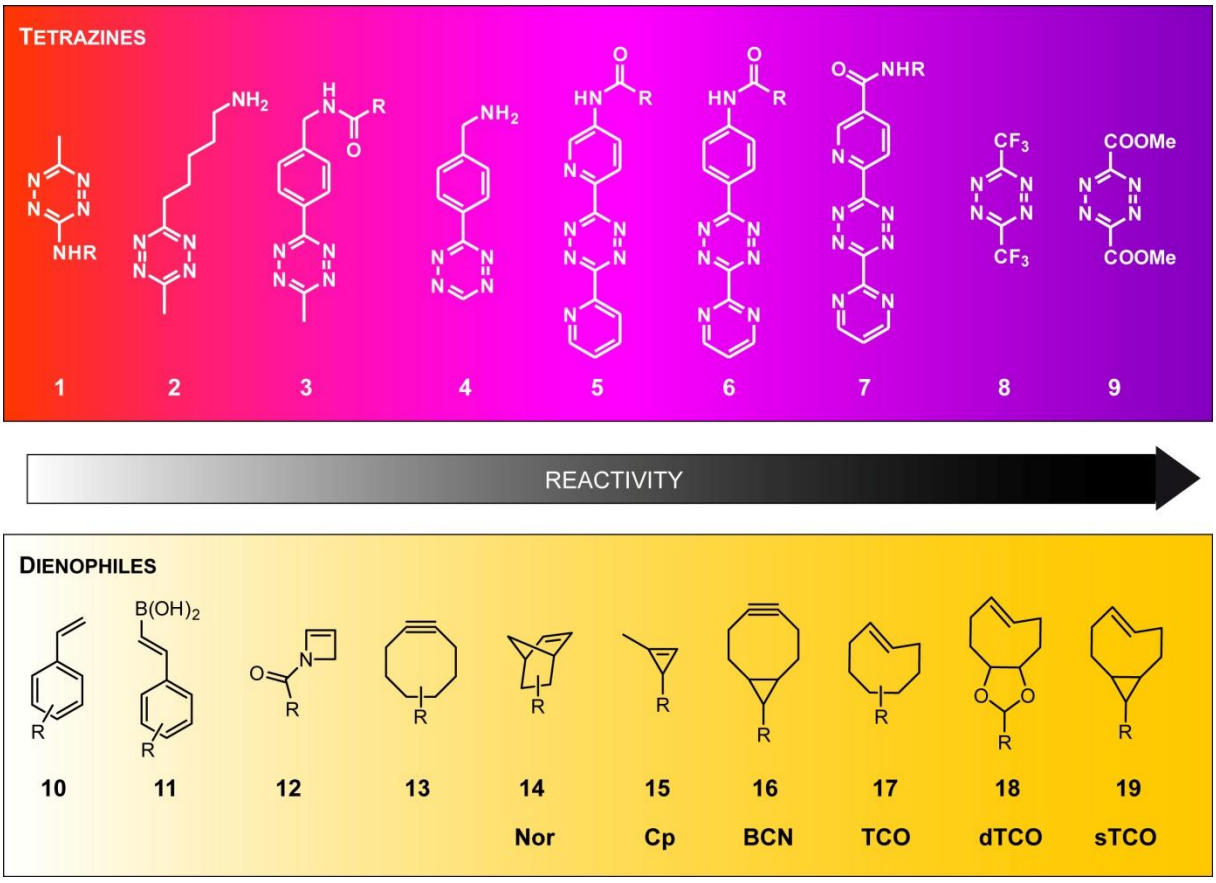


Figure I.12: Reactivity of tetrazines and dienophiles in an iEDDAC

The reaction rates of iEDDA cycloadditions also depend strongly on the dienophiles used. Unstrained olefins like terminal alkenes (**10**) display slow reaction rates of $\leq 10^{-2}$ $M^{-1}s^{-1}$ with tetrazines.[152] A boronic acid substituent (**11**) in vinylogous position can increase these rates

up to 27 $M^{-1}s^{-1}$, but this seems to be only the case with dipyridyltetrazines possibly due to coordination of the boron atom to the pyridyl nitrogen.[153] With monosubstituted tetrazines they displayed a much lower reactivity than expected. Introduction of ring strain is the main factor influencing the kinetics. Norbornene (**14**), which has a relative ring strain of 27.2 kcal/mol[154], was one of the first dienophiles described for use in biological applications.[149] Vrabel *et al.* performed a study on the influence of different substituents on norbornene on the reaction rates with dipyridyltetrazine. The fastest reaction rate was observed for *exo*-5-norbornene-2-methanol with 5 $M^{-1}s^{-1}$. The *endo* product reacted 4 times slower. Further decrease could be identified in the presence of a carboxyl moiety in place of the donating methanol substituent and also by the incorporation of heteroatoms into the norbornene scaffold, which decrease electron density and lower the involved HOMO.[155] Also cyclooctynes (**13**) have been used as dienophiles in iEDDAC with tetrazines, but due to the lower ring strain of 20 kcal/mol compared to norbornene, they react more sluggishly.[121, 147e, 156] With monosubstituted tetrazines they react slightly faster.[156-157] DFT calculations propose these reactions follow a normal electron demand DA cycloaddition instead of inverse-electron demand, due to the electrophilicity of the cyclooctyne.[157c]

Cyclopropene (Cp, **15**) is the smallest cyclic alkene and possesses an increased ring strain (54.5 kcal/mol[154]), which results in faster reaction rates. The first indication that the placement of substituents has an influence on the reactivity of Cp was demonstrated by Thalhammer *et al.* in the 1990s. For 3,3-dimethylcyclopropene they observed a 1000-fold decrease of the reaction rate, compared to 3-methylcyclopropene.[146] Cps without C1 substituent display a higher reactivity, but were found to degrade overnight. The introduction of a methyl substituent in position C1 as reported by Devaraj and coworkers in 2012 is important to facilitate stability under physiological conditions to reach a balance between displaying a certain degree of reactivity towards tetrazines (up to 2.8 $M^{-1}s^{-1}$), while being stable enough for a longer period of time.[158] Kamber and colleagues confirmed these results and reported the selective modification of model tetrazines with 1,3-disubstituted cyclopropene in the presence of 3,3-disubstituted, characterizing them orthogonal to each other. Electron-donating moieties like hydroxyl- or aminomethyl groups on the Cp moiety demonstrate an enhancement of the rate constants of about two orders of magnitude in contrast to electron-drawing moieties like carboxyl substituents.[147f]

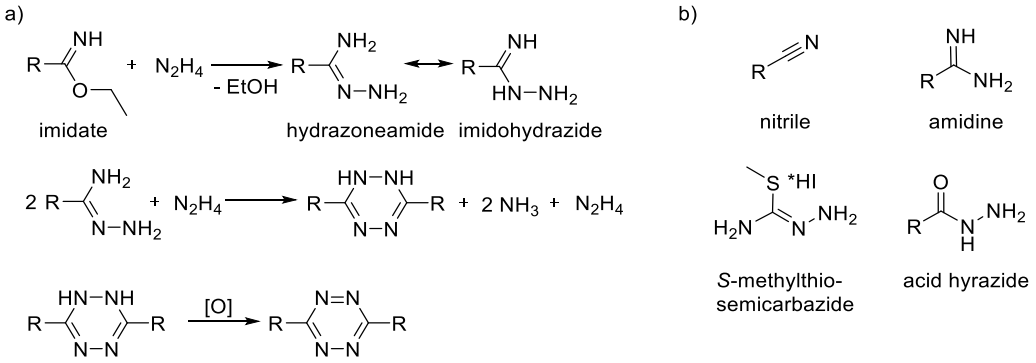
Bicyclo[6.1.0]nonyne (BCN, **16**), initially developed by Dommerholt *et al.* as a novel ring-strained alkyne for metal-free cycloaddition reactions with azides and nitrones, was found to react ~1000 times faster (10^2-10^3 $M^{-1}s^{-1}$) with tetrazines than the fastest norbornene derivative.[58a, 159] The ring strain of BCN is increased by the fusion of a cyclopropane ring. Even faster reaction rates were observed between dipyridyltetrazines and TCO (**17**, 10^3-10^4 $M^{-1}s^{-1}$), in which the double bond is forced into the less stable *trans* configuration of the 8-membered ring.[147a, 160] The first application in water, cell media and cell lysate of the tetrazine ligation using TCO was reported in 2008 by Fox and coworkers.[147a] They developed an even more reactive dienophile bicyclo[6.1.0]non-4-ene-9-ylmethanol (sTCO, **19**), by fusing a cyclopropane ring in *cis* to TCO forcing it to adopt a highly strained half-chair conformation.[147g] sTCO displays reaction rates with different tetrazines in the range of 10^5-10^6 $M^{-1}s^{-}$[158a, 147g, 161], however inactivation of sTCO was observed in several *in vivo*

applications by isomerization to the more stable and thus less reactive isomer.[56b, 58a, 151b] A dioxolane fusion instead of the cyclopropane ring was able to stabilize the resulting strained dTCO (**18**) under physiological conditions, while retaining its high reactivity.[162]

Chemical manipulation of the electronic properties of 1,2,4,5-tetrazines by substituting positions 3 and/or 6 and introduction of ring strain and EDGs into the reacting dienophile enable to modulate the reaction rates of iEDDAC cycloadditions under physiological conditions by 7-8 orders of magnitude. This allows to meet the demands of various applications in a biological context and to design orthogonal bioorthogonal reactions.

I.4.2 Synthesis strategies for tetrazines

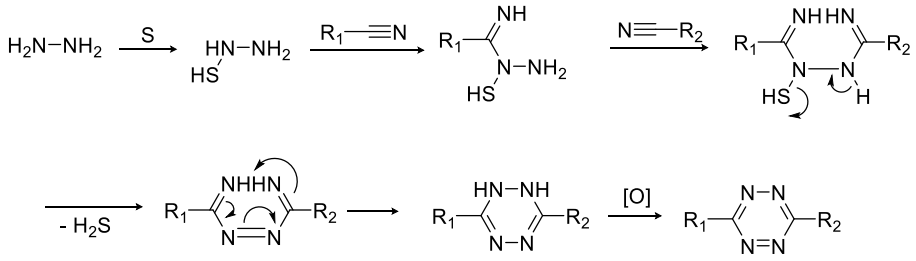
1,2,4,5-tetrazines are known since the beginning of the 20th century. First reports date back to 1893 and 1897, when Adolf Pinner published a synthesis from imidoesters and hydrazine, which is still known as the Pinner synthesis (Scheme I.2a).[163] Hydrazine attacks the imidoester in a nucleophilic manner under release of alcohol and formation of a monohydrazidine, an amidrazone, which can be described by two tautomeric structures, hydrazoneamide and imidohydrazide. Excess of hydrazine leads to formation of the dihydrotetrazine by condensation of two amidrazones and release of two ammonia molecules upon ring formation. In the following step, the tetrazine is obtained by oxidation,[164] utilizing nitrous acid[165], nitrous gases,[166] ferric chloride,[167] bromine,[168] hydrogen peroxide,[166] DDQ[169] or sodium nitrite[143a] in acidic media, the latter being most widely applied.



Scheme I.2: a) Mechanism of 1,2,4,5-tetrazine by Pinner synthesis. b) Various starting materials used for the formation of tetrazines.

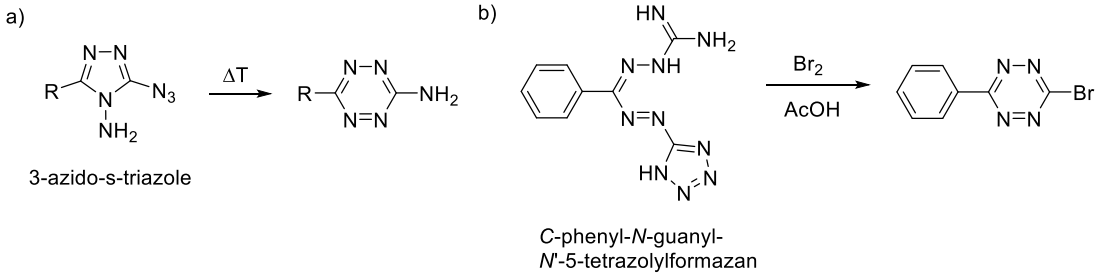
The same mechanism is assumed for the synthesis of 1,2,4,5-tetrazines (also known as s-tetrazines) from hydrazine hydrate and electrophilic compounds such as nitriles,[143b, 170] the currently most widely used method to synthesize tetrazines. The reaction differentiates only in the starting materials used for the formation of the amidrazone. Other reactants apart from nitriles that produce symmetric 1,2,4,5-tetrazines include thioamides,[171] acid hydrazides,[172] ethyl diazoacetate,[173] diaminoguanidine and *S*-methylthiosemicarbazide (Scheme I.2b).[165] The different starting materials enabled access to symmetric tetrazines with various substituents, although in poor yields. A first improvement in this direction was the addition of elemental sulphur to the reaction mixture as a catalyst, which was proposed to activate hydrazine for nucleophilic attack on the first nitrile (Scheme I.3).[174] Subsequently, the formed amidrazone

can attack on a second nitrile, followed by cyclization to 1,2-dihydrotetrazine, which can then be oxidized to the tetrazine.



Scheme I.3: Proposed mechanism in literature[174] for the sulfur-induced synthesis of 1,2,4,5-tetrazines.

A number of aromatic substituted dihydrotetrazines could be generated with yields from 70 to 95 %. However, the yields of aliphatic tetrazines could not be improved with this method.[175] The synthetic accessibility of asymmetric 1,2,4,5-tetrazines has proven to be more difficult. First attempts included thermal decomposition of 3-azido-s-triazoles to the corresponding tetrazines[176] and the transformation of *C*-phenyl-*N*-guanyl-*N'*-5-tetrazolylformazan into 3-bromo-6-phenyl-1,2,4,5-tetrazine by treatment with bromine in acetic acid (Scheme I.4).[168] These conditions are not only quite hazardous, but the reactions in the latter case reactions have to be performed behind a blast shield due to the risk of detonation.[177]

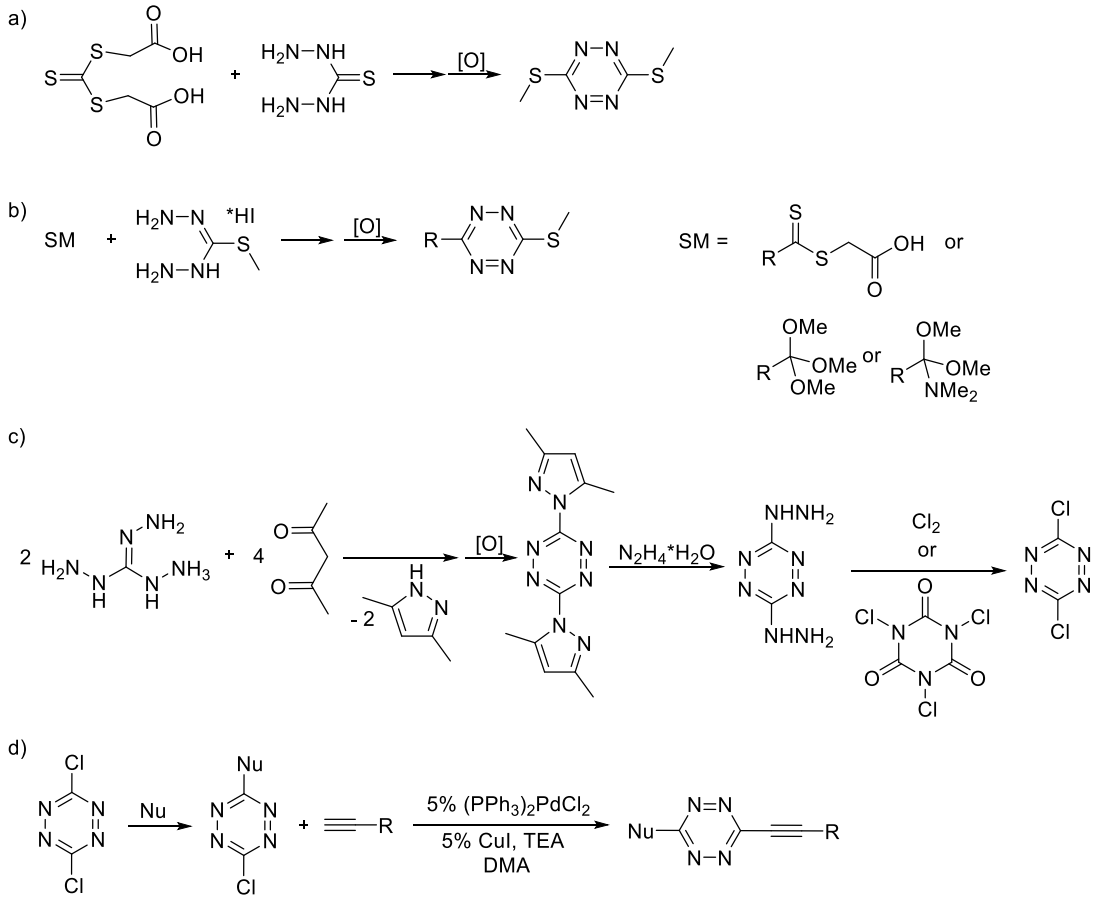


Scheme I.4: First approaches to synthetically access asymmetric 1,2,4,5-tetrazines from a) 3-azido-s-triazoles b) *C*-phenyl-*N*-guanyl-*N'*-5-tetrazolylformazan.

More promising methods modified the original Pinner synthesis of symmetric tetrazines. 3-Alkyl(or hydro)-6-Aryl-tetrazines were obtained from the reaction between benzimidates, different substituted amidines and hydrazine hydrate (Scheme I.2a,b).[167, 178] The yield of asymmetric tetrazines in these reactions is strongly influenced by the size and steric hindrance of the alkyl substituted amidines and the reactivity of alkyl imidates.[167]

Another approach applies the functionalization of 3-*S*-methylthio-[177, 179] or 3,6-bis(*S*-methylthio)-1,2,4,5-tetrazine[180] via nucleophilic substitution of the *S*-methylthio moieties. Thiocarbonohydrazide acts as precursor and can be reacted with bis(carboxymethyl) trithiocarbonate (Scheme I.5a),[181] or in the methylated form with (phenyl-carbonothioyl)sulfanyl acetic acid,[177, 179] orthoesters[182] or amide acetals[182] to obtain the thiomethylated tetrazines (Scheme I.5b). Hiskey expanded the synthetic procedures of tetrazines with a new method to prepare bisamino- and bishydrazino-1,2,4,5-tetrazines[166], which can be further modified with chlorine[183] or trichloroisocyanuric acid[184] for a more convenient handling (Scheme I.5c). Triaminoguanidine and 2,4-pentanedione react in aqueous

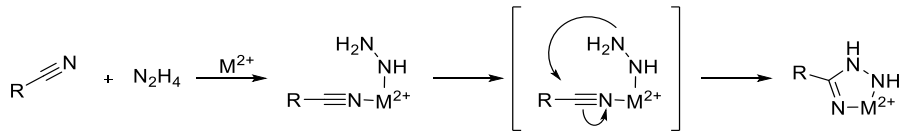
media to generate 3,6-bis(3,5-dimethylpyrazolyl)-1,2,4,5-tetrazine after oxidation with nitrogen dioxide.[166] Through the selective displacement of one chlorine atom in 3,6-dichloro-1,2,4,5-tetrazines by different nucleophiles, diverse tetrazines modified in position 6 could be produced. These tetrazines in turn find application in cross-coupling reactions with terminal alkynes, using Sonogashira conditions (Scheme I.5d).[185] 6-Morpholino substituted 3-*S*-methylthio-s-tetrazines are suitable starting materials for Suzuki cross-coupling reactions with boronic acids or Stille cross-couplings with organostannane derivatives of aromatic ring systems.[186]



Scheme I.5: Various synthetic approaches towards tetrazines for further asymmetric modification.
a) Synthesis 3,6-bis(*S*-methylthio)- or b) 3-*S*-methylthio-subsituted tetrazines as precursors for asymmetric modification by nucleophilic substitution of the *S*-methylthio moieties. c) Preparation of bischloro-1,2,4,5-tetrazines, that can be transformed d) by attack of a nucleophile and subsequent cross-coupling reactions.

Most of the aforementioned methods produce 3,6-substituted 1,2,4,5-tetrazines with one moiety donating electron density to the aromatic heterocycle as a result of the nucleophilic substitution reactions used for their synthesis. This leads to a loss in reactivity of the tetrazine towards dienophiles. Only synthetic procedures applying Pinner synthesis lead to tetrazines with alkyl or aromatic substituents. Nevertheless, this method has proven especially difficult in accessing asymmetric tetrazines, due to a statistical mixture of symmetric and asymmetric products. Devaraj and colleagues found that the addition of transition metal compounds during the synthesis of tetrazines from nitriles and hydrazine improves the yields of known symmetric and asymmetric tetrazines, making novel tetrazines synthetically accessible.[187] The transition metal ions are thought to coordinate the nitrile to promote amidrazone formation

(Scheme I.6), but the mechanism is not completely understood. Zinc(II) salts are thought to activate less reactive alkyl nitriles, while Ni(II) salts give better yields with aromatic nitriles.



Scheme I.6: Possible mechanism for transition metal catalyzed tetrazine formation

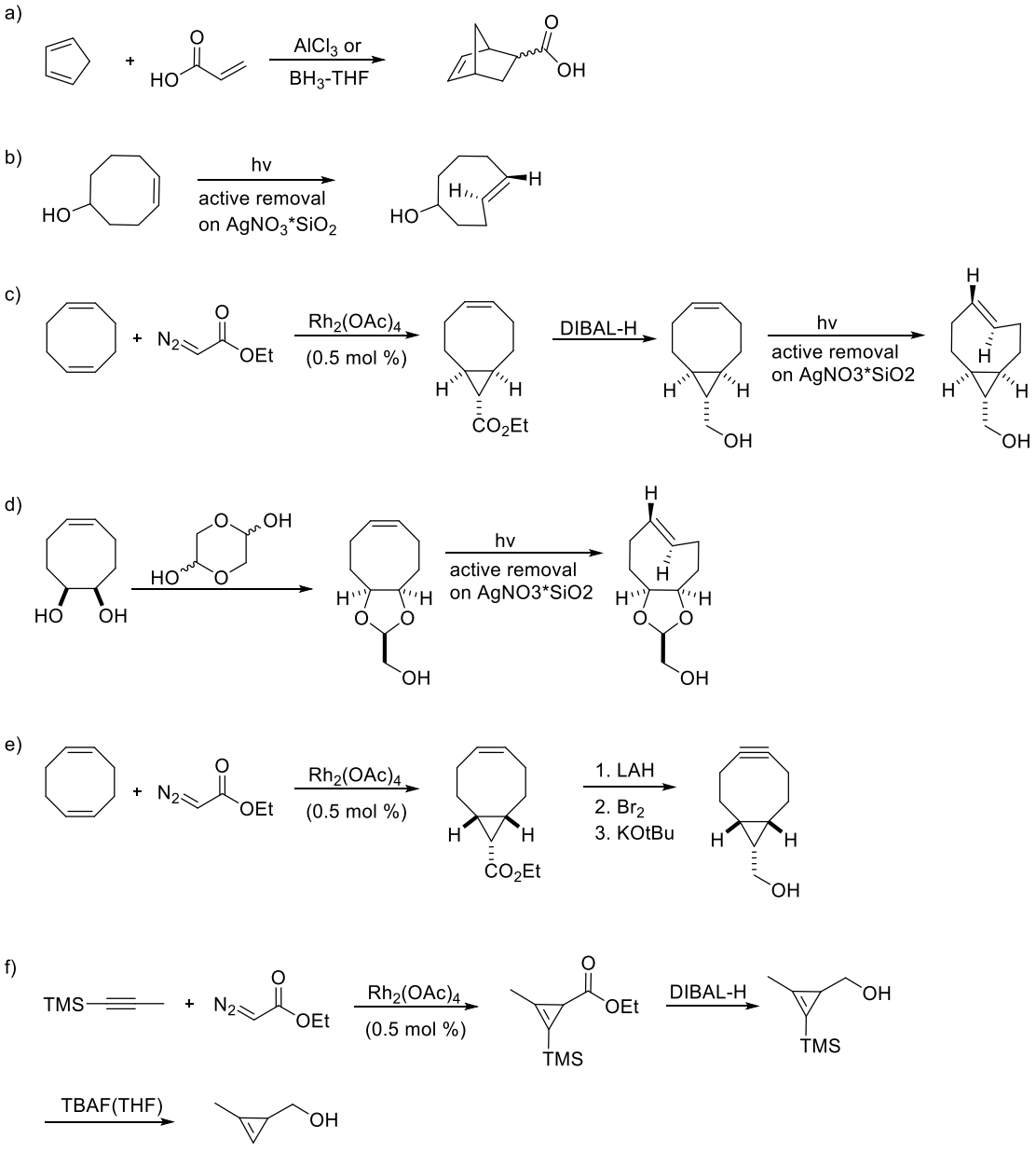
Regardless of the challenges to synthesize asymmetric tetrazines that still remain, via this approach a variety of tetrazine probes for fluorescence and PET imaging have been synthesized.[188]

I.4.3 Synthesis strategies for dienophiles

The introduction of strain energy is a very important feature in designing and tuning reactivity for bioorthogonal reactions and has found widespread application in organic synthesis. Cycloalkenes and –alkynes are highly reactive dienophiles and dipolarophiles in bioorthogonal reactions, which lead to the development of a number of different strained alkenes and alkynes for various biological applications.[119]

The aforementioned alkene and alkyne derivates (norbornene, cyclopropene, trans-cyclooctene and bicyclo[6.1.0]nonyne) are the most widely used dienophiles for iEDDAC reactions with tetrazines, due to their small size (Nor and Cp) and/or their fast reaction rates (BCN, TCO). Norbornene derivatives (Scheme I.7a) can be easily produced from cyclopentadiene and alkenes, bearing electron-drawing groups, utilizing a Diels-Alder reaction with normal electron demand, typically facilitated by Lewis acids (e.g. a borane-THF complex).[189] The synthesis of Cp, TCO and BCN derivatives is a bit more complex. Fox and coworkers published a gram-scale synthesis of TCO (Scheme I.7b) by photoisomerization of *cis*-cyclooctene at 254 nm in continuous flow through a silver nitrate modified silica gel column. The silver(I) ions selectively coordinate the *trans*-isomer, while the *cis*-isomer eluates from the column and can be recycled. Active removal of the *trans*-isomer is achieved by extraction of the silica gel.[190] The generation of bicyclic TCO derivatives further enhances the ring strain. sTCO (Scheme I.7c) is synthesized by rhodium(II)-catalyzed cyclopropanation of cycloocta-1,5-diene from diazoacetate, followed by photoisomerization of the double bond to the *trans*-isomer.[147g] The more stable dTCO bears a dioxolane ring fused to TCO. Cycloocta-1,5-diene is oxidized to the corresponding cyclooct-5-ene-1,2-diol, which is then condensed with glycolaldehyde dimer to the dioxolane. Finally, the obtained bicycle is photoisomerized to the reactive *trans*-isomer (Scheme I.7d).[162] BCN (Scheme I.7e) was developed for application as 1,3-dipolarophile in a strain promoted click reaction by Delft and coworkers. The cyclopropane ring is fused to cycloocta-1,5-diene by rhodium(II) catalysis, followed by building of the alkyne moiety by bromination and subsequent elimination.[159] 1,3-disubstituted cyclopropenes were reported in 2012 by Devaraj and coworkers as small dienophiles for iEDDAC. The cyclopropene ring is formed by a [2+1]-cycloaddition between ethyl diazoacetate and 1-(trimethylsilyl)-propyne using rhodium(II) acetate as the catalyst

(Scheme I.7f).[191] In the second step the obtained cyclopropene-bearing ester is reduced to an alcohol for further functionalization, as the carboxyl substitution displayed slower kinetics due to its electron-withdrawing properties. The trimethylsilyl moiety is introduced to reduce volatility of the products for more comfortable handling and can be easily removed by a tributylammonium fluoride (TBAF) solution in tetrahydrofuran.[192]



Scheme I.7: Synthetic routes towards strained dienophiles. a) Nor, b) TCO, c) sTCO, d) dTCO, e) BCN and f) Cp.

I.4.4 Biological applications

The fast kinetics of the Diels-Alder reaction with inverse electron demand between tetrazines and various dienophiles and their stability under physiological conditions makes it suitable to address diverse biological problems. In 2008, two independent groups around Joseph Fox[147a] and Scott Hilderbrand[149] demonstrated the new bioorthogonal bioconjugation reaction by labeling proteins *in vitro* and *in vivo*. Fox and coworkers illustrated the tetrazine ligation between trans-cyclooctene (TCO) and 3,6-dipyridyl-1,2,4,5-tetrazine to be fast, quantitative, occurring under physiological conditions and producing only nitrogen as a non-toxic byproduct, fulfilling all the criteria for a successful, bioorthogonal reaction. To show successful labeling in context of a biomolecule, they functionalized thioredoxin via a cysteine-reactive maleiimide with a TCO moiety. After subsequent labeling *in vitro* with a 3,6-dipyridyl-1,2,4,5-tetrazine derivate, ESI-MS confirmed the dihydropyridazine conjugation product.[147a] Hilderbrand and coworkers used the iEDDAC between a norbornene-antibody-conjugate and fluorophore-modified 3-(4-benzylamino)-1,2,4,5-tetrazine to visualize a receptor on the cell surface of SKBR3 human breast cancer cells after pretargeting of the antibody via fluorescence imaging.[149] Since these initial publications 1,2,4,5-tetrazines and their use in Diels-Alder reactions with inverse electron-demand with strained alkenes and alkynes received a great deal of new attention and has been applied to all kinds of biomolecules for fluorescence and positron emission tomography (PET) imaging of live cells and live animals. Modified sugars were used to study glycosylation of cell surfaces and both DNA and RNA were labeled via tetrazine ligation. A new aspect in the last years was the used of iEDDAC as reaction to activate enzyme function by chemically decaging masked amino acids or release of drugs from masked prodrugs. These applications will be discussed in further detail in the next chapters.

I.4.4.1 Fluorescence labeling

A further improvement to the use of fluorescent proteins and primary antibody-fluorophore-conjugates for fluorescence imaging was the implementation of the bioorthogonal iEDDAC reaction to allow multistep labeling. Receptors, for example the HER2/neu receptor on the surface of human breast cancer cells or the EGFR receptor on a lung cancer cell line, were pretargeted via an antibody, equipped with a norbornene or a TCO moiety. In a second step these were labeled with tetrazine-bound fluorophores.[149, 193] Tetrazine-BODIPY (borondipyrromethene) conjugates were found to display quenched fluorescence, which can be activated by an iEDDAC reaction with a strained dienophile. Fluorescence was observed to be 15 to 20-fold higher in the ligation product. These dyes were applied to the visualization of microtubules in live kangaroo rat kidney cells pretargeted by taxol, modified with TCO.[194] The introduction of a rigid linker into tetrazine-BODIPY or tetrazine-coumarin dyes further quenched the fluorescence up to a 400-1600- or 4000-fold, respectively.[188a, 188d, 195] Advantages of such dyes are improved signal-to-noise ratios due to a decreased background fluorescence, which renders washing steps unnecessary. This could be observed imaging the distribution of phospholipids in human breast cancer cells, using a cyclopropene tagged phospholipid, with a tetrazine-BODIPY FL (BODIPY exhibiting green fluorescence) conjugate. Cyclopropene-tetrazine ligation exhibits slower reaction rates, but Cp is small in

size and displays a greater stability than TCO, which loses reactivity due to isomerization to the unreactive *cis*-cyclooctene.[196]

An alternative strategy involves the application of lipoic acid labeling to covalently attach a derivative of lipoic acid to the protein of interest and subsequently label it via iEDDAC. Ting and coworkers attached a TCO modified short fatty acid to receptor proteins on the cell surface, as well as intracellular proteins like actin and vimentin and visualized them via chemoselective labeling with tetrazine dyes.[197] Wombacher and coworkers used a similar two-step labeling approach to introduce norbornene as well as tetrazine moieties onto proteins in cell lysate and on the cell surface in living cells and react them with fluorophores bearing the corresponding moiety for an iEDDAC reaction.[33-34]

In the last few years, Chin and workers, as well as other groups have developed various PylRS variants to incorporate analogues of lysine bearing norbornene (NorK), cyclopropene (CpK), cyclooctyne (SCOK), bicycle[6.1.0]nonyne (BCNK) and different *trans*-cyclooctyne (TCOK, TCOK*) moieties into proteins via GCE in *E. coli* and mammalian cells (Figure I.13).[58a, 147c, 198] These strained alkene- and alkyne-containing uAAs are chemoselectively labeled with tetrazine-fluorophore conjugates on proteins and in live cells in an iEDDAC reaction with reaction rates in the range of 10-10^4 $M^{-1}s^{-1}$. This combination of GCE and iEDDAC allowed the expression and labeling of proteins in *E. coli* and mammalian cells both intracellularly and on the cell-surface. An optimized PylRS system for the incorporation of BCNK and TCOK together with a modified labeling protocol, made it possible to visualize EGFR on the surface of HEK cells and intracellular protein targets actin and vimentin in live COS-7 cells with a tetrazine-SiR dye using super-resolution microscopy.[58b]

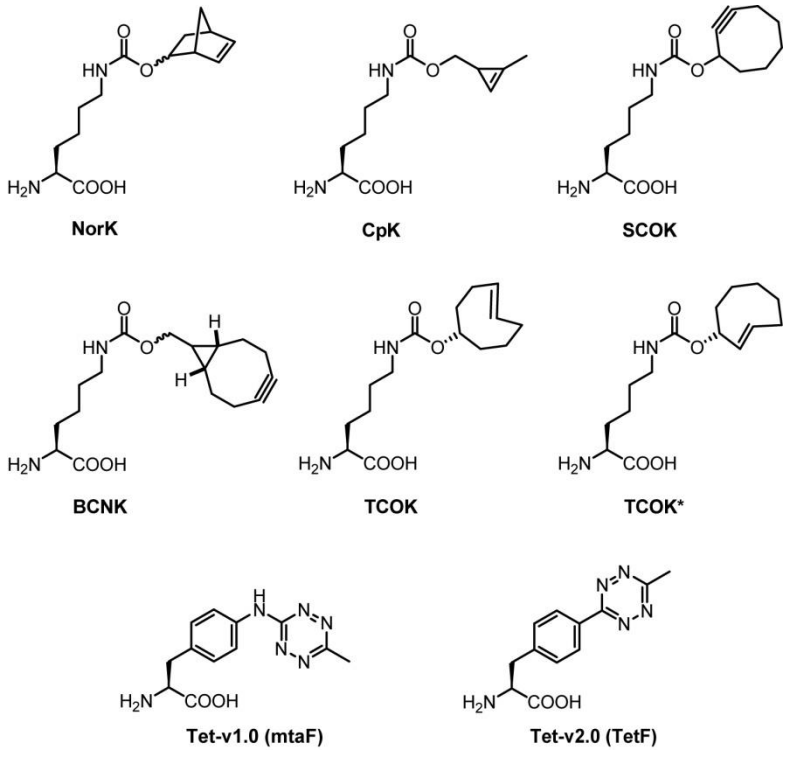


Figure I.13: Various uAAs for iEDDAC labeling bearing strained alkenes/alkynes or tetrazine moieties.

So far, no PylRS/$tRNA_{CUA}$ pair accepts unnatural amino acids bearing tetrazine functionalities as substrates. Two different tetrazine amino acids (Figure I.13) were incorporated into GFP expressed in *E. coli* using evolved TyrRS/$tRNA_{CUA}$ pairs from *M. jannaschii* and labeled with sTCO-containing dyes *in vitro* and *in vivo.* The tetrazine amino acid developed by Seitchik *et al.* suffered from slower reaction rates, due to the amino group directly attached to the tetrazine moiety. Incorporation of tetrazine amino acids at position 150 in GFP was observed to have a quenching influence on its natural fluorescence, which was regained by labeling with TCO.[56b] This can also be used to determine reaction kinetics between the reaction partners directly on the protein. Mehl and coworkers later reported the synthesis of a more rigid tetrazine amino acid without a donor substituent, which therefore displayed better reaction rates.[151b]

I.4.4.2 Controlling enzyme activity via iEDDA cycloadditions in live cells

IEDDA cycloadditions have not only been used to label protein site-specifically for fluorescence imaging, but the bioorthogonal tetrazine ligation can also be used to selectively manipulate enzyme activity in living cells. Chin and coworkers developed a method based on genetic code expansion and bioorthogonal reactions called bioorthogonal ligand tethering (BOLT).[199]

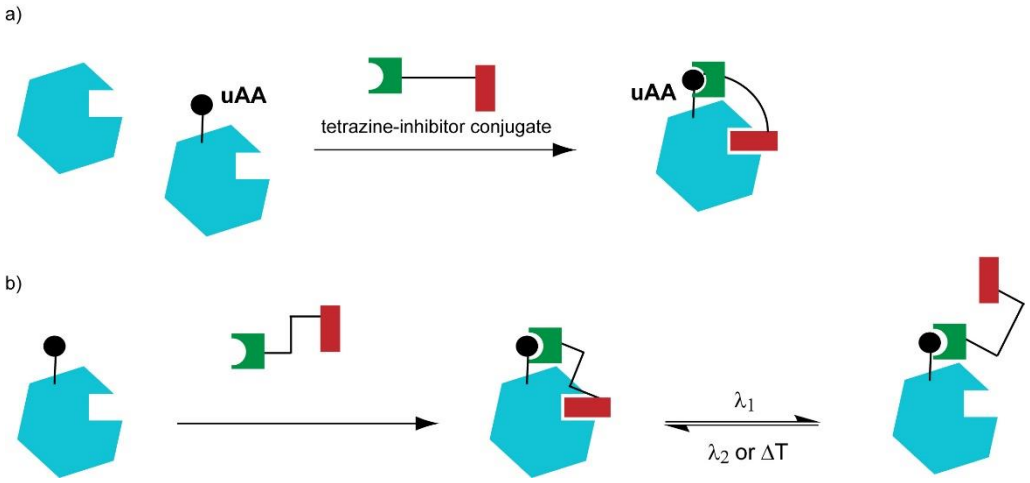


Figure I.14: Enzyme activity is controlled and manipulated via bioorthogonal ligand tethering (BOLT) a) proximity-driven inhibition of proteins via a low affinity inhibitor b) reversible manipulation of the protein via a photo-switchable moiety in the linker[199]

This approach allowed the selective inhibition of mitogen activated protein kinases MEK1 over MEK2, which has not been possible using small molecule inhibitors, as MEK1 and MEK2 share 82 % sequence identity.[200] MEK1 and MEK2 are key players in the MAP kinase signaling pathway involved in regulation of eukaryotic growth and development.[201] BCNK or CpK were incorporated close to the active site of MEK1 in response to the amber codon and tetrazine conjugates of small molecule inhibitors of MEK1 covalently linked to MEK1 via iEDDAC. This lead to the rapid inhibition (iBOLT) of the tethered MEK1, while endogenous MEK1 or MEK2 in the cell are not affected (Figure I.14a). [202] To modulate the activity of the target protein by light-induced inhibitor binding or unbinding, an photo-isomerizable moiety was introduced into the linker of the tetrazine-inhibitor conjugate.[199] Upon tethering to MEK1

light of different wavelength were used to switch between *cis* and *trans* conformation of the photoswitch, enabling Chin and coworkers to control MEK1 activity in live mammalian cells by light (Figure I.14b).[199]

I.4.4.3 Radionuclide labeling for PET Imaging

The tetrazine ligation with its fast kinetics is an ideal choice for molecular probes bearing positron-emitting radionuclides (^{11}C, ^{13}N, ^{15}O and ^{18}F) for positron emission tomography (PET) or nuclides for single photon emission computed tomography/computed tomography (SPECT/CT), which are otherwise synthetically difficult to access and have a short half-life like fluorine-18 ($t_{1/2} \approx 110$ min) and carbon-11 ($t_{1/2} \approx 20$ min). Especially, ^{18}F is complicated to introduce into complex molecules due to the poor nucleophilicity of the fluoride. Synthetic strategies to endow TCO,[203] sTCO,[204] norbornene[205] as well as tetrazine[206] scaffolds with ^{18}F or ^{11}C were reported by different groups.[207] PET imaging *in vivo* and live animals was performed with tetrazine modified Poly(ADP-Ribose)-Polymerase 1 (PARP1) inhibitors, which were labeled with ^{18}F-bearing dienophiles in PARP1-overexpressing cancer cells or PARP1 mouse models.[208] Tetrazine conjugates of a cyclic RGD-peptide and the vascular endothelial growth factor (VEGF)-SH protein were modified with ^{18}F-TCO for microPET imaging of tumor–bearing mice.[209] Robillard and coworkers were the first to apply the iEDDAC for pretargeted tumor imaging in living mice. They developed a method, where a tumor marker pretargeted with a monoclonal antibody (mAb)-TCO conjugate, followed by modification of the radiolabeled tumor-bound mAb via tetrazine ligation. In this case an antigen which is overexpressed in a number of different cancer types was pretargeted in living xenografts with colon cancer. Tetrazine-DOTA complexing ^{111}In was then administered before SPECT/CT imaging. This pretargeted two-step labeling method allowed the application of a higher (therapeutic) radiation dose compared to directly labeled mAbs.[210] Lewis and coworkers expanded this pretargeting approach to modularly synthesize probes bearing different radionuclides for PET and SPECT imaging of HER2 expressing and colorectal cancer cells.[211] In addition to antibodies, supramolecular nanoparticles[212] have been used to pretarget universal tumor phenotypes as well as a pH (Low) Insertion Peptide (pHILIP) that directs radiolabeled liposomes for the delivery of radiodiagnostics and radiotherapeutics to tumor tissue.[213]

I.4.4.4 Labeling of nucleic acids

Nucleic acids like DNA or RNA are the fundamental basis of all kingdoms of life and play an essential role for storing and transferring genetic information. To understand their function concerning catalysis or dynamics in live cells is of great interest.[214] Bioorthogonal modification would enable labeling, visualization and tracking of nucleic acids in living systems. Early attempts altered nucleobases for immunohistochemical staining.[215] Other approaches to transform nucleobases were based on *N*-hydroxysuccinimide (NHS)-ester chemistry before the application of bioorthogonal reactions such as CuAAC. These approaches suffer from limitations, either the necessity of a high excess of label or the non-applicability *in vivo* due to high concentrations of Cu(I).[216] To overcome these limitations,

iEDDAC was applied for nucleic acid labeling. Therefore, dienophile-modified nucleotides were used in solid-phase synthesis to site-specifically introduce them into DNA and RNA oligonucleotides, which were post-synthetically labeled with tetrazine-fluorophore conjugates.[161a, 214, 217] Kath-Schorr and coworkers were able to label *in vitro* synthesized norbornene-modified siRNA with a tetrazine-dye conjugate after transfection into mammalian cells.[218]

Apart from chemical methods, DNA and RNA modified with dienophiles could be prepared by *in vitro* transcription. One approach uses a norbornene-modified guanosine initiator nucleotide[214], which is attached on the 5'-end of RNA by a T7 RNA polymerase recognizing a defined promoter, which in turn has a G nucleotide as +1. Base (Figure I.15a).[219] Nucleotides have also been modified with a TCO moiety at position 5 on the pyrimidine ring, resulting in a cytidine triphosphate (CTP) for labeling of RNA[217b] or a thymidine triphosphate (TTP) analogue for DNA (Figure I.15a).[220] The group of Kath-Schorr was able to site-specifically introduce a norbornene[221] and a cyclopropene moiety[222] into RNA applying unnatural base pairing between an unnatural base pair developed by Romesberg and coworkers.[223] This enabled labeling with tetrazine-fluorophore or tetrazine-biotin conjugates post transcription. Wang and coworkers performed *in vitro* cell imaging using a known aptamer modified via incorporation of TCO-TTP in HeLa cells.[224]

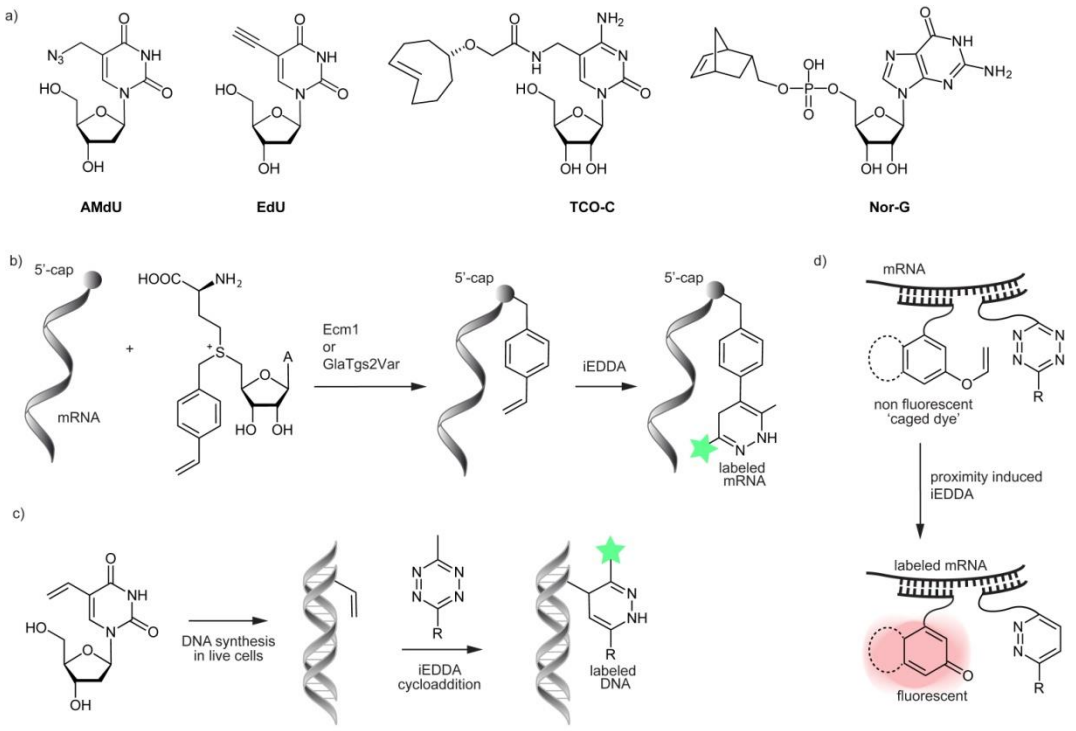


Figure I.15: Labeling of nucleic acids.
a) different dienophile-modified DNA and RNA nucleobases (AMdU = Azidomethyldesoxyuridine, EdU = 5-ethynyl-2'-desoxyuridine, TCO-C = TCO modified cytidine) and nucleotides (Nor-G = norbornene modified guanosine initiator) b) site-specific labeling of the 5'-cap of mRNA c) 5-vinyl-2'-desoxyuridine (VdU) probe for DNA labeling in live cells d) template-directed mRNA detection with a fluorogenic probe.

The Rentmeister group reported a method to site-specifically modify the 5'-cap typical of eukaryotic mRNA with bioorthogonal handles for two-step labeling (Figure I.15b). Highly promiscuous methyltransferase Ecm1 from *Encephalitozoon cuniculi* or an engineered

trimethylguanosine synthase variant (GlaTgs2-Var1) from *Giardia lamblia* transfer 4-vinylbenzyl and other moieties from a modified S-adenosylmethionine analogue to different positions on the 5′-cap structure of mRNA.[225] The labeling was performed in cell lysate from eukaroytes and might be used *in vivo* in the future to modify mRNA. Combination of the two enzymes, which display a different regioselectivity allows the modification with orthogonal bioorthogonal handles and subsequent labeling for FRET studies or with different reporters.[226] To study cellular DNA, Rieder and Luedtke developed a small probe (5-vinyl-2'-deoxyuridine, VdU), which was metabolically incorporated into genomic DNA of HeLa cells and visualized by tetrazine ligation with a fluorescent dipyridyltetrazine derivative (Figure I.15c).[227] Even though the reaction between terminal alkenes and tetrazines displays very slow reaction rates compared to strained alkenes, the labeling was still faster than a strain-promoted alkyne-nitrone cycloaddition (SPANC). Application of orthogonal probes VdU and 5-ethynyl-2'-deoxyuridine (EdU) allowed time-resolved, multicolor labeling of DNA synthesis[227b].

Devaraj and coworkers applied the iEDDAC reaction for template-directed DNA ligation (Figure I.15d). They modified oligonucleotides with tetrazine quenched fluorophore conjugates and cyclopropene groups, which can form a highly fluorescent product upon hybridization of the oligonucleotides to a complementary DNA template strand.[228] An optimized strategy uses oligonucleotides bearing a 7-azabenzonorbornadiene derivative as dienophile, which undergoes a Retro Diels-Alder reaction after initial iEDDAC ligation, leading to a transfer reaction between dienophile and fluorogenic tetrazine. This enabled reaction turnover and amplification of the fluorogenic signal, which was employed to detect DNA and microRNA templates in cell lysate and live cells.[229] In 2016 the same group developed oligonucleotides with fluorogenic probes in the near-infrared for the detection and imaging of mRNA *in vitro* and in live cells. The fluorescence is activated by DNA or RNA template directed tetrazine decaging of a vinyl ether moiety on the probe.[230]

I.4.4.5 iEDDAC for labeling of glycans and lipids

Glycosylation is the enzymatic attachment of glycans (saccharides) to various macromolecules that takes place post-translationally in all kingdoms of life and is involved in different important cellular processes such as signaling, protein folding and stability, as well as exchange of information between cells. Disruption of glycosylation is involved in the development of various diseases like cancer, which makes it an interesting target to study in live cells or animals.

Metabolic labeling is the most common method used for visualization of glycosylation on the cell surface. Affinity-based strategies using lectins and antibodies suffer from poor tissue penetrance and low affinity.[231] Therefore the group of C. Bertozzi developed a method which introduces an unnatural sugar derivative modified with a bioorthogonal handle via metabolic uptake into the glycosylation pattern. After reaction with the bioorthogonal counterpart, glycans can be imaged.[232] Several groups were able to modify a number of different sugar molecules with different dienophiles to employ in iEDDAC reactions with tetrazines. Prescher and coworkers were the first to introduce a 1,3-substituted cyclopropene moiety via

amide linkage into neuramic acid and visualize cell surface glycosylation of live Jurkat cells after metabolic uptake.[192] The 1,3-Cp group could also be incorporated into mannosamine derivatives to image human breast cancer and colon cancer cell lines .[233] The Wittmann group reported a *N*-acetyl-mannosamine sugar modified with 1,3-Cp via a carbamate linkage, which was reported to display faster reaction rates in a iEDDAC with tetrazines to visualize the cell surface of HEK 293T cells (Figure I.16a).[234] Other bioorthogonal handles attached to mannosamine derivatives for use in metabolic glycoengineering include norbornene moieties[235] and terminal alkenes.[236] A two-step labeling approach was reported by Brindle and coworkers to image cell surface glycosylation in live mice. Their bifunctional probes, carrying a TCO and a functionality for SPAAC or Staudinger ligation, were reacted with an azido-labeled galactosamine derivative, incorporated into the cell surface via metabolic uptake. Administering a tetrazine dye allowed the read-out via fluorescence.[237] The Wittmann group developed an approach to visualize the protein-specific glycosylation in live cells (Figure I.16b). *N*-acetylglucosamine bearing a methylcyclopropene tag was metabolically incorporated into the glycome including the protein of interest, which is expressed as a GFP fusion. After iEDDAC with a tetrazine-fluorophore derivative, FRET can be measured between GFP and the fluorophore and used as a read-out for the glycosylation state of the protein in living cells.[238]

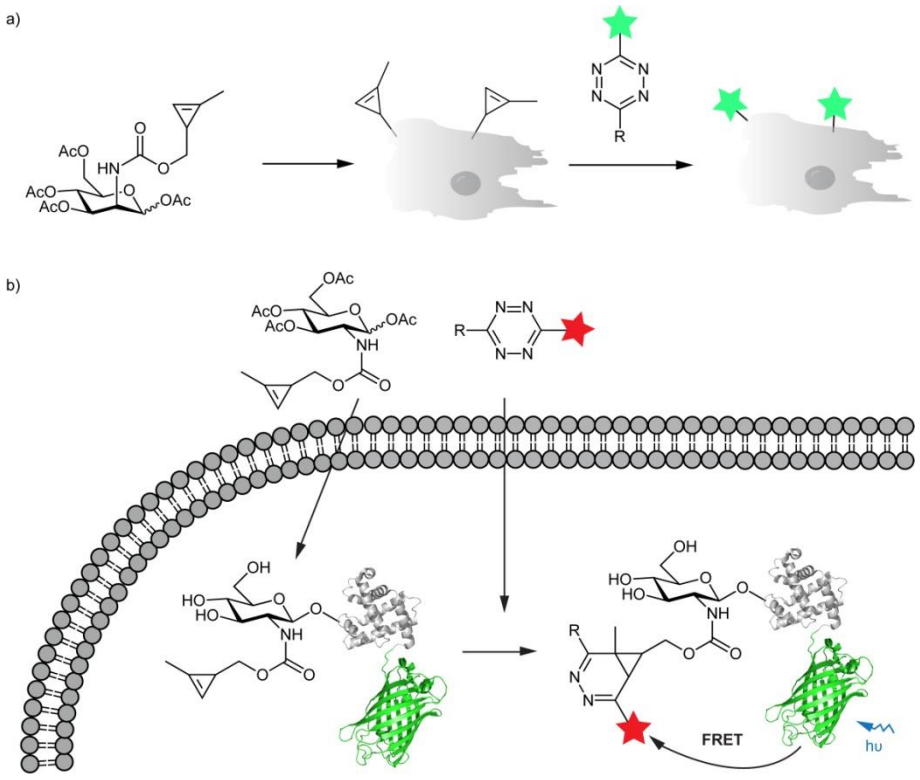


Figure I.16: iEDDAC reactions for the imaging of glycosylation.
a) Labeling of cell surface glycosylation after metabolic uptake of Cp-tagged sugar analogs b) Visualization of glycosylation of a protein of interest (POI) inside living cells through metabolic incorporation and iEDDAC with tetrazine-fluorophore conjugates

Very recently, the same group reported the metabolic labeling of HEK 293T cell surface with mannosamine derivatives bearing three different functionalities for triple orthogonal labeling with iEDDAC, SPAAC and photo-Click reactions.[239] The systemic application of a BCN-

functionalized sialic acid allowed Bertozzi and coworkers to follow sialylation during zebrafish embryogenesis by *in vivo* fluorescence imaging with sub-cellular resolution.[231a] In contrast to the methods using metabolic incorporation described above, Winssinger and coworkers reported an approach to site-specifically and artificially glycosylate a protein. They introduced different glycans via tetrazine ligation into target proteins in *E. coli* via genetically incorporated uAAs TCOK or BCNK.[240]

Devaraj and coworkers reported the synthesis of a 1,3-Cp bearing phospholipid to image the distribution of phospholipids in live cells via fluorogenic tetrazines.[196] The Schepartz group employed the incorporation of a TCO-modified ceramide lipid into the Golgi apparatus to visualize its structure and dynamics by super resolution microscopy in live cells. Therefore, they administered a cell permeable near-infrared dye-tetrazine derivative, which reacted in an iEDDAC with the TCO moiety and allowed prolonged live-cell imaging.[241]

I.4.4.6 Orthogonal bioorthogonal reactions

The application of bioorthogonal reactions *in vivo* has enabled to uncover information on structure and function of the labeled and manipulated biomolecules in their natural environment. However, as biomolecules are part of a multi component network in living systems and therefore need to be observed as part of those, bioorthogonal chemistries which can be applied in parallel are of great interest. The development of bioorthogonal chemistries that are orthogonal to each other is very challenging as many of the known bioorthogonal reactions rely on the same functionalities, for example azides can react in Staudinger ligations, CuAAC and SPAAC.[242] By now a number of reactions have been identified that are orthogonal to each other and have been applied *in vivo* for labeling of two different biomolecules or one biomolecule with two different functionalities for example to study protein folding via FRET.[243]

Tetrazines and strained alkenes like cyclopropenes, norbornenes and TCOs react together in an iEDDA cycloaddition, but they do not show any reactivity towards azides (Figure I.17a,b), which renders them orthogonal to CuAAC and SPAAC reactions,[161a, 220, 236-237, 244] which was employed for the double labeling of DNA,[220] sugars[234, 236] and proteins.[245] Orthogonality of the iEDDAC reaction can be modulated by electronic or steric manipulation of the reaction partners. Introduction of EWGs or EDGs into position 3 and/or 6 into tetrazines strongly influences their reactivity as described above. Strained alkenes/alkynes can also be manipulated by changes in steric or electronic properties. Hilderbrand and coworkers reported the introduction of two benzo moieties into the cyclooctyne scaffold that diminished reactivity towards 3,6-disubstituted tetrazines, but readily reacted with azides (Figure I.17b). The combination of this SPAAC reaction and the iEDDAC with TCO enabled selective, multi target imaging of co-cultured human breast cancer cells and human epidermoid carcinoma cells.[246] Kinetic studies on iEDDAC reactions by the Chin group revealed that iEDDAC reactions can be performed in parallel without displaying any interference with each other (Figure I.17c). Therefore, they incorporated a norbornene derivative of lysine and a deactivated tetrazine amino acid into calmodulin using an evolved ribosome that decodes the

amber codon and a quadruplet codon to introduce the uAAs. The amino-substituted electron-rich tetrazine reacts only sluggishly in an iEDDAC and does not display any reactivity towards the norbornene moiety which has only a moderate ring strain. The two uAAs were sequentially labeled with BCN and an electron-poor dipyridyltetrazine.[247] Bonger and coworkers introduced a boronic acid substituent in vinylogous position of the dienophile, which displayed enhanced reaction rates towards pyridyl- or dipyridyltetrazines due to coordination of the boronic acid to the nitrogen atom of the pyridyl substituent. Pyrimidyl substituents display significantly lower reaction rates, which is probably due to a lower basicity and therefore reduced coordination of the nitrogen atoms to the boron atom. This enabled modification of proteasome with norbornene- and vinylboronic acid-conjugates of different proteasomal inhibitors and sequential labeling with pyrimidyl- and pyridyl-substituted tetrazine-fluorophores respectively (Figure I.17d).[153]

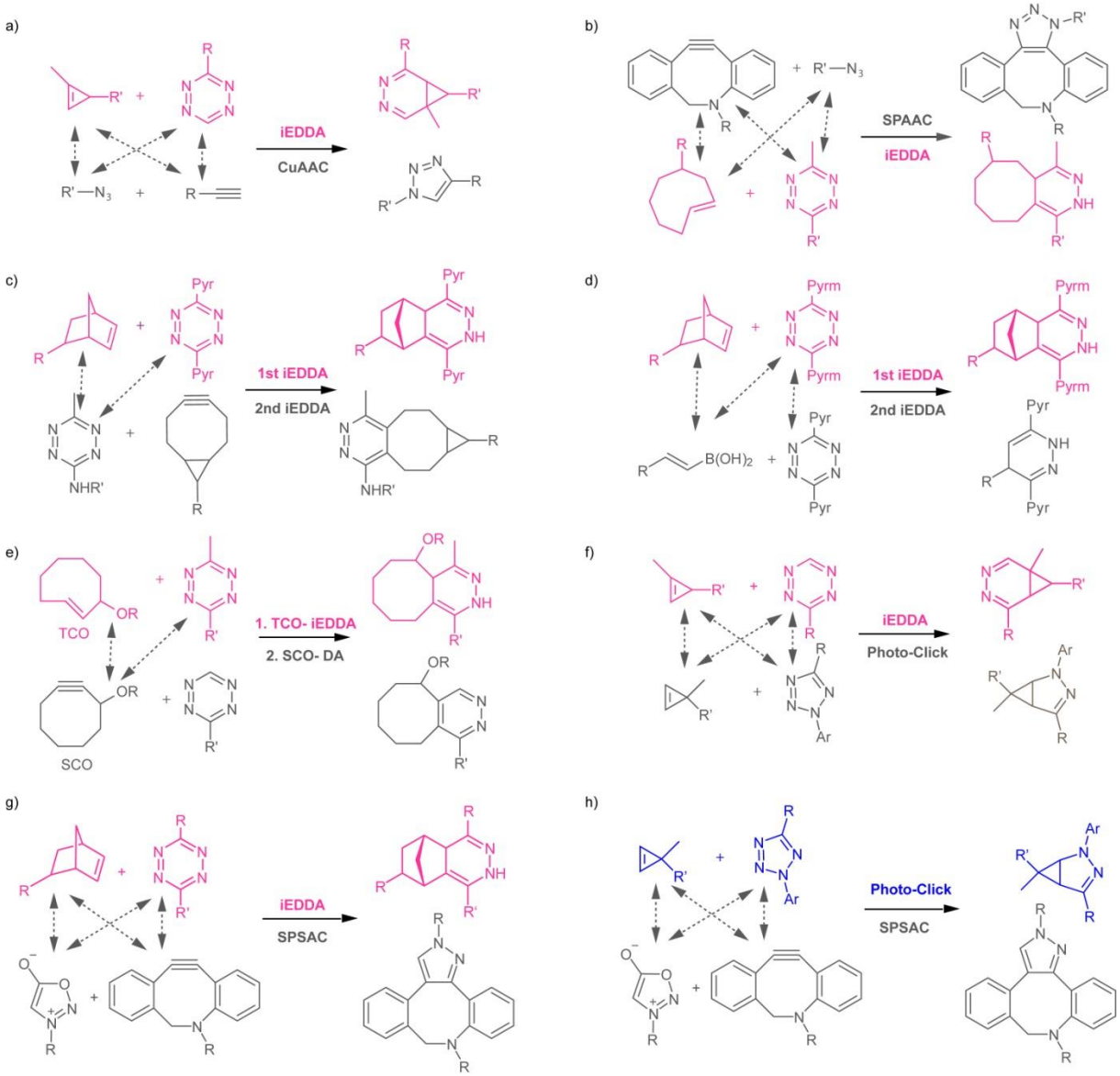


Figure I.17: Orthogonal bioorthogonal reactions.
a) iEDDAC and CuAAC b) SPAAC and iEDDAC c,d) two sequential iEDDACs e) iEDDAC with TCO and normal DA with SCO f) iEDDAC and Photo-Click g) iEDDAC and SPSAC h) Photo-Click and SPSAC

Lemke and coworkers reported a modified cyclooctyne dienophile bearing a propagylic carbamate linkage (SCO), which reacts only with monosubstituted tetrazines due to the steric hindrance of disubstituted ones, while BCN and TCO react with nearly all tetrazines (Figure I.17e). The approach was used for pulse-chase dual labeling of proteins on the cell surface and virus-like particles.[157b] To understand the underlying principles of this orthogonality the group performed DFT calculations and further experiments and found that the carbamate group renders SCO more electrophilic and leads to a more favorable interaction between its LUMO and the HOMO of the tetrazine turning the reaction into a DA cycloaddition with normal electron demand. The calculations showed that methyl-substituted tetrazines would also react in a normal DA due to the electron-donating properties of the methyl group, but here steric clashes with the carbamate moiety hinder the reaction. Stronger electron-donating substituents on the monosubstituted tetrazines could therefore enhance selectivity for SCO over TCO or BCN.[157c] As described above, the reactivity of cyclopropene is strongly influenced by the substitution in positions 1 and 3. Prescher and coworkers exploited this to develop two orthogonal bioorthogonal reactions. 1,3-disubstituted cyclopropenes react in an iEDDAC with tetrazines, while 3,3-disubstituted cyclopropenes are inert towards tetrazines and are suitable reaction partners in Photo-Click reactions (Figure I.17f). There they undergo a 1,3-dipolar cycloaddition with nitrile imines, which are generated *in situ* from tetrazoles.[147f] Utilizing computational screenings Murphy and coworkers discovered that *N*-phenyl sydnones (Figure I.17g) are suitable reaction partners in [3+2] cycloadditions with biarylazacyclooctynone (BARAC) and dibenzoazacyclooctyne (DIBAC).[248] The reaction rates were observed to be 30 times faster that the cycloaddition of sydnones with BCN reported by the Chin group.[249] This new reaction is orthogonal to iEDDAC, which they confirmed by labeling experiments of different model proteins *in vitro*. Predictions indicate that sydnones are also orthogonal to 3,3-disubstituted cyclopropenes, which would allow a labeling combination with the Photo-Click reaction (Figure I.17h).[248]

I.4.4.7 Tetrazine decaging reactions

Traditionally, bioorthogonal reactions between two partners form new covalent bonds and have mainly been used to bioconjugate macromolecules *in vitro* and *in vivo*.[39b] In recent years another approach employing bond cleavage reactions has emerged.[65d] The concept is well known using UV light as a trigger to photo-release of deprotected, so called “caged” functionalities, but suffers from the phototoxicity of UV light and difficulty to penetrate tissue or animal samples.[250] Therefore, the development of decaging reactions employing small molecules as triggers was promoted as an alternative. First approaches report the deprotection of allyloxycarbonyl or propargyloxycarbonyl moieties under rhodium or palladium catalysis.[251] Applying the iEDDAC reaction between tetrazines and strained alkenes/alkynes as bond cleavage reaction was first reported by Robillard and coworkers. Doxorubicin was released as a free amine from the carbamate linkage in the allylic position of TCO upon reaction with dimethyltetrazine. This proceeds via a pyridazine conjugation product, followed by the formation of an exocyclic double bond with the elimination of CO_2 and an amino-substituted doxorubicin.[252] This strategy was also employed to pretargeted tumors in mice bearing colon carcinoma xenografts via an immunoreactive CC49-TCO-Dox antibody-drug-

conjugate and lead to the observation of a high tumor uptake after iEDDAC-mediated cleavage reaction.[253] Prodrug activation of doxorubicin can also be performed by tetrazine-modified magnetic iron oxide nanoparticles, which are also fluorescently-labeled. The uptake and conversion into the chemotherapeutic drug was followed by *in vivo* fluorescence imaging of breast cancer cells.[254]

A second strategy besides the pretargeted prodrug release comprises the chemical activation of enzymes by a tetrazine triggered bond cleavage reaction (Figure I.18). Chen and coworkers therefore incorporated TCOK*, a caged mimic of lysine, site-specifically into the active site of protein replacing a natural lysine important for catalytic activity. The TCO moiety is attached via a carbamate linkage in its allylic position to the ε-amino group of lysine, which leads to deprotection to the free lysine by induction of the cleavage reaction by the addition of dimethyltetrazine (Figure I.18a). Using this approach the Chen group was able to rescue the catalytic activity of intracellular firefly luciferase in HEK293T cells[65d] and manipulate kinase activity living systems.[255] In 2016 the same group introduced a coumarin-based fluorogenic assay to analyze influence of different substituents on tetrazines on the iEDDAC decaging reaction. Electron-withdrawing groups enhance the iEDDA cycloaddition, yet sterically hinder the elimination step. The best results were observed *in vitro* and in living cells with unsymmetrical tetrazines bearing a combination of EWGs and EDGs in positions 3 and 6 (Figure I.18b).[256]

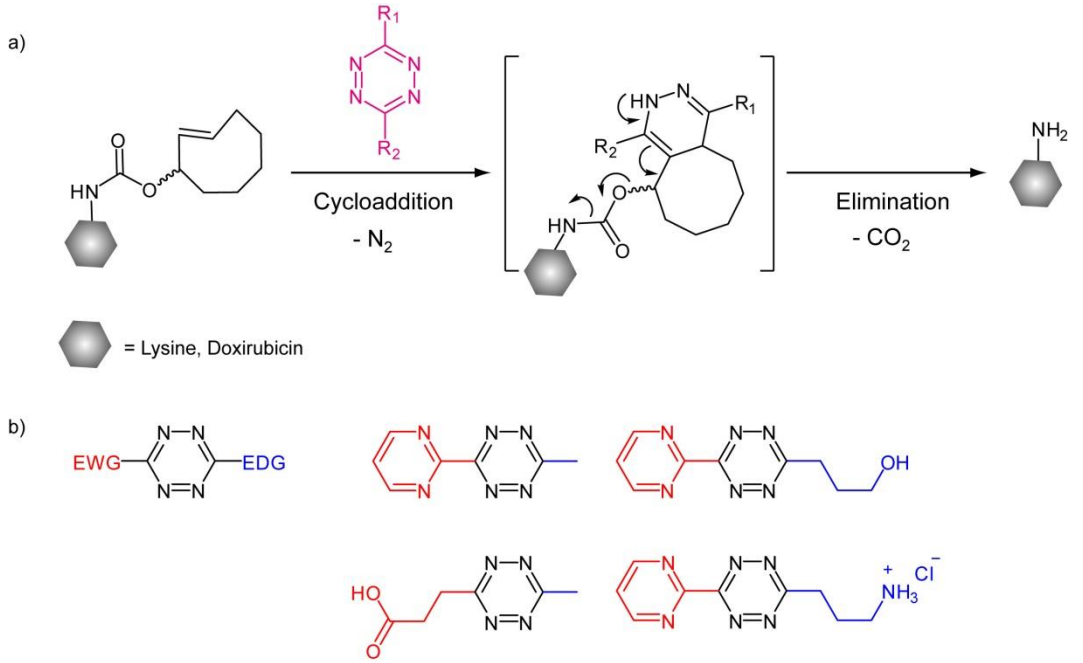


Figure I.18: a) Tetrazine decaging reaction b) Optimized tetrazines for faster decaging reactions; EWG increases kinetics of cycloaddition, while EDG enhances elimination reaction.

I.4.4.8 Application in material science

In the last few years the tetrazine ligation has also found applications in the development of new materials or the modification of materials with new functionalities. For example a covalently bonded hybrid nanocomposite was generated from single walled carbon nanotubes[257] as well as a sp^2-hybridized 2D carbon surface (highly ordered pyrolytic graphite)[258], which can act as dienophiles in an iEDDAC with tetrazine-modified gold nanoparticles (AuNPs). The AuNPs are introduced to stabilize and allow further functionalization of the nanocomposite. This reaction usually requires pre-treatment and harsh conditions, which could be evaded by application of the tetrazine ligation.[257-258] O'Reilly and coworkers inserted norbornene moieties into polystyrene polymer chains, which were reacted with homobifunctional tetrazine linkers to obtain polymer nanoparticles via single chain collapse.[259]

The modification of solid surfaces via tetrazine ligation for their potential application in electronic devices or microarrays has been reported by the Wittmann group as well as Ravoo and colleagues. The Wittmann group immobilized tetrazines via NHS-chemistry on a glass substrate to prepare synthetic and natural carbohydrate microarrays via iEDDAC for possible utilization in glycomics.[260] Ravoo and coworkers adapted the immobilization strategy for the generation of differently patterned surfaces via microcontact chemistry (μCC). With their approach they were able to site-specifically functionalize a surface with different types of carbohydrates as well as poly(methylacrylate) brushes. These were assembled by grafted polymerizations after immobilization of the initiator via tetrazine ligation.[261]

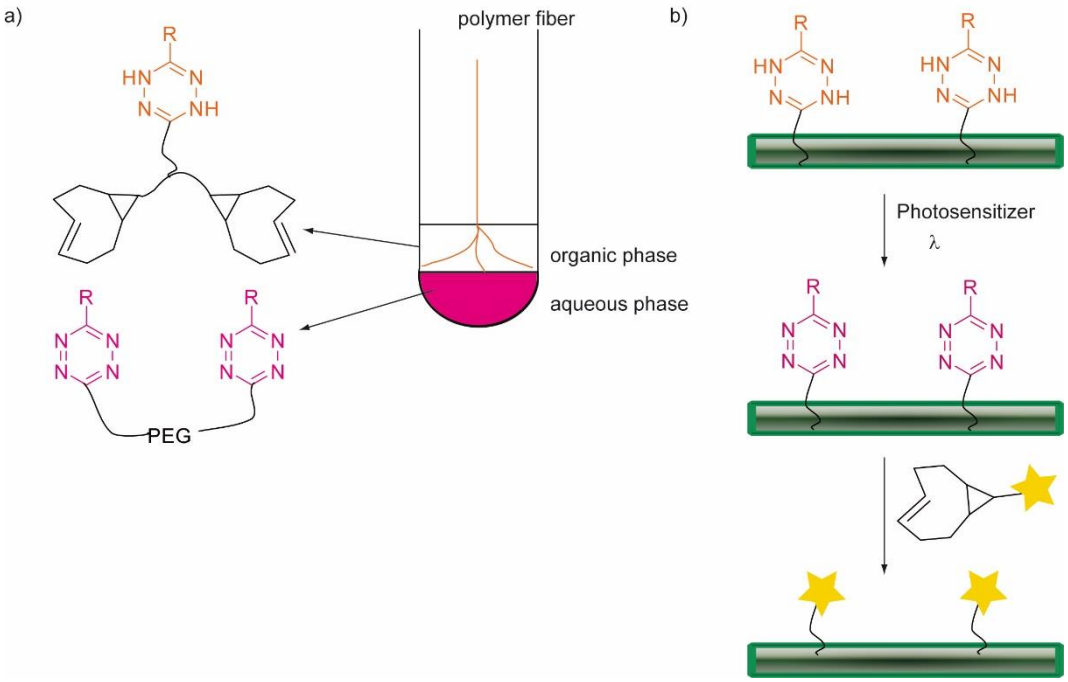


Figure I.19: Application of tetrazine ligation in polymerization reactions.
a) interfacial polymerization of two homobifunctional monomers to obtain multiblock copolymer fibers[262] b) modification with stable dihydrotetrazines enabled the triggering of another iEDDAC with photosensitizers and visible light or HRP for labeling with divers probes post-polymerization

The introduction of strained alkenes like norbornene and tetrazines into a polymer side chain allowed post-polymerization iEDDAC to modify polyesters[263] or form covalently crosslinked alginate hydrogels as delivery systems for cells or bioactive molecules.[264] Tetrazine ligation cannot only be utilized for post-polymerization modification, but also applied to generate hybrid multiblock copolymer microfibers by interfacial polymerization (Figure I.19a). Therefore, Jia and coworkers synthesized two different homobifunctional monomers, containing either TCO and a hydrophobic linker or tetrazines and PEG-linkers, which were dissolved in two immiscible solvents. The multiblock copolymer which was formed at the interface between those two solutions could be drawn into microfibers with semicrystalline structures. The introduction of a peptide into the tetrazine monomer allowed application of the resulting fibers as cell adhesive structures in a biological context.[262a] An advancement of this strategy was the attachment of dihydrotetrazines, which could be oxidized to the tetrazine by catalytic stimuli after formation of the copolymer and subsequently modified with TCO-containing molecules, such as fluorophores or peptides. The oxidation could be catalyzed by long wavelength photosensitizers like methylen blue and rose bengal or in an enzymatic manner using horseradish peroxidase under peroxide-free conditions (Figure I.19b).[262b] The oxidation to form tetrazines from dihydropyridazines can also be stimulated and controlled by utilizing electrochemical potential. In 2014 the Devaraj group reported the surface modification of a electrode with TCO conjugates of ferrocene or horseradish peroxidase by changing the redox state of immobilized dihydrotetrazine.[188c]

CHAPTER 1

Photo-induced inverse-electron demand Diels-Alder cycloaddition for the modification of proteins

The work in the following chapter has been submitted for publication as:

S.V. Mayer,[#] A. Murnauer,[#] M.-K. von Wrisberg, M.-L. Jokisch and K. Lang. Photo-induced and rapid labeling of tetrazine-bearing proteins via cyclopropenone-caged bicyclononynes. **2019**. Manuscript submitted for publication. ([#] these authors contributed equally)

The chapter is strongly based on the submitted publication, but includes also negative results.

Chapter 1: Photo-induced inverse-electron demand Diels-Alder cycloaddition for the modification of proteins

1.1 Aim

The aim of this project was to develop a light-triggered tetrazine ligation with fast reaction kinetics to confer spatiotemporal control over the bioorthogonal labeling reaction. For this, a new strained and light-activatable dienophile was designed, based on a cyclopropenone protected dibenzocyclooctyne structure (DIBO) as structural scaffold, since DIBOs are able to be decaged to the corresponding cycloalkyne at 365 nm, conditions that should be amenable to living cells. Introduction of additional ring strain by fusion of a cyclopropane ring to the DIBO scaffold (Figure 1.1) should increase reactivity towards 1,2,4,5-tetrazines, as DIBO only displays rather slow kinetics in iEDDAC with a 6-hydrogen substituted tetrazine (0.06 $M^{-1}s^{-1}$).[246] Synthesis of cyclopropenone-caged **di**(**m**ethoxy**b**enzo)-bicycl**o**[6.1.0]nonyne (photo-DMBO) compounds (Figure 1.1) was performed by Anton (Toni) Murnauer and is described in detail in his Master's Thesis.[265]

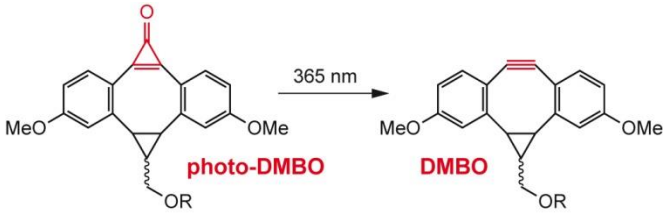


Figure 1.1: photo-DMBO scaffold

As reaction partners for photo-DBMO new lysine-based uAAs bearing tetrazine moieties were designed (Figure 1.2), since the two tetrazine uAAs known to be incorporated into proteins using GCE are derivatives of L-phenylalanine and therefore substrates of a mutant TyrRS from *M. jannaschii* (Tet-v2.0RS),[56b, 151b] which is only orthogonal in *E. coli*. To enable the expansion of tetrazine amino acids to eukaryotic cells, finding a mutant PylRS from *M. barkeri* for the incorporation of tetrazine uAAs is a crucial part of the project, since the PylRS/t RNA pair is also orthogonal in mammalian cells. Most uAAs incorporated by various PylRS mutants are lysine derivatives, which are structurally similar to the natural substrate of the wt PylRS.

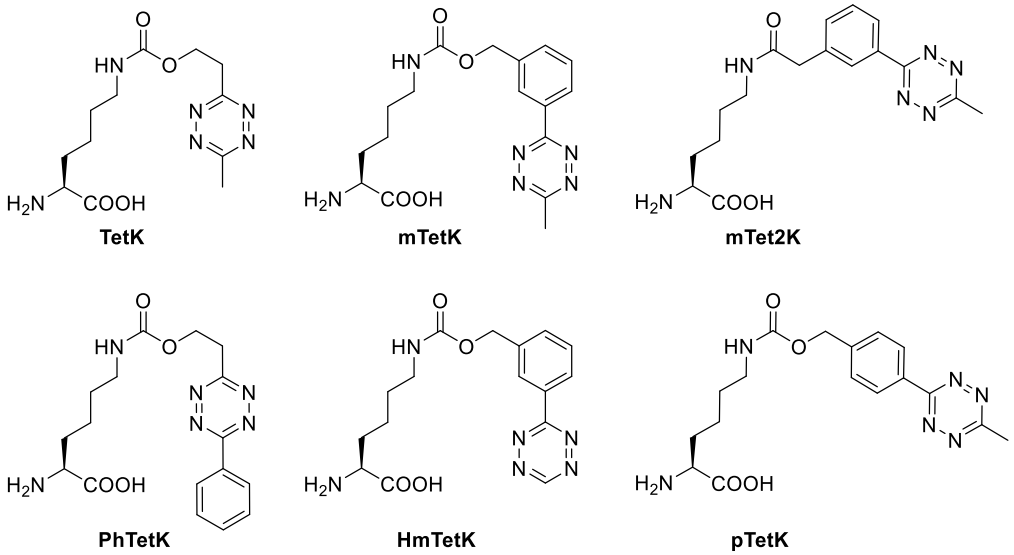


Figure 1.2: Structures of lysine-based tetrazine amino acids

After being successful in finding a PylRS mutant for any of the new tetrazines, they will be incorporated into a variety of model proteins and labeling analyzed via iEDDAC with known dienophiles. Then, the newly designed photo-inducible tetrazine ligation with photo-DMBO (Figure 1.3) will be further characterized on small molecule level and in biological settings, including *in vitro*, in cell lysate and in live *E. coli* cells to establish an approach for the spatio-temporally controlled labeling of proteins in living cells.

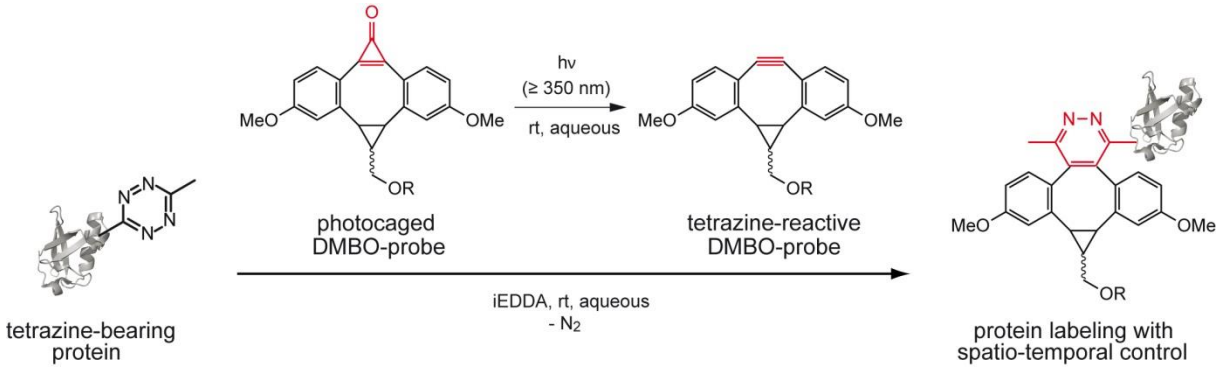


Figure 1.3: Photo-inducible iEDDAC reaction between tetrazine amino acids incorporated into proteins and photo-DMBO compounds upon irradiation at 365 nm.

1.2 General background: Light-inducible bioorthogonal reactions

Bioorthogonal reactions are a powerful tool to endow biomolecules *in vitro* and in living systems with various probes to study biological processes. If one of these reaction partners is masked with a "photocage" and only reacts upon activation with light of a certain wavelength this adds the factor of spatiotemporal control over these bioorthogonal reactions. Other masking groups can be removed by the addition of metals[251a, 251c], small molecules[65d, 252] or enzymes.[66]

There are only a few photo-activatable ligation reactions that have been applied in biological settings. One of these reactions known as Photo-click reaction is based on the formation of a nitrile imine species *in situ* upon irradiation with UV light from a tetrazole, which can also be described as 1,3-dipole that can react with alkenes to yield a stable pyrazoline product in a 1,3-dipolar cycloaddition (Figure 1.4a). [266]

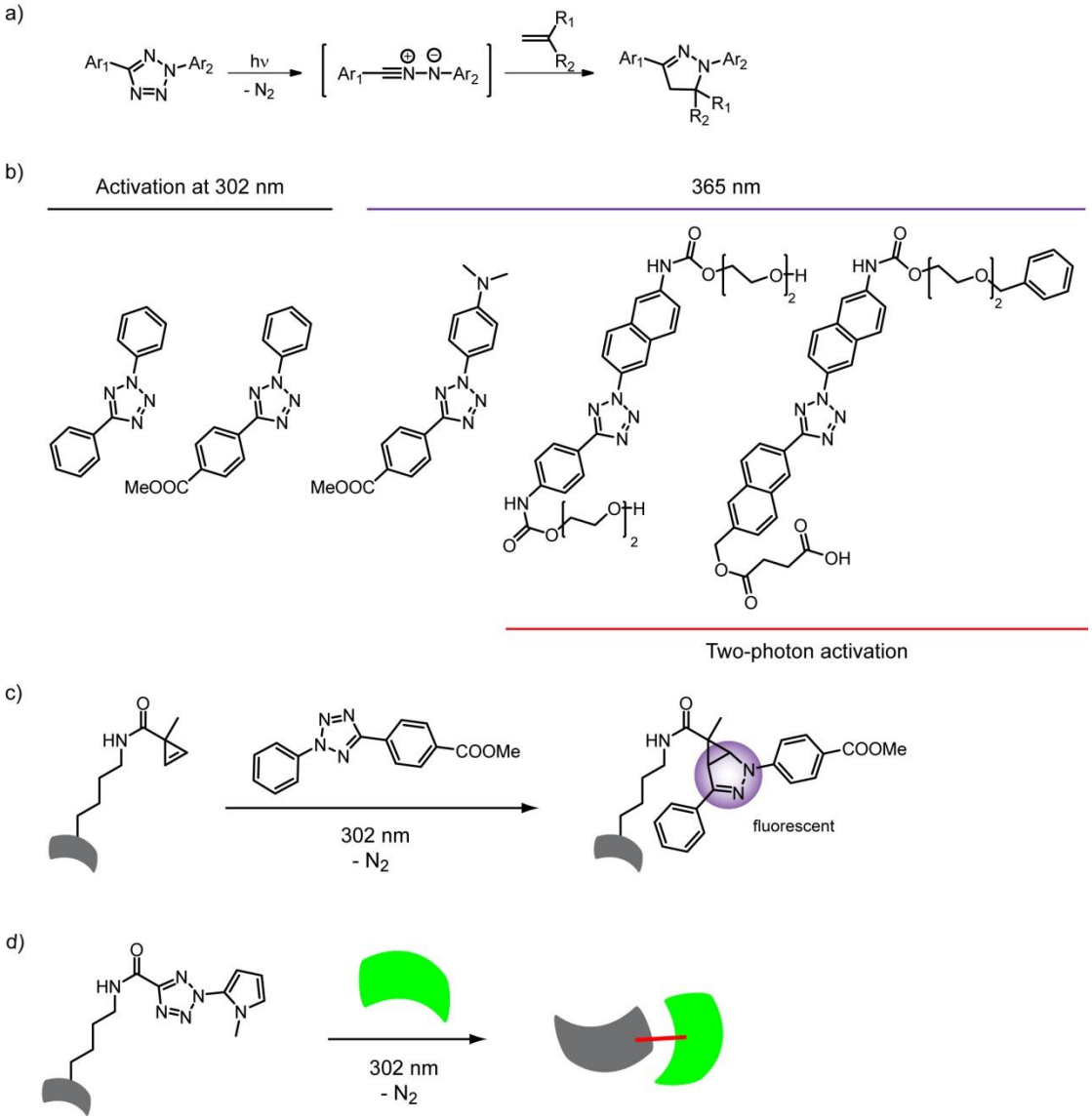


Figure 1.4: Photo-click reaction for bioorthogonal labeling.
a) Photo-activated 1,3-dipolar cycloaddition between a 2,5-diaryl-tetrazol and alkenes. b) Substituents on the aryl rings of 2,5-diaryl-tetrazoles shift wavelength of activation from 302 nm to two-photon activation c) Genetically encoded *N*$^{\varepsilon}$-(1-methylcycloprop-2-enecarboxyamido)-lysine reacts with tetrazole upon activation by light to form fluorescent pyrazoline products. d) 2-aryl-5-carboxytetrazole uAA for photocrosslinking. Grey and green shapes denote a biomolecules.

This reaction was first described around 50 years ago between a diaryltetrazole and methyl crotonate.[267] The first application within a biological context employs this 1,3-dipolar cycloaddition upon activation at 302 nm to synthesize a MDM2 inihibitor.[268] As UV light is harmful to living cells, tetrazoles that can be activated at wavelength > 350 nm have been developed to be employed for bioorthogonal labeling of biomolecules.[269] Another advantage besides activation by light, are the fluorogenic properties of the reaction, as the resulting pyrazoline displays fluorescence, which can be used to visualize successful labeling (Figure 1.4c).[269d, 270] The kinetics for the Photo-Click reaction are determined by the cycloaddition step which is slower than the photoconversion of tetrazole. It is in the range of 1-10^4 $M^{-1}s^{-1}$,[269d, 271] with the fastest reaction rates observed between a highly strained spirocyclic alkene, spiro[2.3]hex-1-ene, and a tetrazole that is activated at 302 nm.[272] Modification of the tetrazole to shift activation > 350 nm leads to a 5-10 fold decrease in the k_2 value.[273] The most recent reports show the development of new tetrazoles that can be triggered by two-photon excitation at wavelengths ~700 nm that are more suitable to living systems (Figure 1.4b),[274] however tetrazoles are not completely bioorthogonal and 2-aryl-5-carboxytetrazole uAAs have been genetically encoded for photocrosslinking of proteins (Figure 1.4d).[275]

Another possibility to photocage one of the bioorthogonal reaction partners is to mask an alkyne moiety that reacts in click reactions with azides as cyclopropenone functionality. Upon irradiation of the cyclopropenone by light of certain wavelength CO is extruded to result the alkyne. The reaction proceeds via an intermediate that can be described by to resonance structures - a semi-carbene and a semi-zwitterion (Figure 1.5a).[276] The substituents R and R' strongly influence the wavelengths that induce the decaging reaction. Alkyl-substituted cyclopropenones decage at 250 nm to the alkyne, while two aryl substituents shift the wavelength up to 365 nm (Figure 1.5a).[277] These compounds have high quantum yields ranging from 36 to 78 % and the resulting decarbonylation proceeds quite fast, which necessitates irradiation only for a short amount of time to yield the alkyne quantitatively.[278] These characteristics make the application *in vivo* highly interesting and by masking strained alkynes as cyclopropenones the SPAAC reaction was turned into a photo-activatable bioorthogonal reaction. Its first report was in 2009, when Popik and coworkers labeled azide-bearing glycans on the surface of living cells with a cyclopropenone protected dibenzocyclooctyne derivative (photo-DIBO) only upon irradiation at 350 nm for 1 min (Figure 1.5b).[278] The introduction of an oxygen atom into the dibenzocyclooctyne scaffold resulted the cyclopropenone protected oxa-dibenzocyclooctyne (photo ODIBO, Figure 1.5c), which so far showed the highest reaction rates (29-45 $M^{-1}s^{-1}$ with alkyl or aromatic azides, respectively) measured in a SPAAC reaction with azides after induction by light. The enhancement was observed to be twice as fast as the corresponding SPAAC reaction with the highly activated cyclooctyne BARAC.[279] In an aqueous environment the reaction kinetics between azides and strained alkynes were observed to be 4-28-fold higher compared to organic solvents, which might be due to the higher polarity of water or its influence on the donor-acceptor interaction.[279] Photo-ODIBDO enabled the light-controlled labeling of azide-bearing RNA in living cells.[280] The combination of an aza-dibenzocyclooctyne (ADIBO, Figure 1.5c) and a photo-DIBO moiety in a heterobifunctional linker was used to perform sequential SPAAC reactions to ligate two azide tagged molecules; for example azide-modified bovine serum albumin (BSA) with azido-fluorescein.[281] Another approach towards a

double click reaction was the application of a bisreactive molecule, the so called Sondheimer diyne (Figure 1.5c), where both alkyne moieties can react in a SPAAC reaction.[282] To optimize the low stability of the diyne in aqueous solutions Popik and coworkers reported a bis-cyclopropeneone masked dibenzo[a,e]cyclooctadiyne (photo-DIBOD, Figure 1.5c), which also enabled triggering the ligation reaction by light to label azido-BSA with azido-fluorescein or immobilize azide-bearing substrates on azide-functionalized 96-well plates.[283] The second click reaction proceeds much faster with values of 34 $M^{1}s^{-1}$ than the first SPAAC (0.02 $M^{-1}s^{-1}$) reaction. Therefore, a mono-cyclopropene protected dibenzocyclooctadiyne (MC-DIBOD, Figure 1.5c) was designed, which was employed for sequential labeling of purified proteins.[284] The first SPAAC reaction is fluorogenic as its monotriazole product shows fluorescent properties, which was used for fluorescence labeling of purified proteins.[285]

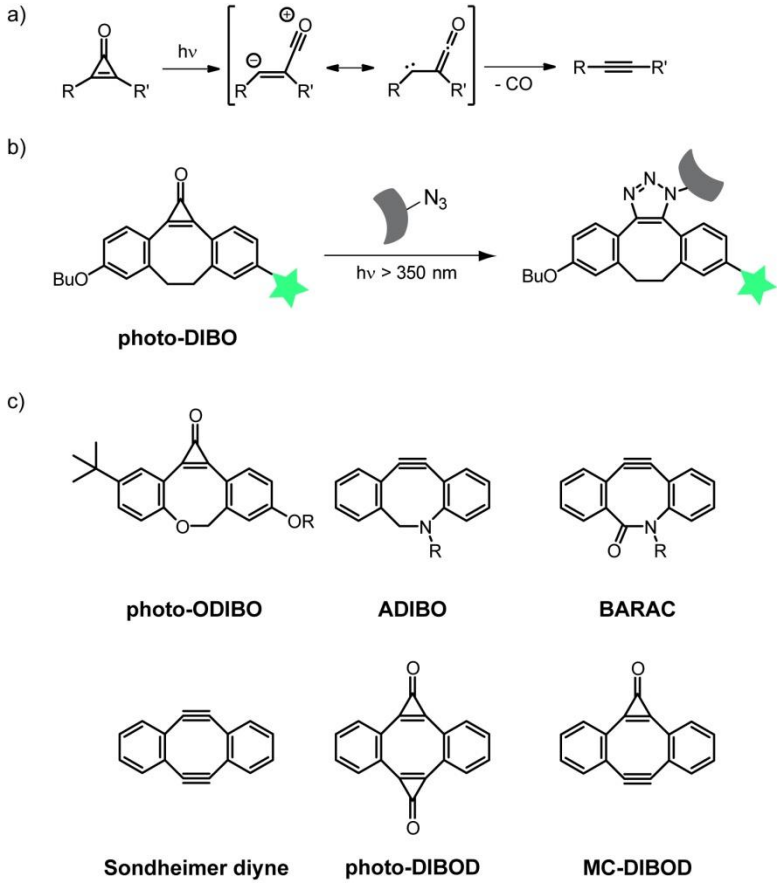


Figure 1.5: Photo-SPAAC reactions for bioorthogonal labeling
a) Photo-induced decarbonylation of cyclopropenones to the alkynes. b) Labeling of azide-tagged biomolecules in a SPAAC reaction upon activation with light. Grey shape denotes a biomolecule and the green star represents a biophysical probe (e.g. fluorophore). c) Different cyclopropenone-caged dibenzocyclooctyne derivatives.

In recent years a few different light-inducible iEDDAC reactions have been reported. Fox and coworkers described a light/enzyme-triggered redox-activation of dihydrotetrazines to tetrazines (Figure I.19)[262b]. Klán and coworkers were inspired by Popik and developed a cyclopropenone-caged dibenzosilacyclohept-4-yne that can react with azides as well as tetrazines upon light activation (Figure 1.6a). The reaction rate for the iEDDAC reaction of 260 $M^{-1}s^{-1}$ they reported, however, was measured using dicarboxy-1,2,4,5-tetrazine, one of the fastest tetrazines known, which is however not stable in aqueous buffer.[286] Less reactive tetrazines used for biological applications would therefore display rates 2-3 orders of

magnitude lower.[138a] The Loughlin group chose a different approach and developed 3-*N*-substituted spirocyclopropenes that bear bulky, photo-cleavable substituents (6-nitroveratryloxycarbonyl) in position C3, which abolishes the reaction with tetrazine due to steric hindrance and a electron-withdrawing difluoromethylen group at the second C3 position to stabilize the scaffold (Figure 1.6b). Removing this substituent with light triggers the iEDDAC reaction, showing however rather low reaction kinetics with monosubstituted 1,2,4,5-tetrazine (second order rate constant of 0.0006 $M^{-1}s^{-1}$). Replacement of the difluoromethylen moiety by a ketone group enhanced reactivity 270-fold (0.11 $M^{-1}s^{-1}$) or 125-fold (0.05 $M^{-1}s^{-1}$), when reacted with 3,6-di-2-pyridyl-1,2,4,5-tetrazine, which is however rather slow for an iEDDAC reaction and not useful for quantitative protein labeling.[138a]

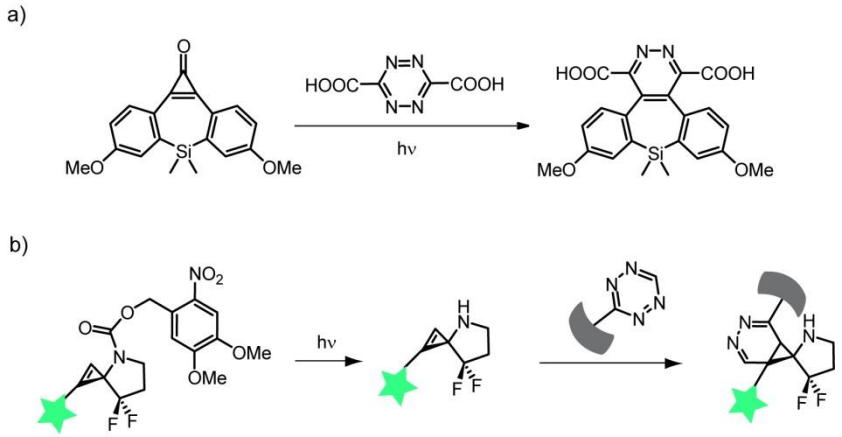


Figure 1.6: Photo-inducible iEDDAC reactions between tetrazines and a) a cyclopropenone-caged dibenzosilacycloheptyne or b) a 3-*N*-substitued spirocyclopropene. Grey shape denotes a biomolecule and the green star represents a biophysical probe (e.g. fluorophore)

As the reported approaches for photo-induced iEDDAC suffer from slow reaction rates (caged cyclopropenes) or are limited by the availability of air-stable dihydrotetrazines the development of a novel, photo-activatable tetrazine ligation with faster reaction kinetics would be highly demanding.

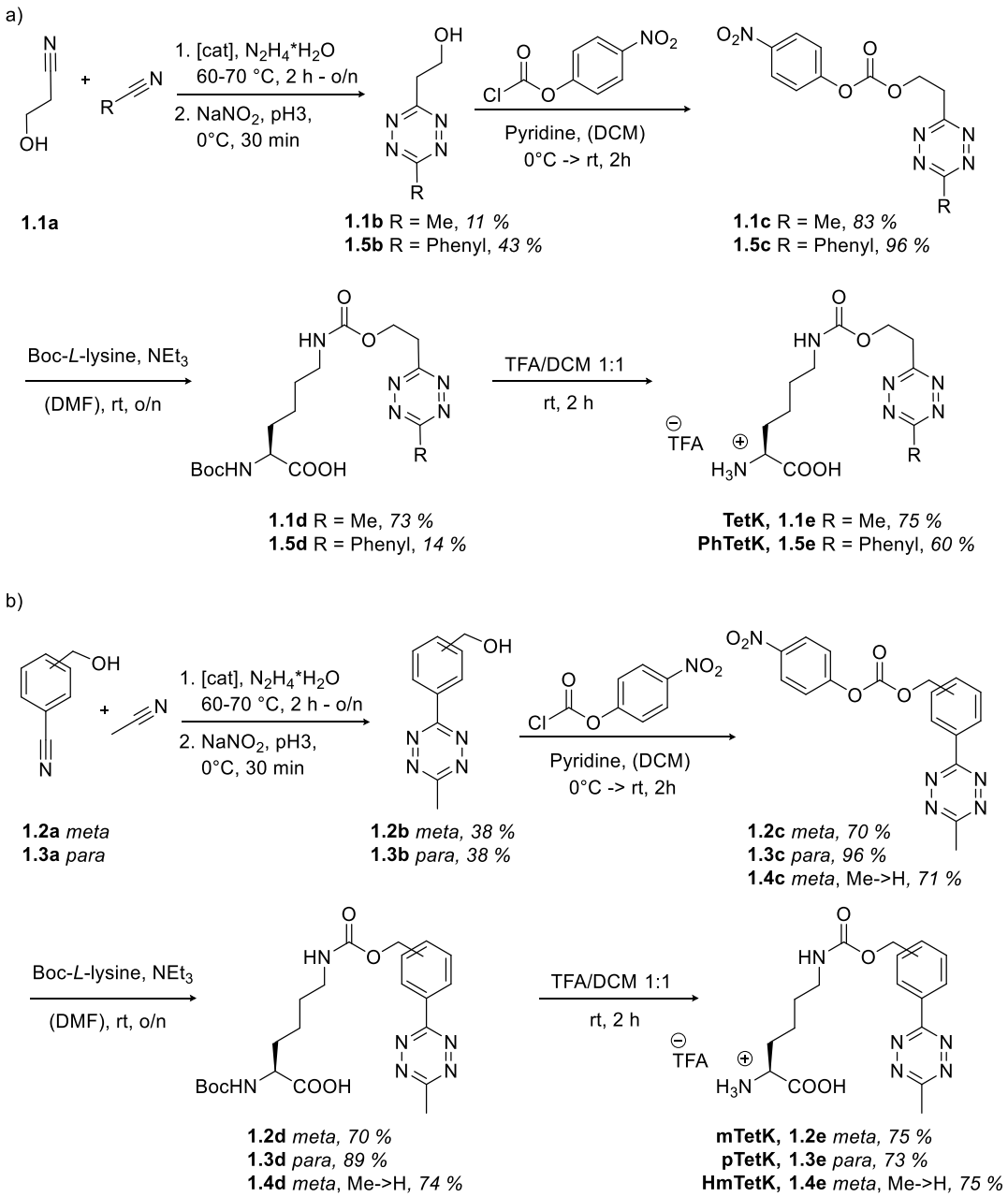
1.3 Results and Discussion

1.3.1 Synthetic procedures for lysine based tetrazine amino acids

1.3.1.1 General synthetic route

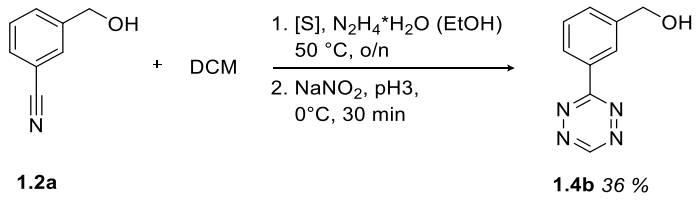
The first idea to enhance the chances of finding a PylRS mutant for incorporation of tetrazine-bearing amino acids into proteins was to design the uAAs as derivatives of L-lysine to resemble the natural substrate. Secondly, we wanted to synthesize a pool of structurally diverse tetrazines to modify the lysine scaffold (Figure 1.2) to further increase the chance of finding a synthetase, since it is difficult to predict preferences of PylRS. In all but one case (mTet2K) the tetrazine moiety was attached to the ε-amino group of L-lysine via a carbamate linkage, as these linkages are known to be stable in *E. coli* and mammalian cells and seem to be preferred over amide linkages by Mb PylRS variants. *E. coli* endogenously expresses the deacylase CobB, which was shown to cleave amide bonds between the ε-amino group of L-lysine and unnatural moieties for bioconjugation.[63] To find a sweet spot between reactivity of the tetrazine moieties and their stability under physiological conditions, a combination of aryl and alkyl substituents (mTetK, mTet2K, pTetK, PhTetK) was chosen for most structures with the exception of one dialkyl tetrazine (TetK) and one uAA bearing an aryl moiety as single substituent (HmTetK).

For the majority of the newly designed tetrazine derivatives of lysine, the synthesis was achieved in 4 steps (Scheme 1.1). First, the corresponding tetrazine, bearing a primary alcohol was generated. Therefore, hydroxypropionitrile or *meta*- respectively *para*-hydroxymethyl-substituted benzonitrile was reacted in a condensation with acetonitrile or benzonitrile and hydrazine hydrate facilitated by zinc(II)- or nickel(II) salts. The corresponding unsymmetrically substituted tetrazines were obtained in 11-43 % yield, which is in agreement with literature.[187, 287] The low yields are due to a number of possible side reactions to the two corresponding symmetrically substituted tetrazines, as well as the formation of tetrazoles.[182] For a few of the tetrazines, starting materials like *meta*- or *para*-hydroxymethyl benzonitrile (**1.2a, 1.3a**) or benzonitrile were difficult to separate from the product. Purification by RP-HPLC on the step of the deprotected uAA gave mostly pure products in satisfying yields.



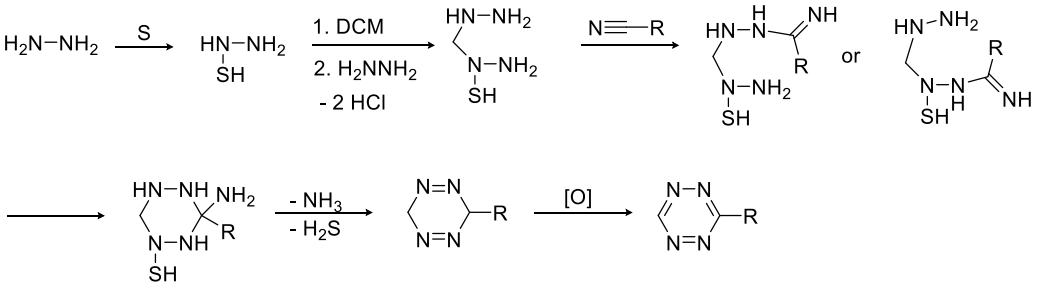
Scheme 1.1: Synthesis of lysine-based tetrazine amino acids. a) Tetrazine uAAs with an alkyl linker. b) Tetrazine uAAs with an aryl linker.

3-(3-(Hydroxymethyl)phenyl)-1,2,4,5-tetrazine (**1.4b**) was not synthetically accessible via the standard tetrazine reaction conditions. Here, the condensation of 3-(hydroxymethyl)-benzonitrile with hydrazine hydrate was achieved with an equimolar concentration of dichloromethane, facilitated by sulfur (Scheme 1.2).[288]



Scheme 1.2: Formation of 3-(3-(Hydroxymethyl)phenyl)-1,2,4,5-tetrazine

Audebert and coworkers, who developed this synthesis approach using DCM towards 3-monosubstituted unsymmetrical 1,2,4,5-tetrazines, also proposed a possible mechanism. Sulfur activated hydrazine attacks and adds itself to DCM under loss of hydrogen chloride. Addition of a second hydrazine molecule leads to formation of 1-(hydrazinylmethyl)-1-mercaptohydrazine, which can in turn attack 3-(hydroxymethyl)benzonitrile and generate the corresponding 3,6-dihydrotetrazine by elimination of ammonia and hydrogen sulfide, which is further oxidized to obtain the 1,2,4,5-tetrazine.[288]

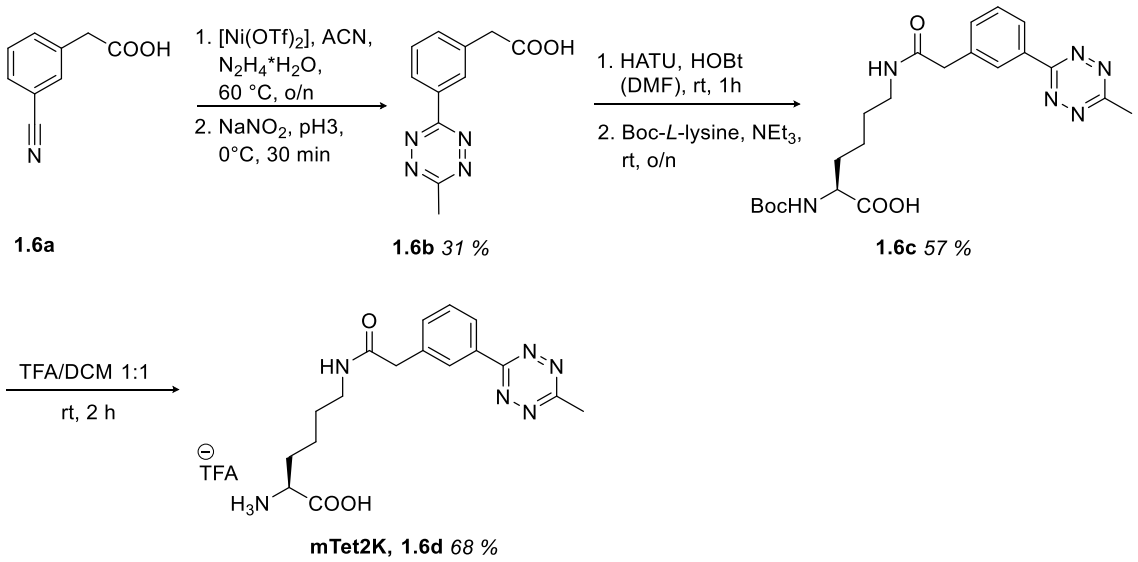


Scheme 1.3: Proposed mechanism of tetrazine formation with DCM[288]

Tetrazines, bearing a hydroxyl moiety, were subsequently activated with 4-nitrochloroformate as 4-nitrophenyl carbonates in 70-96 % yields. Hydroxyethyl-substituted tetrazines (**1.1b, 1.5b**) are prone to elimination of water to the vinyl-tetrazine upon addition of base; therefore base was added equimolar to the tetrazine as final component to the reaction to avoid elimination. This was successful in case of 3-(2-Hydroxyethyl)-6-methyl-1,2,4,5-tetrazine (**1.1b**) compared to 3-(2-Hydroxyethyl)-6-phenyl-1,2,4,5-tetrazine (**1.5b**), where eliminated side product was still being observed in the reaction mixture.

The activated tetrazines were reacted with α-amino Boc protected L-lysine to modify the side chain with a carbamate linkage. Attack of the ε-amino group on the carbonate lead to release of 4-nitrophenol and the formation of carbamate-linked tetrazine amino acids in 60 to 89 % yield. BocPhTetK (**1.5e**) was obtained in low yields of 14 %, which might be due to the difficulties of starting material removal in the first steps and the problems concerning elimination as mentioned before. Boc deprotection was performed in a 1:1 mixture of TFA in DCM with a few drops of water to scavenge the resulting isobutene, followed by precipitation in cold diethyl ether, which gave the final amino acids as TFA salts in 60-75 % yield.

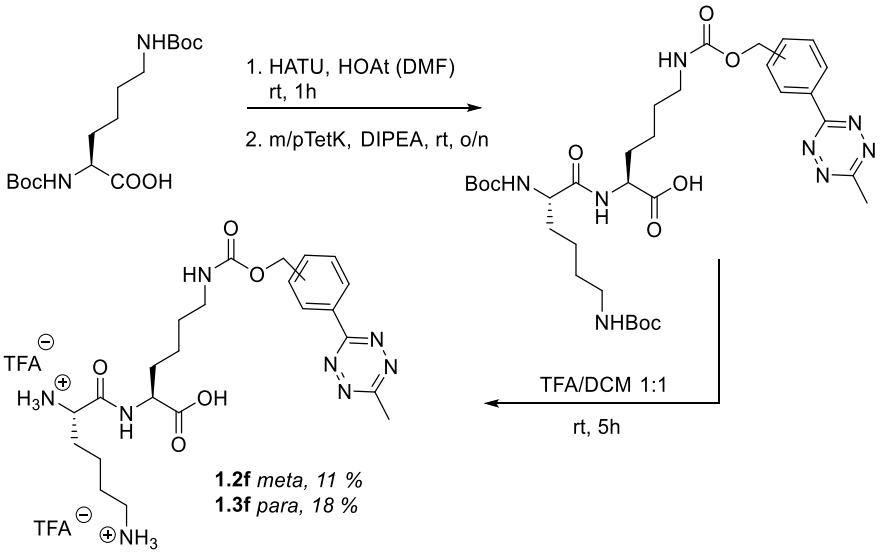
Synthesis of mTet2K (**1.6d**) displaying an amide linkage at the ε-amine was realized in 3 steps (Scheme 1.4). The activation of the carboxyl moiety on the tetrazine as active ester was performed with HATU and HOBT for 1 h prior to the addition of Boc-L-lysine giving the product in only 57 % yield. This might be due to the use of HATU and HOBt in 0.9 equivalents instead of a slight excess compared to the tetrazine to avoid activation of the free carboxyl terminus of Boc-L-lysine and formation of dipeptides as side products. Generation of the tetrazine and the deprotection reaction went according to the general synthesis Scheme 1.1.



Scheme 1.4: Synthesis of mTet2K

1.3.1.2 Synthesis of dipeptides via solution based synthesis and SPPS

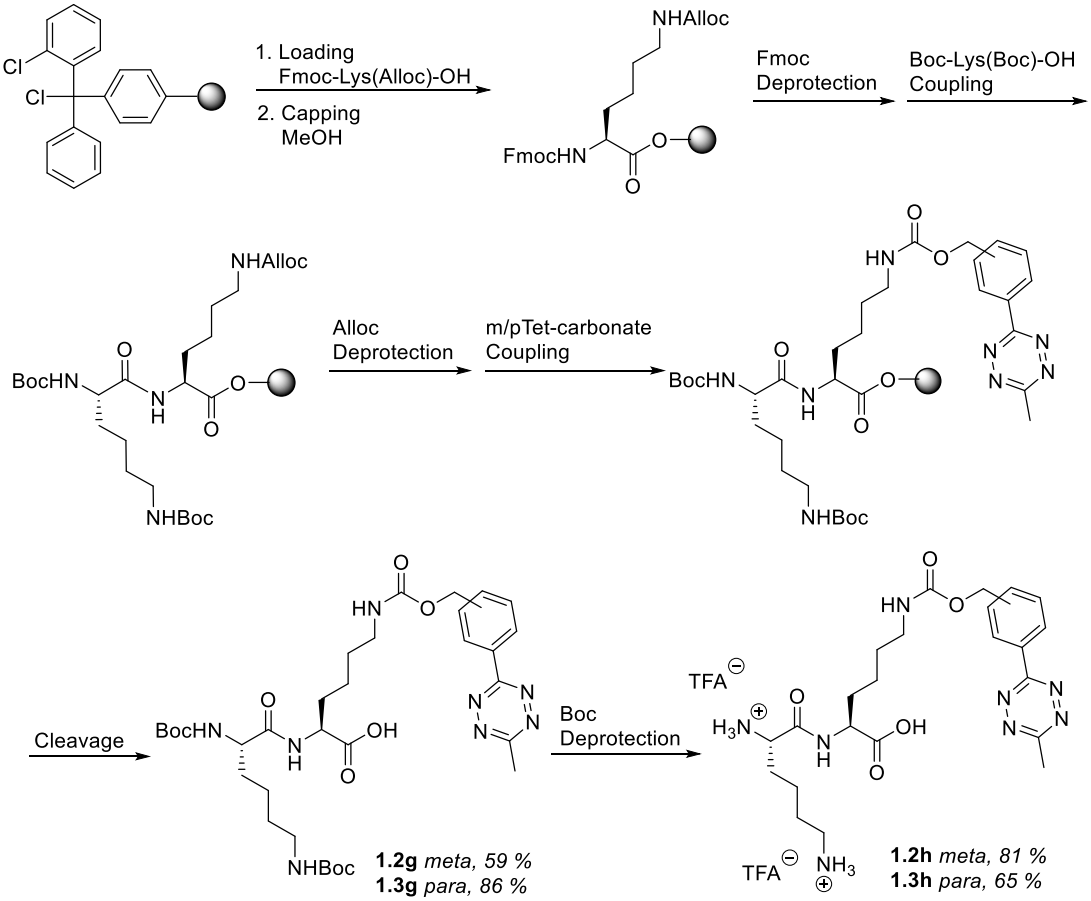
The final tetrazine amino acids were used to prepare stock solutions (100 mM) in water or DMSO with a TFA content of 200 mM, which were used for expression test experiments. As m- and pTetK heavily precipitated in aqueous media and solutions, dipeptides with L-lysine attached in the *N*-terminal position were prepared, we hoped for better solubility and uptake as described by Lou *et al.* for phosphotyrosine-based uAAs.[289] In the cytosol the *N*-terminal lysine is cleaved from the dipeptide by endogenous peptidases and the uAA is set free, ready to be incorporated.



Scheme 1.5: Synthesis of dipeptides in solution

The first attempts to synthesize K-m/pTetK (**1.2f, 1.3f**) were carried out in solution, as displayed in Scheme 1.5. α- and ε- Boc protected L-lysine was activated with HATU and HOAt for 1 h, followed by addition of m- or pTetK, which resulted the corresponding dipeptides after Boc deprotection in 11-18 % over two steps. The low yields arose from a

number of side reactions on the unprotected *C*-terminus of m- or pTetK. Therefore, a synthesis on solid support was developed (Scheme 1.6). There, 2-chlorotrityl resin was loaded with α-Fmoc and ε-Alloc protected L-lysine, followed by Fmoc deprotection and coupling with α- and ε-Boc protected L-lysine. After Alloc deprotection the tetrazine moiety was attached onto the free amino group by nucleophilic attack of the amine on the tetrazine carbonate, forming a carbamate linkage. Cleavage from the resin and Boc deprotection yielded K-m/pTetK (**1.2h, 1.3h**) in 22-52 % after 8 steps. The dipeptides were further purified by RP-HPLC. The resulting purified dipeptides displayed enhanced water solubility and could be prepared as 100 mM stock in MQ-water without the addition of TFA.



Scheme 1.6: Synthesis of dipeptides via solid phase

1.3.2 Site-specific incorporation of tetrazine amino acids into proteins

With a portfolio of various flexible tetrazine-modified amino acids at hand, the next goal was to site-specifically incorporate them into proteins *in cellulo*. A large variety of structurally diverse UAAs have been genetically encoded in response to an amber codon introduced into a gene of interest by employing PylRS/$tRNA_{CUA}$ pairs from *Methanosarcina* species.[290] Many of the lysine-based uAAs that can be incorporated via PylRS mutants contain an alkyl side chain with a bulky and often hydrophobic moiety that is attached to lysine via a carbamate linkage.[291] None of the so far developed PylRS variants accepts however lysine derivatives with bulky polar side chains such as in the newly synthesized tetrazine uAAs (Figure 1.2). Guided by structural analyses of the *C*-terminal catalytic centre of wt PylRS and its mutants,[292] we screened a panel of >30 different *M. barkeri* PylRS mutants (Table II.4) for their ability to direct the selective and site-specific incorporation of TetK, PhTetK, mTetK, K-mTetK, HmTetK, pTetK and K-pTetK into *C*-terminally His-tagged superfolder green fluorescent protein (sfGFP) containing a premature amber codon at position 150, which proved to be unsuccessful in finding a PylRS mutant accepting any of the tetrazine amino acids (Figure 1.7, not all data shown). N^{α}-(*tert*-Butoxycarbonyl)-L-lysine (BocK), which is known to be recognized and processed by the wild type PylRS synthetase (D4)[70c], was used as a positive control for the production of full-length sfGFP with molar mass of approx. 27 kDA. Cultures where no uAA was added served as a negative control. In those samples only truncated sfGFP (MW ~ 14 kDA) was produced, showing none of the natural amino acids was introduced instead of uAA.

Synthetases D1 and MC1 (Table II.4) showed expression of full length sfGFP-N150TAG-His6 in presence of mTetK, however purification with Ni-NTA beads, followed by electrospray ionization mass spectrometry (ESI-MS) analysis revealed a mass that was not corresponding to sfGFP-N150mTetK-His6, but was 70 Da lower. This revealed to be N^6-(((3-cyanobenzyl)oxy)-carbonyl)-L-lysine (3-CNPhK, b), a side product originating from mTetK synthesis, where (3-hydroxymethyl)benzonitrile is used as a starting material and incomplete removal leads to activation and coupling to lysine. Since the contamination of 3-CNPhK in mTetK was found to be < 5 %, this indicates a high specificity of PylRS mutants D1 and MC1 towards 3-CNPhK as a substrate for aminoacylation.

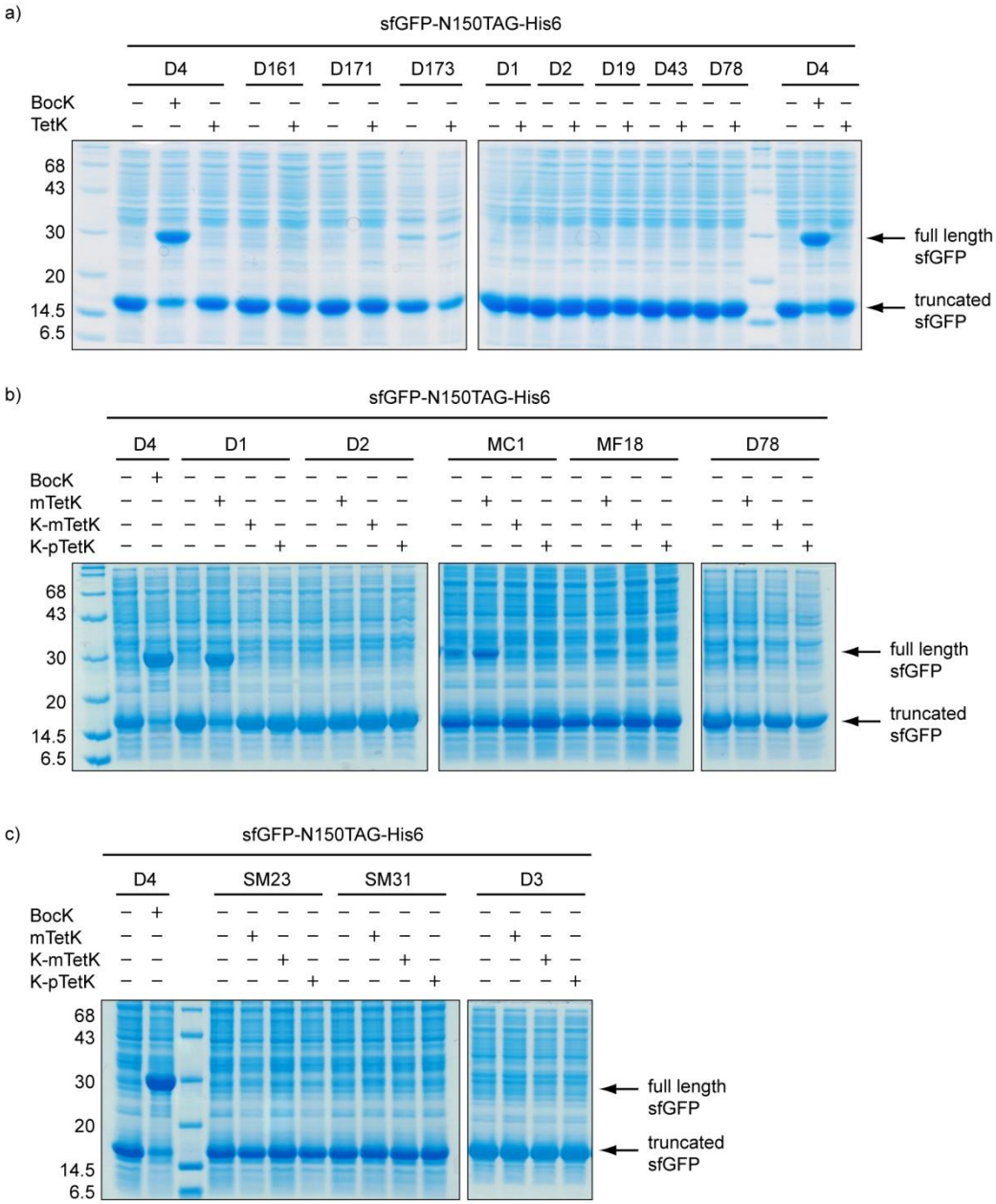


Figure 1.7: Coomassie stained SDS-PAGEs showing the PylRS synthetase screening for the incorporation of lysine based tetrazine amino acids a) TetK and b, c) mTetK, K-mTetK and K-pTetK.

As a next step in finding a synthetase, a directed evolution approach was chosen for the uAAs TetK, mTetK and pTetK. The selection protocol was performed with one round of positive and one round of negative selection, followed by sfGFP fluorescence read-out including a second positive selection step. Directed evolution was performed with libraries AB3, ABshuffle, Nicholas_RS5 and Lib_fw (Table II.33), which did not result a synthetase for the incorporation of any of the tetrazine amino acids TetK, mTetK and pTetK, but rather selected PylRS mutants for the incorporation of the small contamination of 3-CNPhK in mTetK (data not shown). The uAA BCNK, which was already known to be incorporated by a PylRS mutant (Y271M, L274G and C313A, *Mb* numbering, D1),[58a] albeit in relative low yields, was used as a positive control for the selection process with library Lib_fw, since the reported mutant D1 was evolved from this PylRS library. Fluorescence read-out yielded two clones

A11 (Y271A C313V, *Mb* numbering, dubbed MF3) and A8 (Y271G C313V, *Mb* numbering, dubbed MF4), which showed better incorporation than D1 (Figure 1.8).

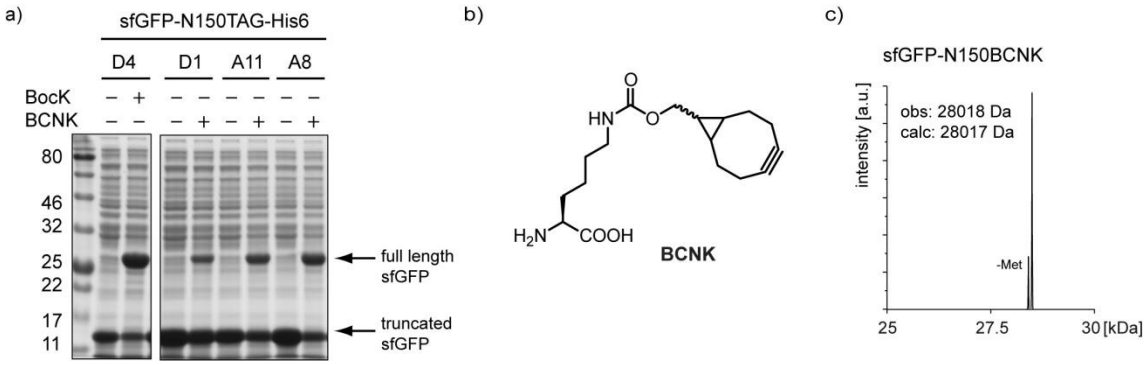


Figure 1.8: Incorporation of BCNK into sfGFP-150TAG-His6.
a) Comparison of incorporation between newly evolved clones A11 and A8 with published synthetase D1. b) Structure of BCNK c) ESI-MS characterization of sfGFP-150BCNK-His6 incorporation with MF3.

Having new *Mb* PylRS mutants MF3 and MF4 at hand, we wanted to test them for the incorporation of the different tetrazine amino acids mTetK and the dipeptides K-mTetK and K-pTetK (Figure 1.9) into sfGFP-N150TAG-His6 as a last attempt in finding a synthetase and were given a positive surprise. MF3 showed production of full-length GFP with mTetK and K-mTetK (Figure 1.9a), and especially the result for K-mTetK looked promising, since the dipeptide had been purified by HPLC, while mTetK was still the batch contaminated with traces of 3-CNPhK. ESI-MS analysis of the purified proteins after expression with mTetK and K-mTetK confirmed the incorporation of mTetK, however as a 70:30 mixture of 3-CNPheK to mTetK for the mTetK batch and a 20:80 ratio for K-mTetK (Figure 1.9b). After this indication that it is possible to remove the 3-CNPhK contamination on the uAA level, mTetK was purified by HPLC. ESI-MS analysis of purified sfGFP-N150mTetK-His6 after expression with the HPLC purified batch (Figure 1.9c,d) displayed only the correct mass for incorporation of mTetK, while no trace of incorporated 3-CNPhK was found.

Even though directed evolution was conducted with Lib_fw for mTetK, the PylRS mutant MF3 was not discovered during the selection process for mTetK, but found later to accept mTetK as substrate. The quality of the results of a directed evolution approach performed in *E. coli* is strongly influenced by the efficiency of the initial library transformation. A library where five positions in the active site are randomized already contains 3.3×10^7 members, which need to be taken up by *E. coli* cells. Their transformation efficiency might already lead to the loss of some of the mutants, which are then submitted to selection pressure. The small, but present contamination of mTetK with 3-CNPhK might have further affected the positive and the read-out selection step. 3-CNPhK is the preferred substrate for MF3 compared to mTetK and therefore might have even hindered the selection process for mTetK. One possibility to select for of PylRS mutants with an increased selectivity for mTetK is to add 3-CNPhK during the negative selection step to remove PylRS mutants accepting 3-CNPhK as substrate.

After successful incorporation of mTetK by PylRS mutant MF3, the other tetrazine amino acids were tested as well. Thereby, MF3 accepts also amino acids TetK and HmTetK for

installment into proteins, as shown by SDS-PAGE and ESI-MS analysis (Figure 1.9c,d). However, ESI-MS characterization of sfGFP-N150HmTetK-His6 showed contamination of HmTetK with 3-CNPhK and it was not possible to separate the mixture by HPLC in this case. Furthermore, sfGFP-N150HmTetK-His6 was observed to degrade rapidly after purification and storage at 4 °C, therefore HmTetK was excluded from further experiments. pTetK, K-pTetK, mTet2K and PhTetK were also not persued any further, since several synthetase screenings and performance of directed evolution with different libraries did not result any PylRS mutant for incorporation of any of them, combined with the success of finding a PylRS mutant, which effectively incorporates uAAs TetK and mTetK into proteins. Therefore, further experiments were performed with these uAAs.

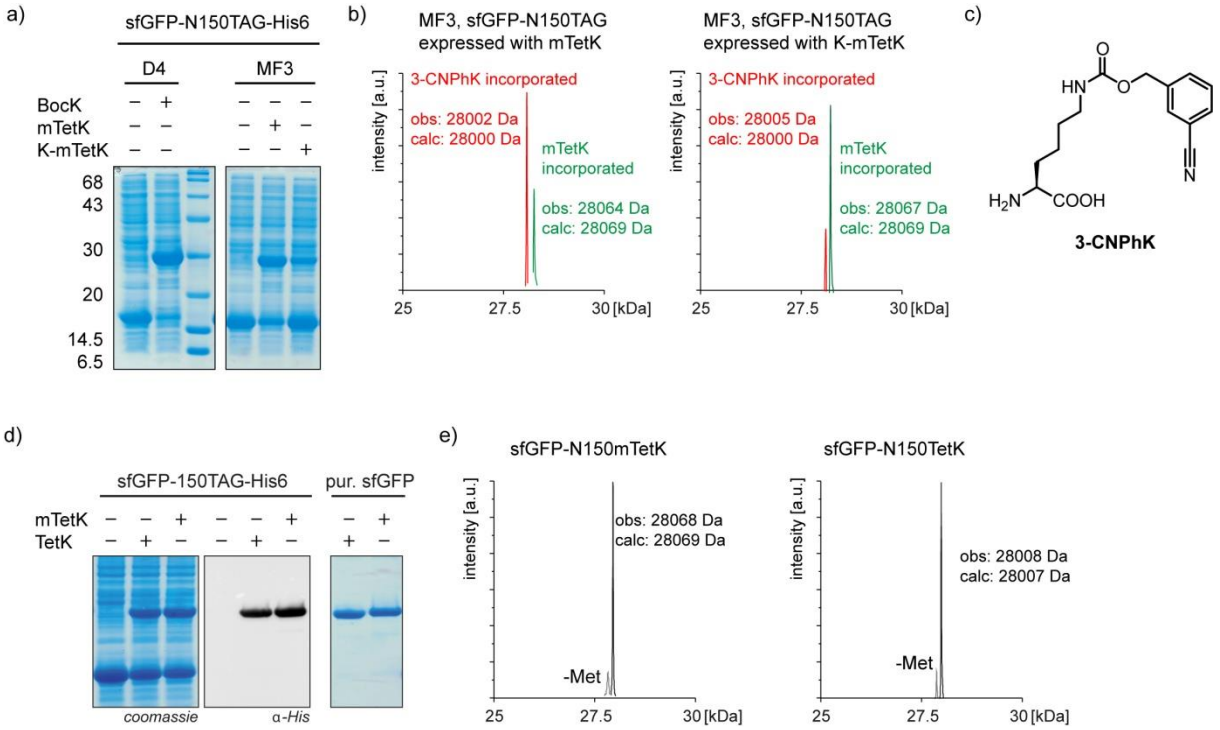


Figure 1.9: Finding a PylRS mutant for the incorporation of mTetK and TetK.
a) Coomassie stained SDS-PAGE of sfGFP-N150TAG-His6 expression using MF3 mTetK and dipeptide K-mTetK b) ESI-MS spectra of purified sfGFP-N150mTetK-His6 contaminated with sfGFP-N150-3CNPhK-His6 for expression with mTetK and K-mTetK. c) Structure of 3-CNPhK d) Coomassie stained SDS-PAGE of sfGFP-N150TAG-His6 expression test of sfGFP-N150TAG-His6 using MF3, HPLC-purified mTetK and TetK d) ESI-MS characterization of purified sfGFP-N150mTetK-His6 (left) and sfGFP-N150TetK-His6 (right).

Expression of sfGFP-N150-His6 was performed with mTetK and TetK as performed in larger cultures and the full-length proteins sfGFP-mTetK-His6 and sfGFP-TetK-His6 were isolated in good yields (Figure 1.9c, ~80 mg/L of culture) and analyzed by SDS-PAGE and α-His6 Western Blot (WB), which confirmed good incorporation efficiencies for both mTetK and TetK. Similarly, other proteins such as myoglobin (Myo) and ubiquitin (Ub) bearing site-specifically introduced amber codons produced good yields of protein in presence, but not absence of mTetK and TetK. The selective and site-specific incorporation of tetrazine-bearing amino acids was further confirmed by ESI-MS of purified proteins (Figure 1.9d, Figure 1.10).

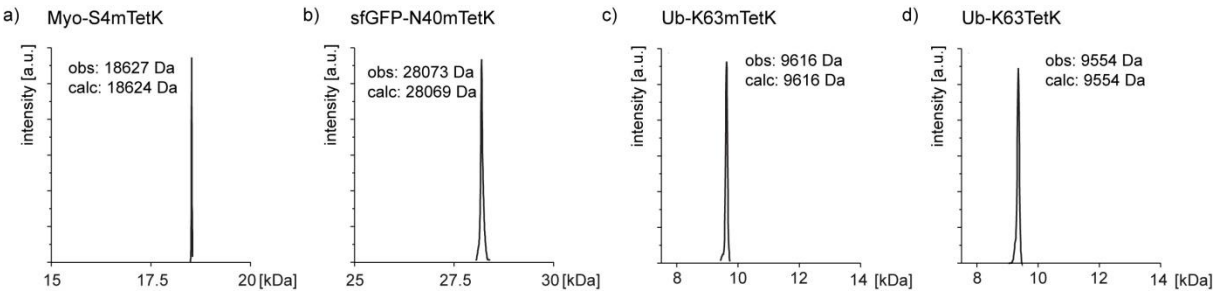


Figure 1.10: ESI-MS characterization of purified proteins
a) Myo-S4mTetK-His6, b) sfGFP-N40mTetK-His6, c) Ub-K63mTetK-His6, d) Ub-K63TetK-His6.

1.3.3 Rapid and quantitative labeling of tetrazine-modified proteins with BCN- and TCO-fluorophore conjugates

Next, the reactivity of site-specifically incorporated methyl-tetrazines in iEDDA cycloadditions was accessed with known dienophiles. To enable a fluorescence read-out of the labeling reaction, tetramethylrhodamine (TAMRA, TMR) was modified with BCN or TCO-moieties (BCN-TMR, TCO-TMR Figure 1.11a). These were then added in an excess to purified proteins bearing mTetK or TetK and the reaction analyzed by ESI-MS and SDS-PAGE based fluorescence imaging (Figure 1.11b,c). As expected, the labeling reaction with BCN- and TCO-TMR was very rapid and quantitatively modified mTetK- or TetK-bearing proteins.

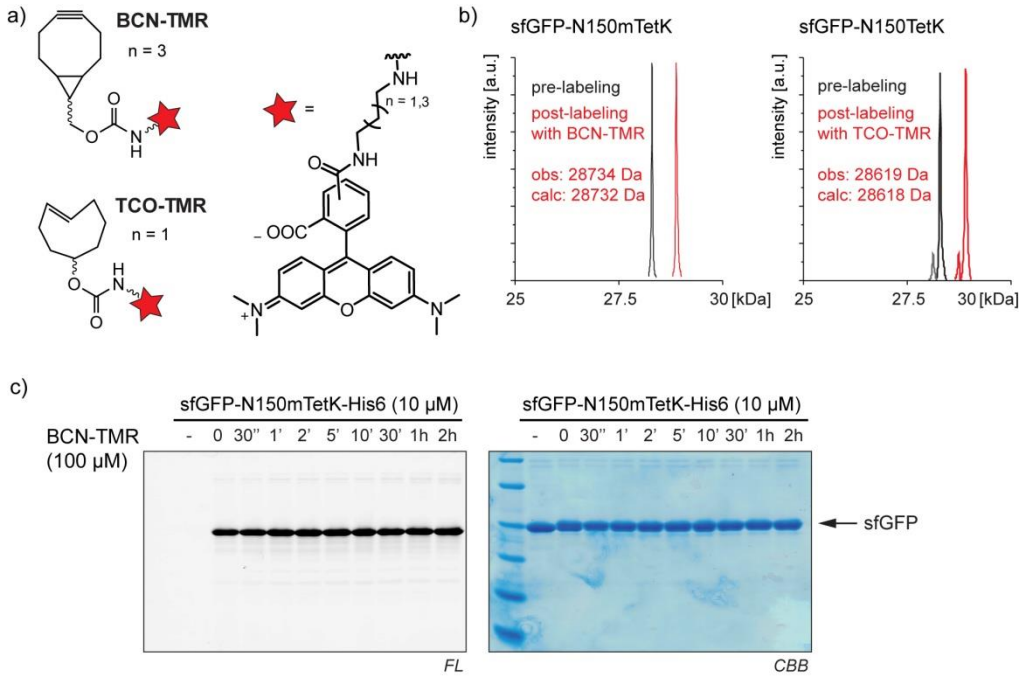


Figure 1.11: Specific and selective labeling of purified tetrazine-bearing proteins with BCN- and TCO-fluorophores.
a) Structures of BCN-TMR and TCO-TMR (b) ESI-MS shows quantitative modification of sfGFP-mTetK-His6 with BCN-TMR and sfGFP-TetK-His6 with TCO-TMR. c) SDS-PAGE fluorescence imaging of purified sfGFP-mTetK-His6 with BCN-TMR displays very rapid and quantitative labeling. CBB: coomassie brilliant blue, FL: fluorescence.

After labeling purified proteins, specificity of this cycloaddition was tested in *E. coli* cell lysate Therefore, sfGFP, Ub or Myo harboring a premature stop codon were expressed in the presence of mTetK or TetK overnight. Cells were then harvested, washed three times with

PBS to remove residual mTetK or TetK, lysed by sonication and then subjected to labeling with low µM concentrations of BCN-TAMRA or TCO-TAMRA. In gel fluorescence confirmed specific and selective labeling of proteins containing mTetK or TetK. No unspecific labeling was observed for cellular proteins or for over-expressed proteins bearing BocK instead of mTetK or TetK (Figure 1.11).

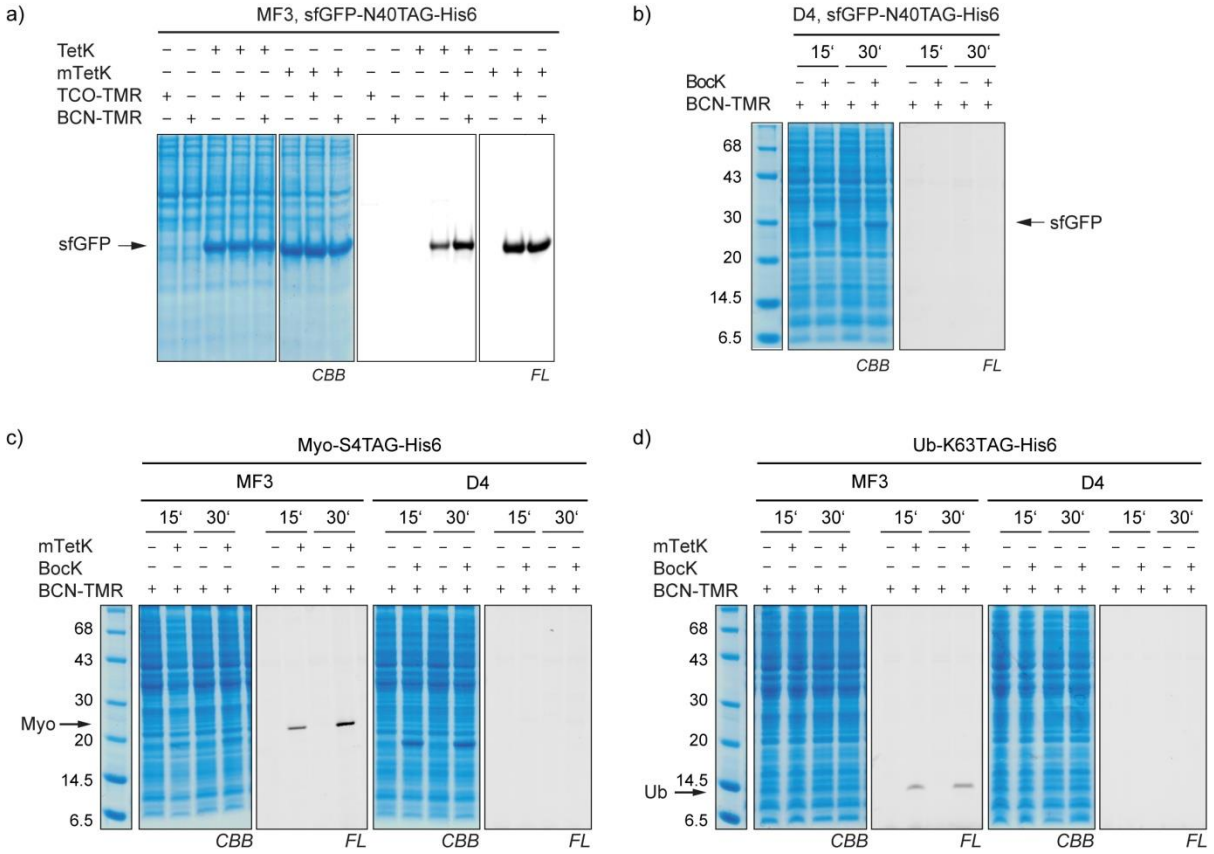


Figure 1.12: Specific and selective labeling of tetrazine-bearing proteins with BCN- and TCO-fluorophores in the *E. coli* proteome.
SDS-PAGE fluorescence imaging of a,b) TetK-, mTetK- and BocK- bearing sfGFP-N40TAG-His6 c) mTetK- and BocK-bearing Myo-S4TAG-His6 d) mTetK- and BocK-bearing Ub-K63TAG-His6 in *E. coli* cell lysate. CBB: coomassie brilliant blue, FL: fluorescence.

As BCN- and TCO-TAMRA were not able to penetrate *E. coli* cells for intracellular labeling of tetrazine-bearing proteins in living *E. coli*, cell-permeable bicyclo[6.1.0]non-4-yn-9-ol (BCN-OH) was used as an alternative for the live cell labeling of sfGFP-mTetK-His6. Purification and subsequent LC-ESI-MS analysis conformed quantitative labeling of the modified protein (Figure 1.13a,b). Outer membrane protein OmpC with an amber codon at position 232 on the cell surface of *E. coli* was chosen as target to perform live cell labeling with fluorophores. Therefore, OmpC232TAG was expressed in the presence of mTetK, the cells washed three times with PBS to remove residual mTetK and then treated with BCN-TAMRA. SDS-PAGE fluorescence imaging showed chemoselective modification of OmpC232mTetK with BCN-TAMRA (Figure 1.13c).

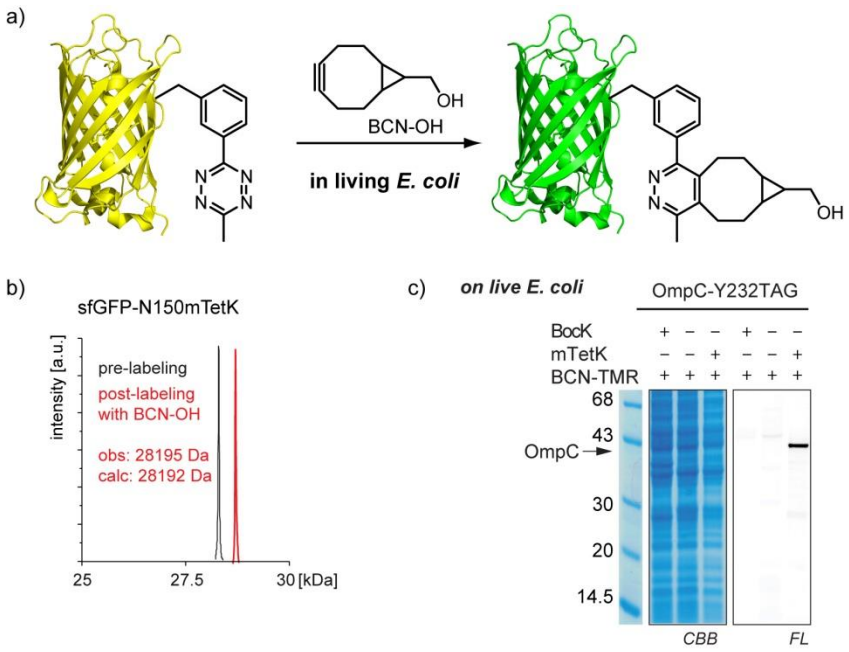
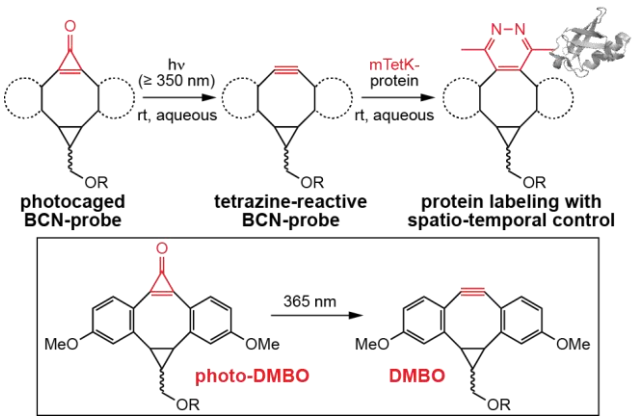


Figure 1.13: Labeling in and on living *E. coli*.
a) Schematic illustration of the labeling reaction between sfGFP-N150mTetK-His6 and BCN-OH b) ESI-MS characterization of sfGFP-N150mTetK-His6 before and after labeling in living *E. coli* with BCN-OH, shows complete modification of the protein after 2 h. c) Selective labeling of OmpC-Y232mTetK with BCN-TMR on the cell surface of live *E. coli.* CBB: coomassie brilliant blue, FL: fluorescence.

1.3.4 Development of a rapid photo-induced iEDDAC

To further extend the application of tetrazine-based bioorthogonal reporters, a photochemically triggered iEDDAC that confers temporal and spatial control to the labeling of target proteins was developed.[293] For this, the photochemical generation of reactive BCN-probes was investigated. Inspired by work from Popik, who described cyclopropenone-caged dibenzocyclooctynes for photo-induced and strain-promoted cycloadditions with azides,[294] a cyclopropenone-caged version of BCN was designed that presents a tetrazine-reactive alkyne moiety upon photochemical decarbonylation. This cyclopropenone-caged dibenzoannulated BCN probe (photo-DMBO, Scheme 1.7) should be amenable to photo-decarbonylation to form DMBO by irradiation above 350 nm.



Scheme 1.7: Photochemical decarbonylation of a cyclopropenone-caged BCN-probe (photo-DMBO) induces reactivity towards tetrazine-bearing proteins, conferring spatial and temporal control to protein labeling.

Synthesis of photo-DMBO-acetate (photo-DMBO-OAc, Figure 1.14a) was performed by Toni and is discussed in detail in his Master's Thesis.[265] Sonogashira cross-coupling of 3-iodoanisol and 3-ethynylanisol yielded the symmetrical acetylene. Lindlar's catalyst was chosen to achieve partial hydrogenation of the acetylen to generate the corresponding *Z*-alkene, which was then cyclopropanated with ethyl diazoacetate. Reduction of the ester group on the newly created cyclopropane ring gave the corresponding hydroxyl group, which was protected by acetylation. Subsequent Friedel-Crafts alkylation with tetrachlorocyclopropene in the presence of $AlCl_3$ formed the corresponding dichlorocyclopropene intermediate, which was hydrolyzed *in situ* to yield photo-DMBO-OAc (Figure 1.14a).[295] Despite its considerable ring strain, photo-DMBO-OAc is stable in aqueous buffers at physiological pH over several days. Photo-DMBO-OAc decages quantitatively to form the corresponding cycloalkyne DMBO-OAc (Figure 1.14b) upon 5 min of irradiation at 365 nm. Photo-decarbonylation was found to work effectively in different organic solvents (MeOH, DMSO), as well as in aqueous buffers. To test reactivity of photo-DMBO-OAc towards mTetK, a 100 µM solution of mTetK was incubated with a five-fold excess of photo-DMBO-OAc. mTetK reacted quantitatively to the cycloaddition compound only upon UV irradiation of the mixture (5 min at 365 nm), while when kept dark formation of product could not be observed even after prolonged incubation (Figure 1.14c), confirming the photo-inducibility of the cycloaddition reaction.

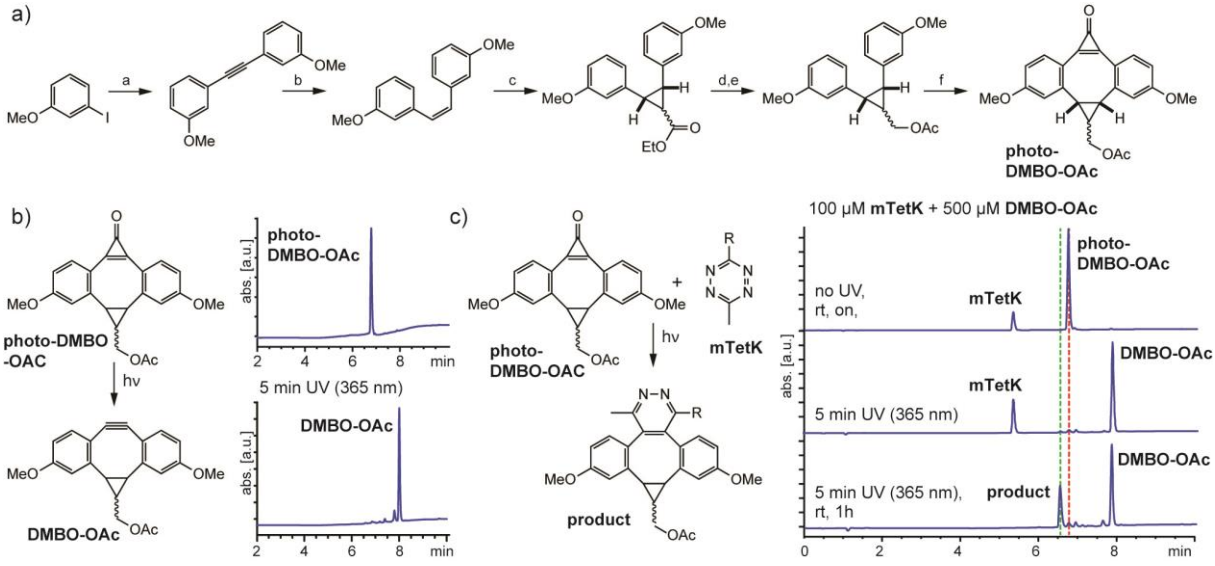


Figure 1.14: Establishing photo-iEDDAC reactivity.
a) Synthesis of photo-DMBO-OAc. Conditions: (a) 1.2 eq 3-ethynylanisole, 3 eq DIPEA, 0.1 eq CuI, 0.05eq $Pd(PPh_3)_4$, in THF, reflux, overnight, (71%); (b) 20% w/w Lindlar catalyst, in hexane, 2h, rt, (77%); (c) 2.5 eq ethyl diazoacetate, 0.06 eq $CuSO_4$, in toluene, 75°C, overnight, (18%); (d) 2 eq $LiAlH_4$, in Et_2O, 0°C-rt, 2h, (68%); (e) 2.6 eq Ac_2O, 0.05 eq DMAP, 4.9 eq NEt_3, in DCM, 0°C-rt, 2h, (70%); (f) 1 eq tetrachlorocyclopropene, 3 eq $AlCl_3$, in DCM, -20°C-rt, 4h, aqueous workup (55%). b) Short UV-irradiation (5 min, 365 nm) converts photo-DMBO-OAc quantitatively to DMBO-OAc. c) Photo-DMBO-OAc reacts with mTetK only when irradiated at 365 nm for 5 min to form iEDDAC product.

1.3.5 Rapid, light-induced protein labeling via photo-iEDDAC

As a next step, the light-induced iEDDAC between photo-DMBO and mTetK was characterized on protein level. Therefore, photo-DMBO was further functionalized by Toni. The introduction of a PEG linker with a primary amino group (photo-DMBO-NH_2, Figure 1.15a) enhanced water-solubility of the photo-DMBO compounds and allowed further modification with fluorophores. Photo-DMBO-NH_2 was quantitatively decaged to DMBO-NH_2 by irradiation at 365 nm for 5 min (Figure 1.15a) and subsequently used for the labeling of sfGFP-N150mTetK-His6. The labeling was observed to be complete within 5 min upon addition of a 25-fold excess of DMBO-NH_2. Also irradiation (10 min, 365 nm) of photo-DMBO-NH_2 in the presence of sfGFP-N150mTetK-His6 lead to complete modification of sfGFP within 30 min as shown by ESI-MS analysis (Figure 1.15b,d). Importantly, incubation in the dark did not yield any cycloaddition product (Figure 1.15c).

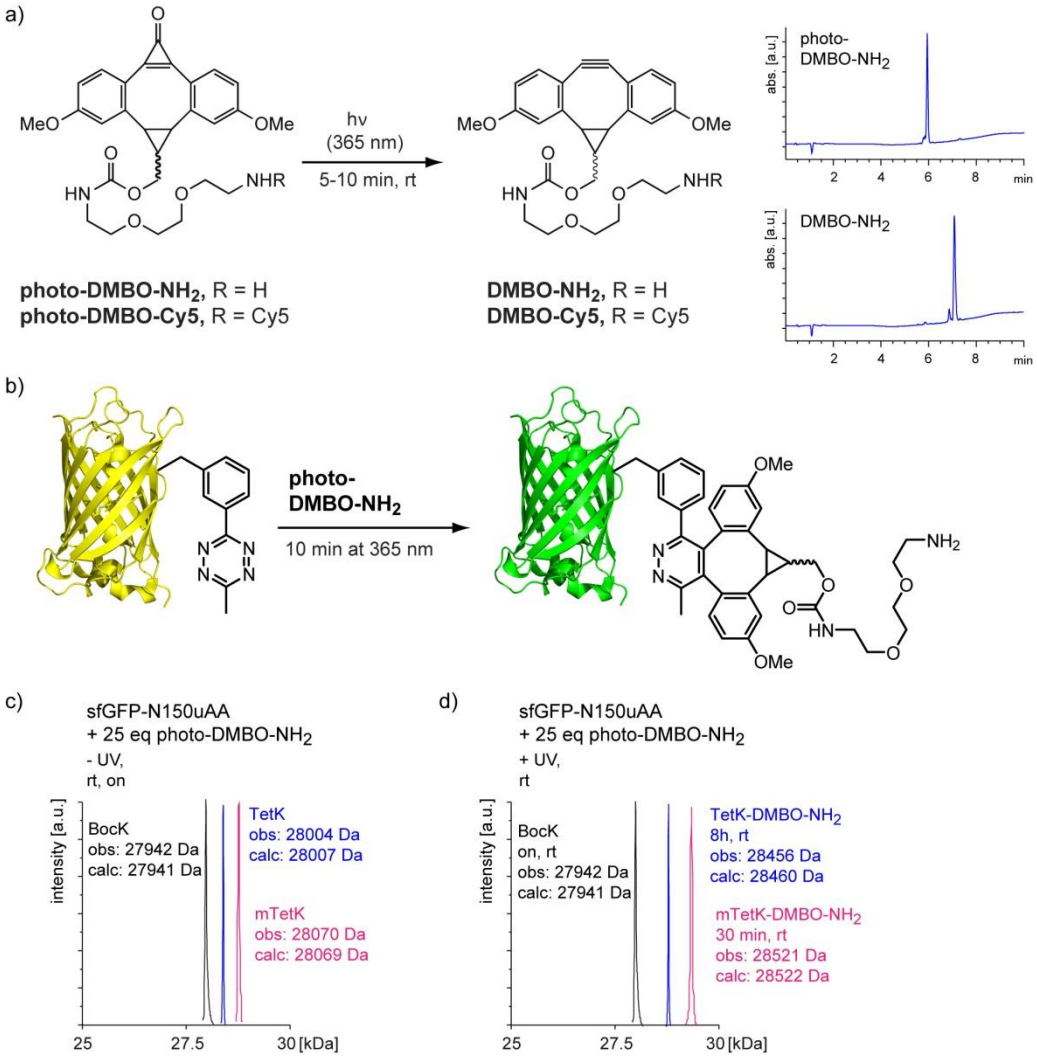


Figure 1.15: Characterization of photo-iEDDAC on tetrazine-bearing proteins.
a) Structures of water-soluble photo-DMBO conjugates. Both photo-DMBO-NH_2 and its Cy5-conjugate photo-DMBO-Cy5 are decarbonylated by short irradiation at 365 nm to quantitatively form DMBO-NH_2 and DMBO-Cy5, respectively. b) Schematic illustration of sfGFP-N150mTetK-His6 labeling with photo-DMBO-NH_2 upon irradiation c) ESI-MS shows no modification of sfGFP-N150uAA-His6 (uAA = Bock (black), TetK (blue), mTetK (pink)) with photo-DMBO-NH_2 when incubated in darkness. d) ESI-MS characterization of quantitative and rapid labeling of sfGFP-N150uAA-His6 (uAA = TetK (blue), mTetK (pink)) with photo-DMBO-NH_2 upon light irradiation. sfGFP-N150BocK-His6 (black) is not modified with DMBO-NH_2.

Furthermore, the incubation of sfGFP-N150BocK-His6 with photo-DMBO-NH_2 (-/+ UV) or DMBO-NH_2 did not lead to any labeling, confirming the selectivity of our photo-iEDDAC approach (Figure 1.15e). Quantitative and selective protein labeling was also observed for TetK-bearing proteins (sfGFP-N150TetK-His6) with a 25-fold excess of photo-DMBO-NH_2 upon activation at 365 nm, however the reaction proceeded more slowly and showed complete modification only after several hours (Figure 1.15d). To evaluate the reactivity of our new photo-DMBO compounds against previously reported cyclopropenone-caged cycloalkynes,[294] Toni synthesized a cyclopropenone-caged dibenzocyclooctyne (photo-DIBO, Figure 1.16a) without additional ring strain[296] and a recently reported dibenzoannulated and cyclopropenone-caged silicon-containing cycloheptyne (photo-Si, Figure 1.16a).[297] Modification of sfGFP-N150mTetK-His6 with either of these two compounds could not be observed upon irradiation at 365 nm even after overnight incubation (Figure 1.16b), although decarbonylation to DIBO could be detected by LC-MS.

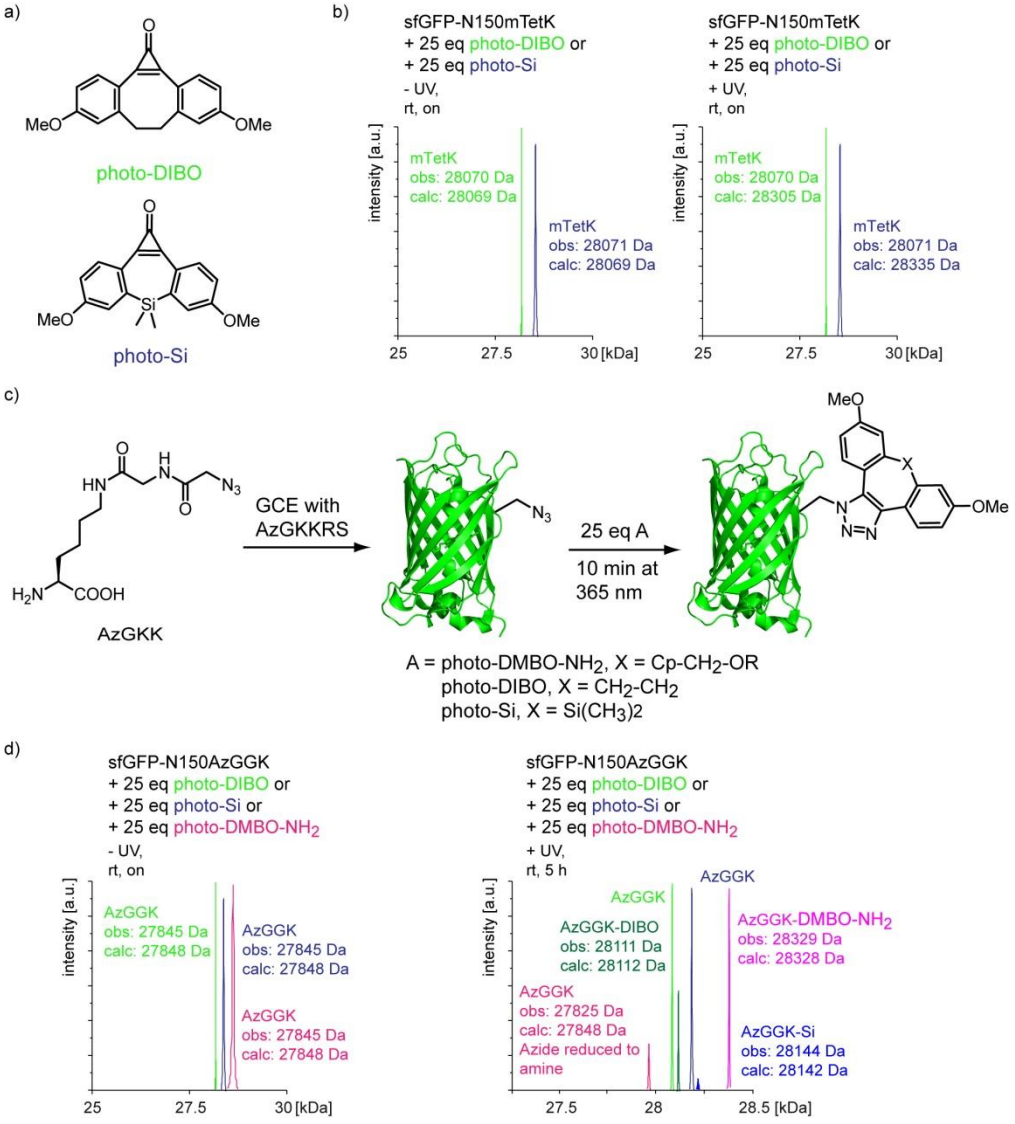


Figure 1.16: Comparison of reactivity of photo-DMBO-NH_2 with published compounds photo-DIBO and photo-Si.
a) Left: Structures of photo-DIBO and photo-Si. b) ESI-MS characterization of reactivity between sfGFP-N150mTetK-His6 (10µM) and 25-fold photo-DIBO or photo-Si without (left) and upon light activation (10 min at 365 nm, right). Samples were analyzed after 5 h or overnight. c) Schematic illustration of the photo-SPAAC reaction between genetically incorporated AzGGK and photo-DMBO-NH_2, photo-DIBO or photo-Si. d) ESI-MS characterization of labeling of sfGFP-N150mAzGGK-His6 (10µM) with 25-fold photo-DMBO-NH_2, photo-DIBO or photo-Si without (left) and upon light activation (10 min at 365 nm, right). Samples were analyzed after 5 h or overnight.

Since we based our DMBO compounds on a DIBO scaffold, which was developed for reactions with azides in SPAAC, as a next step reactivity of DMBO compounds, as well as the reported compounds photo-DIBO and photo-Si towards azides was tested. sfGFP bearing an azide uAA, AzGGK,[100] was subjected to labeling with a 25-fold excess of photo-DMBO-NH_2, photo-DIBO or photo-Si. Labeling with photo-DMBO-NH_2 upon activation at 365nm went to completion after ~3 hours, demonstrating higher reactivity towards tetrazines compared to azides (Figure 1.16c). Photo-DIBO in contrast yielded only 60 % triazole product overnight under the same conditions (Figure 1.16c), while photo-Si showed almost no product formation (< 5 % product), highlighting an enhanced reactivity of DMBO compounds in SPAAC and iEDDAC reactions due to its increased ring strain in addition. Furthermore, azide-modified amino acids have been observed to be reduced to the corresponding amine by endogenous reductases in the cytosol of *E. coli*, making our methyl-tetrazines the more reliable reporters. Encouraged by the short incubation times needed for quantitative iEDDAC labeling between DMBO compounds and sfGFP-N150mTetK-His6, rate constants were determined in a more quantitative way on-protein (Figure 1.17).

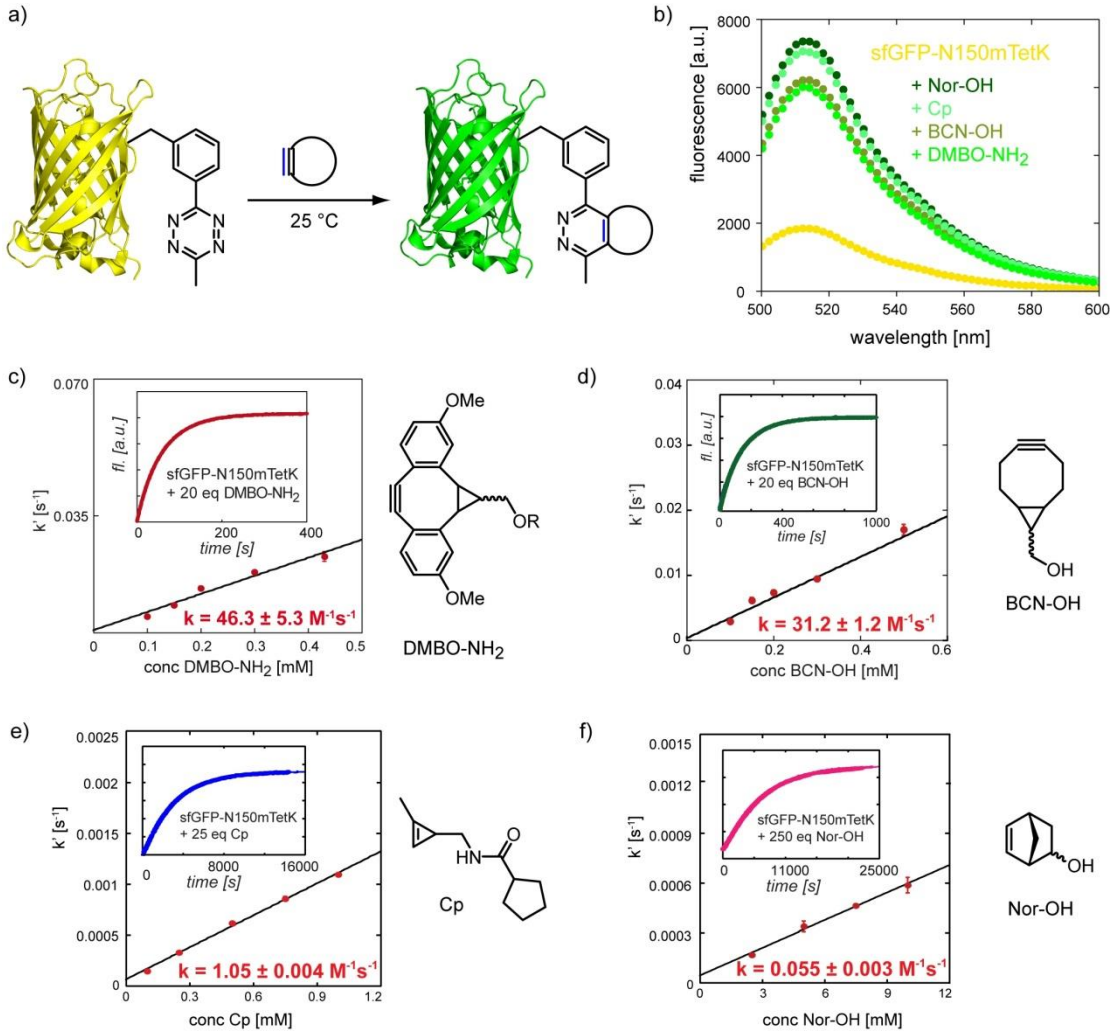


Figure 1.17: Determination of kinetics between sfGFP-N150mTetK-His6 and DMBO-NH_2 and other strained dienophiles (BCN-OH, Cp and Nor-OH).

a) Schematic illustration of the labeling reaction. b) Excitation at 460 nm produces quenched fluorescence for sfGFP-N150mTetK-His6, which is restored after reaction with DMBO-NH_2 and other strained dienophiles. d-f) Determination of second-order rate constants k of sfGFP-N150mTetK-His6 with either DMBO-NH_2, BCN-OH, Cp or Nor-OH. Inset shows fluorescence increase (508 nm) of sfGFP-N150mTetK-His6 upon addition of an excess of DMBO-NH_2 or the other strained dienophiles over time.

Mehl and coworkers noticed quenching of sfGFP fluorescence by genetic encoding of tetrazine amino acids Tet-v1.0 and Tet-v2.0 with Tet-v2.0RS into the proximity of the sfGFP chromophore (N150) and took advantage of this observation to determine iEDDAC reaction rates with sTCO on-protein *in vitro* and *in cellulo*.[56b, 151b] Incorporation of flexible tetrazine uAAs mTetK and TetK into sfGFP-N150TAG-His6 led to the same fluorescence quenching of sfGFP. Its recovery was triggered by the addition of DMBO-NH_2 (Figure 1.17b) and its exponential increase was followed over time. This enabled the quantification of second order on-protein rate constants under pseudo-first order conditions. A rate constant of ~50 $M^{-1}s^{-1}$ was determined (Figure 1.17c), which is nearly twice as fast as iEDDAC between mTetK and BCN-OH and 50 times faster than the corresponding reactions with 1,3-disubstituted cyclopropenes or even or 1000-fold faster with norbornenol (Figure 1.17d-f), dienophiles that are commonly used for labeling of biomolecules.[147c, 192, 198b, 198c] iEDDAC between sfGFP-N150mTetK-His6 and DMBO-NH_2 was also significantly faster than reaction between DMBO-NH_2 and sfGFP bearing the previously reported uAA Tet-v2.0 (TetF)[151b], probably due to greater flexibility of mTetK, resulting in higher on-protein accessibility compared to the more rigid TetF (Figure 1.18).

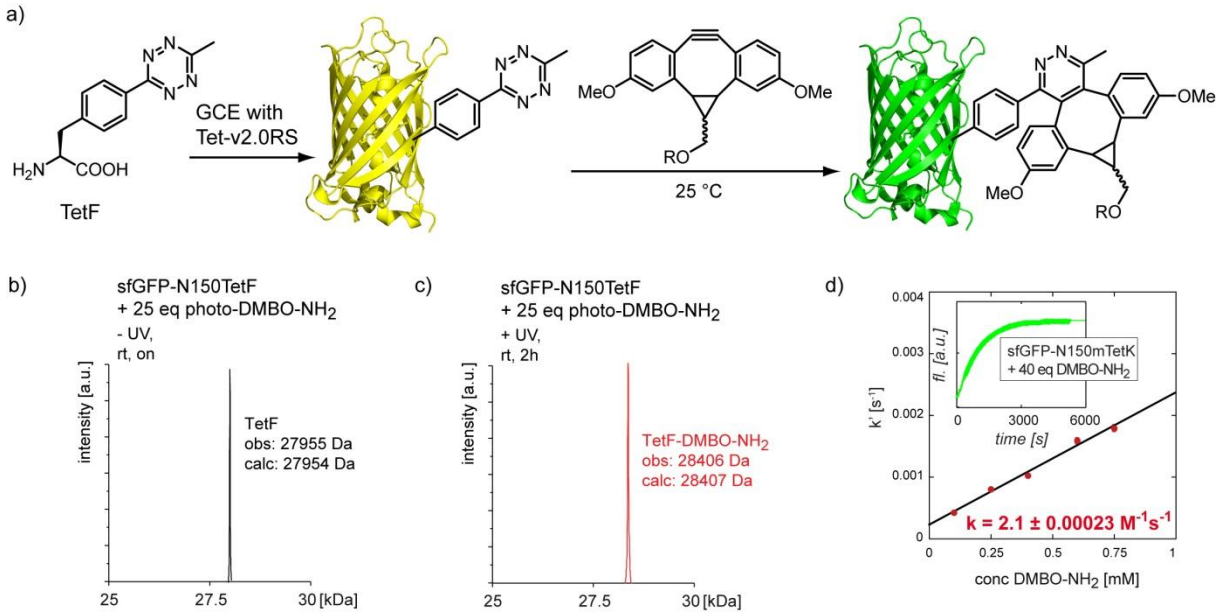


Figure 1.18: Characterization of iEDDAC between sfGFP-N150TetF-His6 and DMBO-NH_2.
a) Schematic illustration of incorporation of TetF via GCE and subsequent labeling with DMBO-NH_2 b) ESI-MS confirms incorporation of TetF into sfGFP-N150TetF-His6 and no modification with photo-DMBO-NH_2 in the dark was observed. c) ESI-MS characterization of the selective photo-iEDDAC between sfGFP-N150TetF-His6 and photo-DMBO-NH_2: only upon activation by light, complete formation of modified sfGFP is observed after 2h. d) Determination of second-order rate constant k of sfGFP-N150TetK-His6 and DMBO-NH_2. Inset shows fluorescence increase (508 nm) of sfGFP-N150TetF-His6 upon addition of a 40-fold excess of DMBO-NH_2 over time.

Our new photo-induced iEDDAC approach allows the quantitative and chemoselective labeling of tetrazine-bearing proteins with photo-DMBO-NH_2 only upon short irradiation at 365 nm, going to completion within a few minutes at low µM concentrations. Due to the fast reaction kinetics our photo-iEDDAC is equivalent with the fastest photo-induced bioorthogonal reactions (photo-SPAAC with photo-ODIBO (~40 $M^{-1}s^{-1}$), [298] and significantly faster than photo-iEDDAC reactions employing caged cyclopropenes (~10^{-2}-10^{-4} $M^{-1}s^{-1}$)[299] or photo-SPAAC reactions with strained dibenzocyclooctynes (~ 10^{-2}-10^{-1} $M^{-1}s^{-1}$).[300]

1.3.6 Photo-iEDDAC confers spatio-temporal control to protein labeling on living *E. coli*

To enable a fluorescent read-out of our photo-iEDDAC labeling, Anton synthesized photo-DMBO-Cy5 by coupling a succinimidyl ester of Cy5 dye to photo-DMBO-NH_2 (Figure 1.19a). Irradiation of photo-DMBO-Cy5 at 365 nm resulted in quantitative decarbonylation to the corresponding DMBO-Cy5 compound within 10 min (Figure 1.19a), which reacted rapidly with purified sfGFP-N150mTetK-His6 after administering in 25-fold excess to the cycloaddition product, confirmed by ESI-MS (Figure 1.19b). SDS-PAGE-based fluorescence imaging following this reaction in a time-resolved manner illustrates the rapid labeling of purified sfGFP-N150mTetK-His6 in more detail (Figure 1.19c). Performing labeling reactions with photo-DMBO-Cy5 in *E. coli* cell lysate shows specific and selective modification of mTetK-bearing proteins upon irradiation at 365 nm with minimal background (Figure 1.19d).

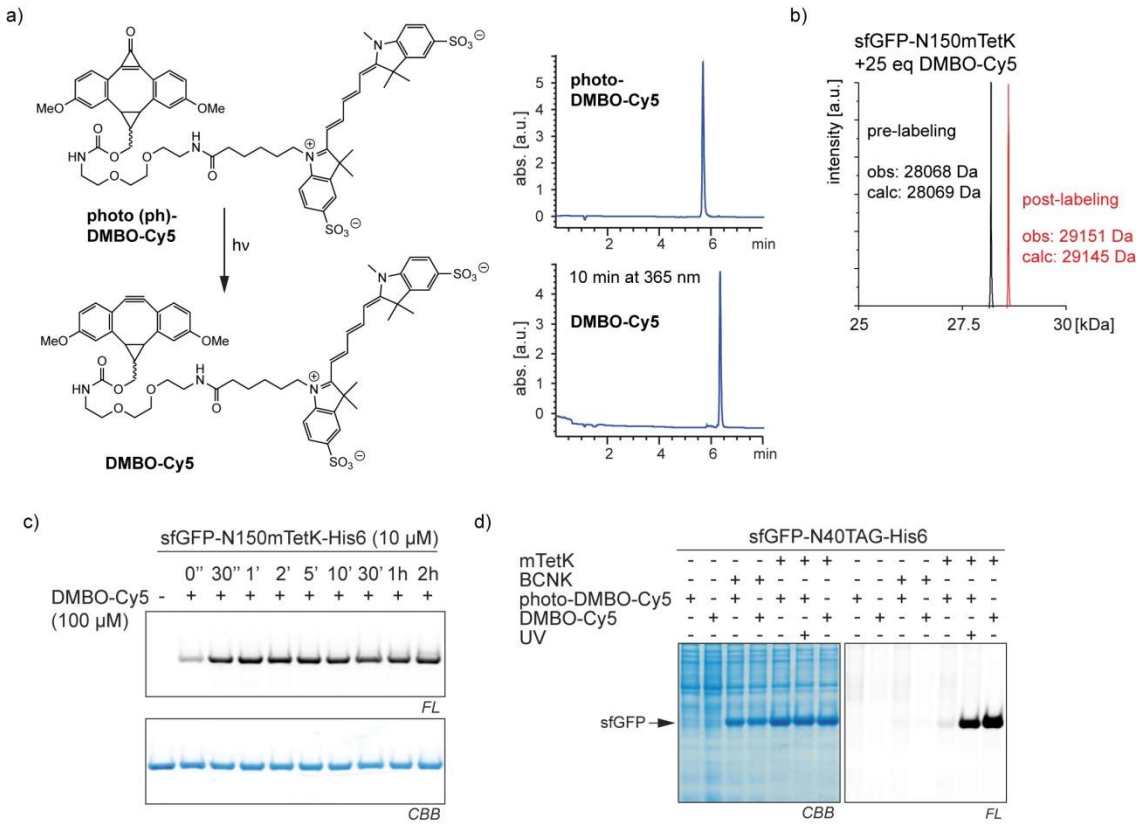


Figure 1.19: Specific and rapid labeling of mTetK-bearing proteins with photo-DMBO-Cy5.
a) Left: Structure of the Cy5 conjugates photo-DMBO-Cy5 and DMBO-Cy5. Right: LC-MS analysis shows quantitative photo-decaging of photo-DMBO-Cy5 at 365 nm. b) ESI-MS analysis proves quantitative labeling of sfGFP-N150mTetK-His6 with DMBO-Cy5. c) SDS-PAGE in gel fluorescence confirms rapid labeling of sfGFP-N150mTetK-His6 with DMBO-Cy5. d) Selective and light-induced labeling of mTetK-bearing sfGFP with photo-DMBO-Cy5 towards the *E. coli* proteome. CBB: Coomassie brilliant blue; FL: fluorescence.

Encouraged by specific modification of mTetK-bearing proteins in cell lysate, we wanted to use photo-DMBO-Cy5 for labeling of living *E. coli* under temporal control by light. For this, OmpC-Y232mTetK was over-expressed on the surface of *E. coli* cells, residual mTetK removed by several washing steps and the resulting culture subjected to labeling with photo-DMBO-Cy5 either in the presence or absence of UV-light. Labeling of OmpC was analyzed by SDS-PAGE fluorescence imaging and fluorescence microscopy of living *E. coli* (Figure 1.20). In gel fluorescence was observed for OmpC-Y232mTetK treated with DMBO-Cy5 or photo-DMBO-Cy5 only upon irradiation at 365 nm. OmpC bearing strained alkyne BCNK

(Figure 1.20b) was used as a negative control and was, as expected, not observed to be modified (Figure 1.20a). Microscopy images show *E. coli* cells either expressing OmpC-Y232mTetK or OmpC-Y232BCNK in the bright field. Emission of red fluorescence was observed only upon treatment of OmpC-Y232mTetK with DMBO-Cy5 as a positive control or photo-DMBO-Cy5 in the presence of UV light in the latter case. To show that OmpC-Y232BCNK is able to be modified with the corresponding reaction partner, it was labeled with a methyl-tetrazine-Cy5 conjugate (Tet-Cy5, Figure 1.20b) and the labeling visualized by fluorescence microscopy (Figure 1.20c). These experiments confirm the selectivity of our novel photo-iEDDAC reaction with mTetK-bearing proteins on the cell-surface of living *E. coli*.

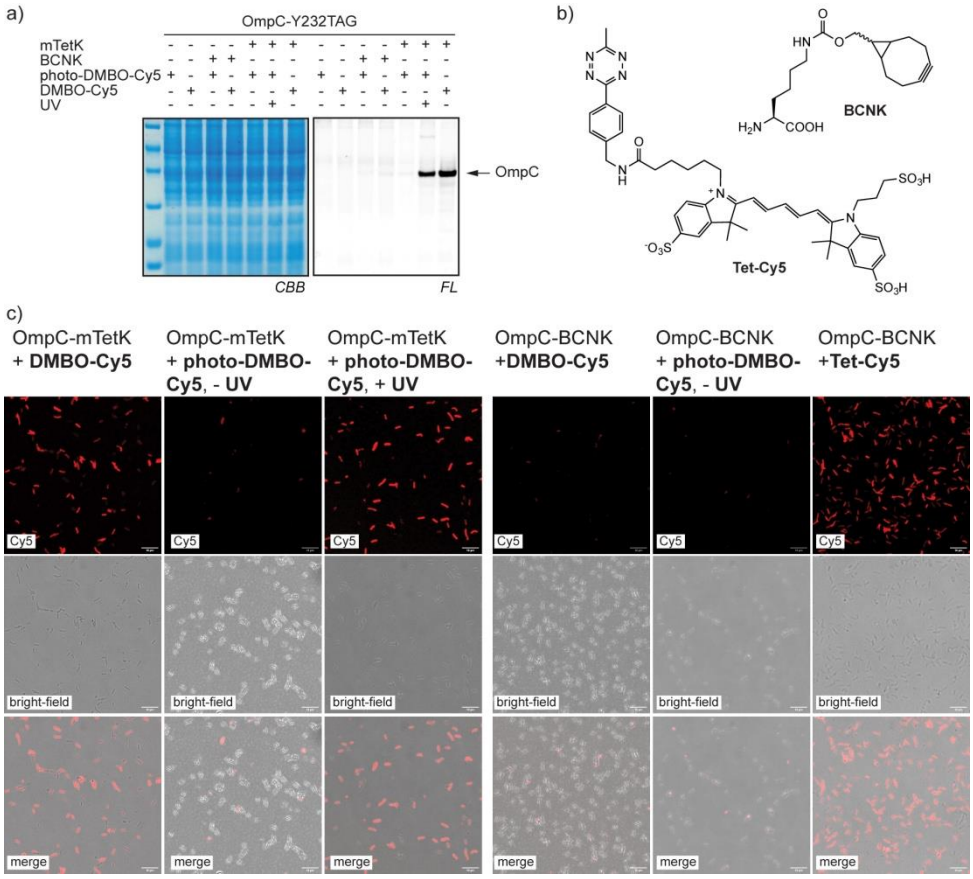


Figure 1.20: Photo-iEDDAC labeling on living *E. coli* cells.
a) In gel fluorescence imaging shows specific and selective labeling of OmpC-Y232mTetK on the surface of living *E. coli* only upon irradiation with 365 nm. b) Structures of BCNK and tetrazine-Cy5 conjugated Tet-Cy5. c) Fluorescence microscopy shows efficient and light-induced labeling of a cell-surface protein OmpC-Y232mTetK with photo-DMBO-Cy5 on living *E. coli*. Upper panels show red fluorescence in the Cy5 channel, middle panels show bright field images, bottom panels show merged images. CBB: Coomassie brilliant blue; FL: fluorescence.

As photo-DMBO remains strictly inactive until induced by UV-irradiation, photo-iEDDAC can be applied in one-pot reactions with iEDDAC for sequential dual labeling of proteins. To show this mutual orthogonality, a heterobifunctional crosslinker was designed to combine two moieties with comparable reactivities in iEDDAC, BCN- and DMBO, in one compound photo-DMBO-BCN for sequential one-pot labeling. This molecule enables controlled reaction with different methyl-tetrazine conjugates. *E. coli* cells expressing mTetK-bearing OmpC on their cell surface were treated with photo-DMBO-BCN, which lead to modification via its BCN moiety. The photo-DMBO functionality in the heterobifunctional compound however, is

inactive towards tetrazine ligation prior to UV-irradiation and cells are labeled with Tet-Cy5 only after light-activation (Figure 1.21), conferring mutual orthogonality to iEDDA cycloadditions with tetrazines and providing control over protein-labeling both in time and space.[301]

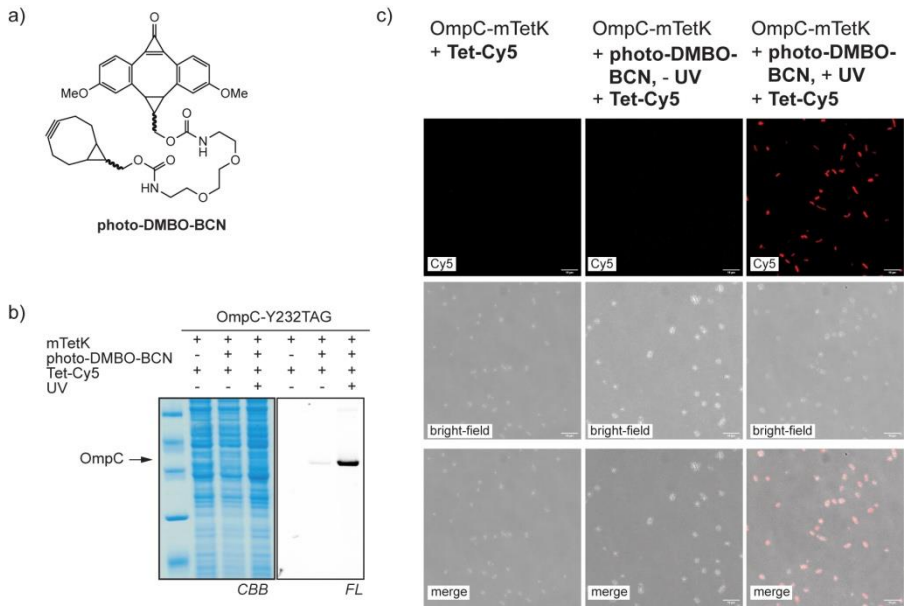


Figure 1.21: Sequential photo-iEDDAC labeling on living *E. coli* cells.
a) Structure of bifunctional linker photo-DMBO-BCN, displaying a BCN and a photo-DMBO moiety. DMBO moiety becomes only available upon UV-irradiation and thereby allows spatio-temporal controlled, sequential labeling of living *E. coli*. b) In gel fluorescence imaging and c) fluorescence imaging of living *E. coli* cells show that labeling of OmpC-Y232mTetK on the cell surface is depending on treatment with photo-DMBO-BCN and its activation with UV light.

1.3.7 Incorporation experiments with tetrazine amino acids in mammalian cells

Having a PylRS mutant at hand that incorporates TetK and mTetK effectively into proteins in *E. coli*, as a next step the expansion of our new tetrazine uAAs to mammalian cells was planned. For this, Marie-Lena Jokisch first tested different mammalian cell lines HEK293T, COS-1, COS-7 and HeLa for their amber suppression efficiency. Therefore, she coexpressed sfGFP carrying a premature amber codon and a C-terminal His6-tag with the wt PylRS in the presence of BocK in these different cell lines and evaluated the efficiency by production of fluorescent, full-length sfGFP. Among these, HEK293T cells showed the best amber suppression efficiency of sfGFP, while all the other cell lines showed a significantly decreased to no fluorescence signal. Expression tests with TetK and mTetK were then performed using HEK293T cells, but did unfortunately did not show incorporation of either TetK or mTetK into sfGFP (Figure 1.22). mTetK could only be used in low mM concentrations, since it is prone to precipitation in aqueous media at concentrations > 0.5 mM. The use of K-mTetK, which was better water-soluble than mTetK did also not give better incorporation results and the used batch was still contaminated with small amounts of 3-CNPhK. The uAAs BocK and BCNK were used as positive controls for the incorporation.

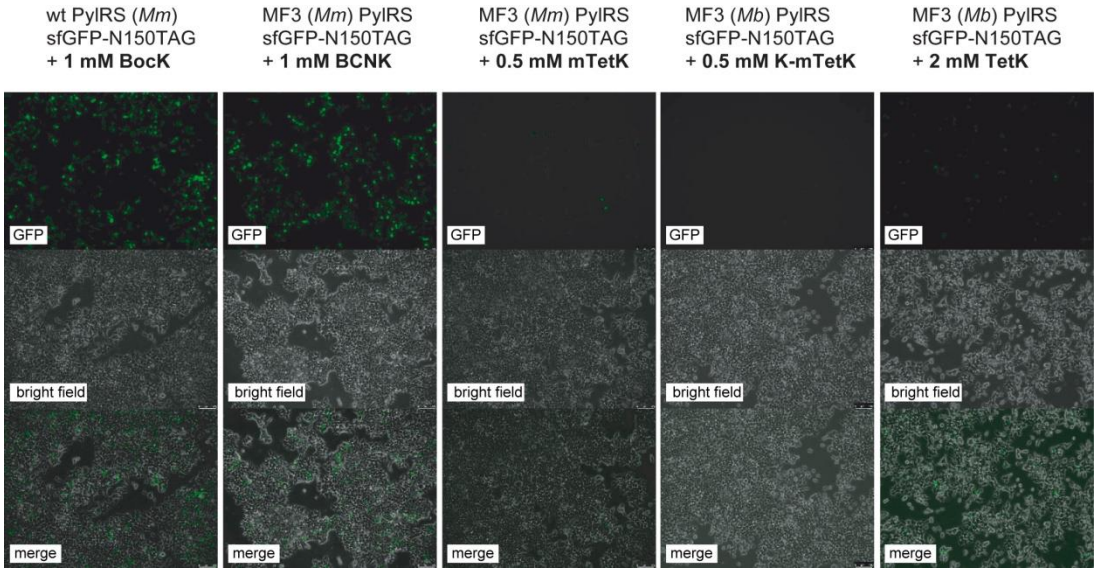


Figure 1.22: Incorporation of different uAAs into sfGFP-N150TAG-His6 in HEK293T cells.

To test if the tetrazine amino acids are not incorporated into proteins, because they are not able to penetrate cells, Marie-Lena then performed cell-permeability assays, where sfGFP-N150TAG-His6 was expressed in the presence of BCNK and then treated with uAAs TetK, mTetK or K-mTetK. Cells were washed, lysed, sfGFP purified via His6-tag and then subsequently analyzed by ESI-MS, which confirmed modification of sfGFP-N150BCNK with TetK or mTetK (Figure 1.23).

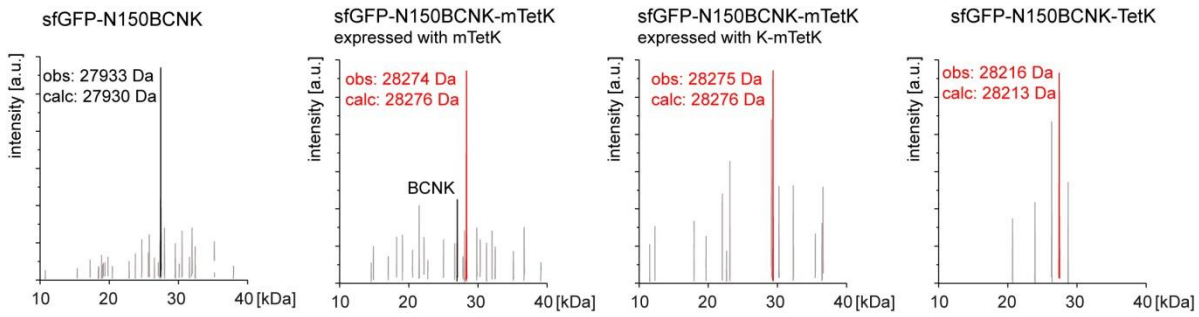


Figure 1.23: Uptake of tetrazine uAAs mTetK, K-mTetK and TetK is analyzed by ESI-MS.

As a next step, we investigated if the incorporation of the tetrazine uAAs into other proteins would give better results. For this, Marie-Lena tested different proteins bearing premature TAG codons for the incorporation of TetK and mTetK including a GFP-mCherry fusion, as well as the epidermal growth factor receptor fusion EGFR-EGFP and the interleukin 4 receptor IL-4R; however, successful suppression was not observed for these proteins. During the poster session at a conference someone shared the hypothesis that tetrazine amino acids are reduced to the corresponding dihydrotetrazines, incorporated into proteins as such and then reoxidized on the protein. We investigated this first in *E. coli*, where a 4-fold excess of 2-(diphenylphosphino)benzoic acid (2-DPBA) was added to cultures expressing sfGFP-N150TAG-His6 in the presence of 0.5 mM TetK and mTetK (Figure 1.24a). The addition of 2-DPBA to *E. coli* cultures negatively influence their growth, as seen for the culture without addition of uAA, where also less truncated sfGFP was produced. Furthermore, the addition of 2-DPBA to cultures where tetrazine uAA was present also did not show production of full-

length sfGFP anymore. To determine, if this is mainly due to the negative influence of 2-DPBA to *E. coli* cells, mTetK was reduced to the corresponding dihydrotetrazine dihydro-mTetK (dh-mTetK) by addition of an excess of 2-(diphenylphosphino)benzoic acid and purified by HPLC. Dh-mTetK was tested for incorporation into proteins in *E. coli*, with and without the addition of 2-DPBA, but full-length sfGFP was observed in neither case (Figure 1.24b). Marie-Lena also tested the combination of tetrazine uAAs and 2-DPBA in mammalian cells and could not observe successful incorporation of TetK or mTetK, suggesting that tetrazine uAAs are incorporated into proteins in their oxidized form.

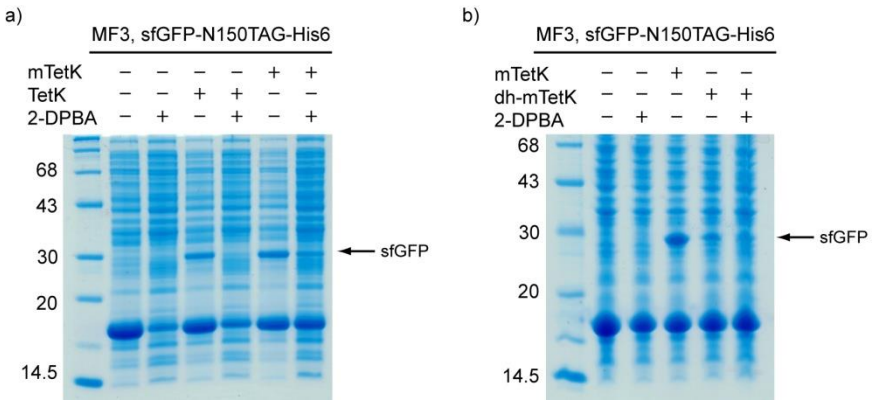


Figure 1.24: Incorporation of dihydrotetrazine uAAs into sfGFP-N150TAG-His6.
a) Coomassie stained SDS-PAGE of incorporation of TetK and mTetK into sfGFP-N150TAG-His6 after *in situ* reduction to corresponding dihydrotetrazines by 2-DPBA. b) Coomassie stained SDS-PAGE of incorporation dihydro-mTetK into sfGFP-N150TAG-His6.

One hypothesis that emerged during these experiments employing 2-DPBA was that the tetrazines uAAs are actually reduced to the corresponding dihydrotetrazines upon uptake and therefore are not incorporated into proteins anymore. To examine this, we applied an approach described by Fox and coworkers that uses the photosensitizer methylen blue in the presence of visible light to induce iEDDAC between dihydrotetrazines and TCO-conjugates.[262b] Marie-Lena tested this approach in mammalian cells. Therefore, cells were treated with 0.5 mM TetK in the presence of 10 µM methylen blue and irradiated at 660 nm, which completely abolished their viability, probably due to a combination of methylen blue and irradiation, as cells tolerated treatment with either one of these two factors.

To date it is not completely understood, why the incorporation of tetrazine uAAs is not successful in mammalian cells. As described before, several hypotheses for this were examined, but could not be confirmed by experimental data. Further experiments in mammalian cells to elucidate the underlying reasons are currently under way.

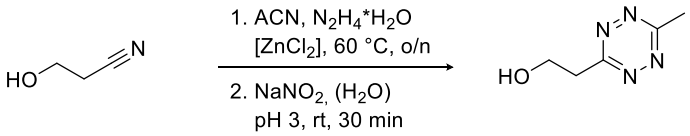
1.4 Summary and Outlook

In conclusion, we have reported a novel light-induced iEDDAC between tetrazines and a cyclopropenone caged dibenzoannulated BCN probe (photo-DMBO). We have demonstrated the efficient site-specific incorporation of methyl-tetrazine modified amino acids TetK and mTetK into proteins in *E. coli* and their efficient, specific and light-triggered labeling with photo-DMBO fluorophore conjugates. Photo-iEDDAC proceeds with on-protein rate-constants of ~50 $M^{-1}s^{-1}$, equivalent to the fastest bioorthogonal photo-induced reactions, enabling protein labeling within minutes at low µM concentrations. These reactions are 2-4 orders of magnitude faster than iEDDAC using cyclopropene or norbornene as dienophiles and nearly twice as fast as cycloadditions between BCN and metabolically stable methyl-tetrazines. In the future other fields such as the fabrication of microarrays,[302] biosensors and the preparation of multifunctional material may benefit from this photo-triggered and rapid iEDDAC reaction. While we have demonstrated the advantages of our approach *in vitro* and in living *E. coli*, the ability to incorporate uAAs in mammalian cells and *C. elegans*,[303] zebrafish[304] and *D. melanogaster*[305] using engineered PylRS/$tRNA_{CUA}$ pairs suggests that it may be possible to extend the photo-induced labeling approach presented herein to site-specifically label proteins in animals, which might be especially attractive in combination with multi-photon activation of photo-DMBO compounds.[306] Exploitation of photo-iEDDAC and genetically encodable tetrazine reporters in mammalian cell is focus of current studies in our lab.

1.5 Experimental

1.5.1 Synthesis of TetK

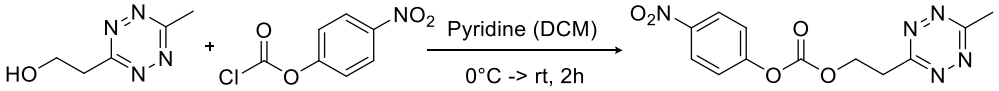
3-(2-Hydroxyethyl)-6-methyl-1,2,4,5-tetrazine[187]



Hydroxypropionitrile (5.0 g, 70.3 mmol, 1 eq) was added to a flask containing zinc chloride (4.8 g, 35.2 mmol, 0.5 eq), followed by acetonitrile (37 mL, 7.03 mol, 10 eq) and hydrazine hydrate (90 mL, 1.7 mol, 25 eq). The resulting reaction mixture was stirred at 60 °C for 36 h. Then, sodium nitrite (24.3 g, 0.35 mol, 5 eq) dissolved in a minimal amount of water was added and the mixture cooled on ice. 6 M HCl solution (300 mL) was slowly added to acidify the reaction mixture to pH 3. It was stirred for 30 minutes on ice until the formation of nitrous fumes was completed, then the water phase was extracted with EtOAc (5x 200 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 before removing the solvent under reduced pressure. The crude residue was purified by flash chromatography (0-60 % EtOAc/pentane), which resulted the product (1.11 g, 11 %) as a red liquid.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 3.06 (s, 3H, CH_3), 3.57 (t, $^3J_{HH}$ = 5.8 Hz, 2H, CH_2), 4.26 (t, $^3J_{HH}$ = 5.8 Hz, 2H, CH_2OH).

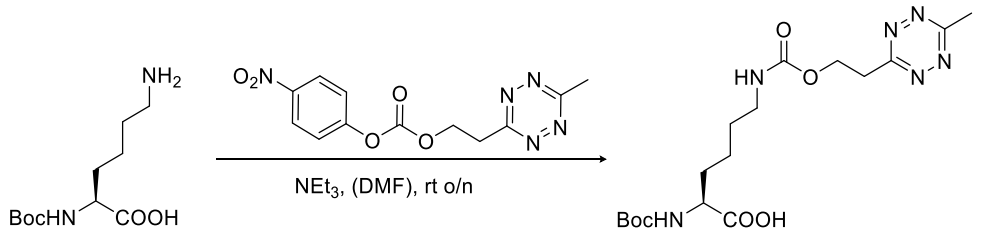
2-(6-methyl-1,2,4,5-tetrazin-3-yl)ethyl-(4-nitrophenyl)-carbonate



3-(2-Hydroxyethyl)-6-methyl-1,2,4,5-tetrazine (0.11 g, 0.79 mmol, 1 eq) was dissolved in DCM (2 mL) and cooled to 0 °C. 4-Nitrophenyl chloroformate (0.35 g, 1.7 mmol, 2.2 eq) was added to the solution in small portions, followed by the addition of pyridine (70 µL, 0.86 mmol 1.1 eq) and then stirred at room temperature for two hours. After removal of the solvent under reduced pressure flash chromatography was used to purify the crude mixture (0-30 % EtOAc/pentane), which resulted the product (0.2 g, 83 %) as a red liquid.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 3.08 (s, 3H, CH_3), 3.78 (t, $^3J_{HH}$ = 6.2 Hz, 2H, CH_2), 4.98 (t, $^3J_{HH}$ = 6.2 Hz, 2H, CH_2O), 7.35 (d, $^3J_{HH}$ = 9.2 Hz, 2H, H_{Ar}), 8.27 (d, $^3J_{HH}$ = 9.2 Hz, 2H, H_{Ar}).

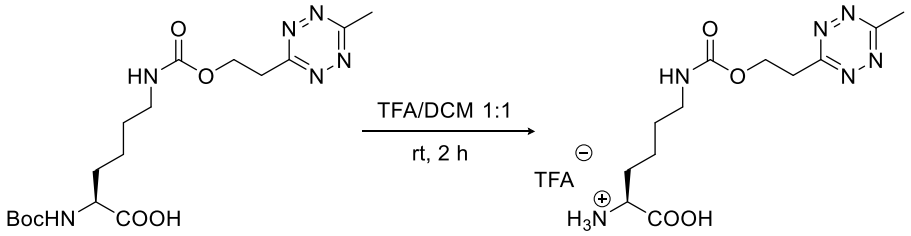
N^2-(*tert*-butoxycarbonyl)-$N6$-((2-(6-methyl-1,2,4,5-tetrazin-3-yl)ethoxy)-carbonyl)-L-lysine (BocTetK)



Boc-L-lysine (1.84 g, 7.45 mmol, 1.2 eq) was dissolved in DMF (20 mL), then NEt_3 (1 mL, 7.45 mmol, 1.2 eq) and 2-(6-methyl-1,2,4,5-tetrazin-3-yl)ethyl-(4-nitrophenyl)-carbonate (1.90 g, 6.21 mmol, 1 eq) were added to the solution, which was stirred at room temperature overnight. Afterwards the solvent was evaporated under reduced pressure and the crude residue purified by flash chromatography (0-10 % MeOH/DCM + 0.1 % AcOH) to result the product as a red oil (1.94 g, 73 %).

^{1}H-NMR (500 MHz, DMSO-d_6): δ = 1.19-1.36 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.37 (s, 9H, Boc), 1.46-1.64 (m, 2H, $CH_{2\text{-}\beta}$), 2.84-2.90 (m, 2H, $CH_{2\text{-}\varepsilon}$), 2.95 (s, 3H, CH_3), 3.52 (t, $^3J_{HH}$ = 6.3 Hz, 2H, CH_2), 3.76-3.84 (m, 1H, $CH_{\text{-}\alpha}$), 4.49 (t, $^3J_{HH}$ = 6.3 Hz, 2H, CH_2O), 6.99 (d, $^3J_{HH}$ = 7.9 Hz, 1H, $NH_{\text{-}\alpha}$), 7.09 (t, $^3J_{HH}$ = 5.8 Hz, 1H, $NH_{\text{-}\varepsilon}$).
MS (ESI), *m/z* calc for $C_{17}H_{28}N_6O_6$ 412.21,
found 435.0 $[M+Na]^+$, 312.8 $[M\text{-}Boc+H]^+$, 411.1 $[M-H]^-$

N^6-((2-(6-methyl-1,2,4,5-tetrazin-3-yl)ethoxy)-carbonyl)-L-lysine trifluoroactate (TetK)

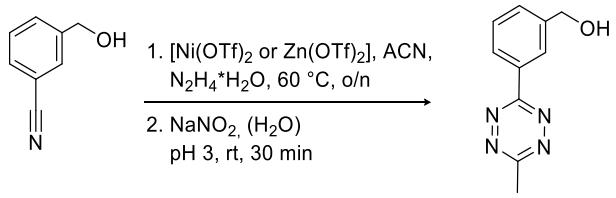


BocTetK (1.90 g, 5 mmol, 1 eq) was dissolved in 25 mL TFA/DCM (1:1) with a drop of H_2O and stirred at room temperature for two hours. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (1.41 g, 75 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.20-1.43 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.64-1.80 (m, 2H, $CH_{2\text{-}\beta}$), 2.84-2.93 (m, 2H, $CH_{2\text{-}\varepsilon}$), 2.95 (s, 3H, CH_3), 3.52 (t, $^3J_{HH}$ = 6.3 Hz, 2H, CH_2), 3.80-3.93 (m, 1H, $CH_{\text{-}\alpha}$), 4.49 (t, $^3J_{HH}$ = 6.3 Hz, 2H, CH_2O), 7.08 (t, $^3J_{HH}$ = 5.0 Hz, 1H, $NH_{\text{-}\varepsilon}$), 8.20 (s, 3H, $NH_{\text{-}\alpha}$).
^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 20.8 (1C, CH_3), 21.8 (1C, C_γ), 28.8 (1C, C_δ), 29.7 (1C, C_β), 34.4 (1C, CH_2), 39.9, (1C, C_ε), 52.1 (1C, C_α), 61.3 (1C, CH_2O), 156.0 (1C, NCOO), 167.1 (1C, $C_{Tet\text{-}6}$), 167.5 (1C, $C_{Tet\text{-}3}$), 171.2 (1C, COOH).
RP-HPLC (C18, 250 x 21.2 mm, 10 mL/min, 1-80 % ACN in 16 min): t_R = 12-13.5 min
MS (ESI), *m/z* calc for $C_{12}H_{20}N_6O_4$ 312.15, found 313.2 $[M+H]^+$, 311.1 $[M-H]^-$

1.5.2 Synthesis of mTetK

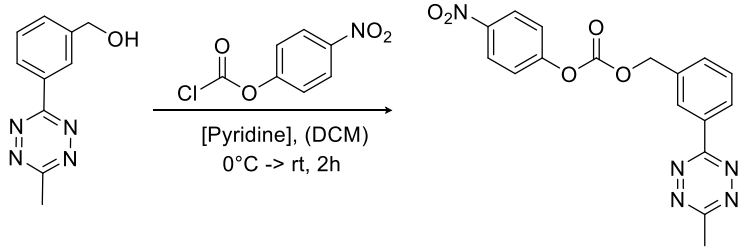
3-(3-(Hydroxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine



3-(Hydroxymethyl)benzonitrile (1 g, 7.5 mmol, 1 eq) was combined in flask with Zn or $Ni(OTf)_2$ (1.4 or 1.3 g, 3.8 mmol, 0.5 eq), then ACN (3.9 mL, 75 mmol, 10 eq) and hydrazine hydrate (9.1 mL, 0.19 mol, 25 eq) were added and the reaction heated to 60 °C for 24-36 h. $NaNO_2$ (2.6 g, 38 mmol, 5 eq) dissolved in H_2O (10 mL) was added and the recation mixture cooled in an ice bath before slowly acidifying to pH 3 using a 6 M HCl solution. After stirring for 30 minutes on ice until the formation of nitrous fumes was complete, the reaction was extracted with EtOAc (5x 80 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude residue was purified by flash chromatography (0-50 % EtOAc/pentane), which resulted the product (Zn:0.53 g, 35 %; Ni: 0.56 g, 38 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 3.00 (s, 3H, CH_3), 4.65 (d, $^3J_{HH}$ = 5.8 Hz, 2H, CH_2), 5.42 (t, $^3J_{HH}$ = 5.8 Hz, 1H, OH), 7.60-7.64 (m, 2H, $H_{Ar\text{-}4,6}$), 8.32-8.37 (m, 1H, $H_{Ar\text{-}5}$), 8.45-8.48 (m, 1H, $H_{Ar\text{-}2}$).
MS (ESI), *m/z* calc for $C_{10}H_{10}N_4O$ 202.09, found 202.7 $[M+H]^+$

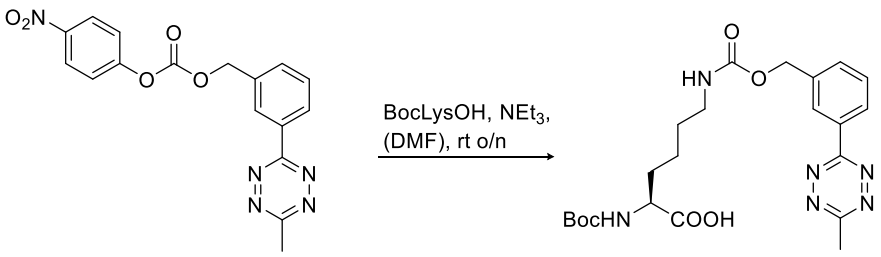
3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate



3-(3-(Hydroxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine (2.2 g, 11 mmol, 1 eq) was dissolved in DCM (50 mL) and cooled to 0 °C. 4-Nitrophenyl chloroformate (2.6 g, 13 mmol, 1.2 eq) was added to the solution, followed by the addition of pyridine (0.88 mL, 11 mmol 1 eq) and then stirred at room temperature for 2 hours. After removal of the solvent under reduced pressure flash chromatography was used to purify the crude mixture (0-60 % EtOAc/pentane), which resulted the product (3.5 g, 88 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 3.02 (s, 3H, CH_3), 5.48 (s, 2H, CH_2), 7.60 (d, $^3J_{HH}$ = 9.2 Hz, 2H, $H_{Ar\text{-}meta NO2}$), 7.75 (dd, $^3J_{HH}$ = 7.7, 7.7 Hz, 1H, $H_{Ar\text{-}5}$), 7.80 (td, $^3J_{HH}$ = 7.7 Hz, $^4J_{HH}$ = 1.5 Hz, 1H, $H_{Ar\text{-}4}$), 8.33 (d, $^3J_{HH}$ = 9.2 Hz, 2H, $H_{Ar\text{-}ortho NO2}$), 8.49 (td, $^3J_{HH}$ = 7.7 Hz, $^4J_{HH}$ = 1.5 Hz, 1H, $H_{Ar\text{-}6}$), 8.57-8.60 (m, 1H, $H_{Ar\text{-}2}$).

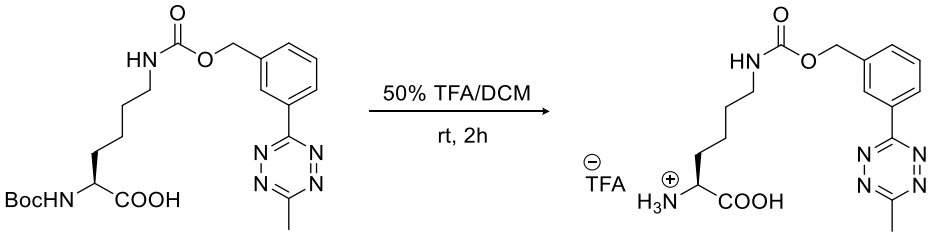
N^2-(*tert*-butoxycarbonyl)-N^6-(((3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine (BocmTetK)



Boc-L-lysine (0.32 g, 1.3 mmol, 1.2 eq) was dissolved in DMF (3.5 mL), then NEt_3 (0.18 mL, 1.3 mmol, 1.2 eq) and 3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate (0.4 g, 1.1 mmol, 1 eq) were added to the solution, which was stirred at room temperature overnight. Afterwards the solvent was evaporated under reduced pressure and the crude residue purified by flash chromatography (0-2 % MeOH/DCM + 0.1 % AcOH) to result the product as a pink solid (0.36 g, 70 %).

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.22-1.54 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.36 (s, 9H, Boc), 1.47-1.69 (m, 2H, $CH_{2\text{-}\beta}$), 2.94-3.00 (m, 2H, $CH_{2\text{-}\varepsilon}$), 3.01 (s, 3H, CH_3), 3.75-3.86 (m, 1H, $CH_{2\text{-}\alpha}$), 5.17 (s, 2H, CH_2), 7.00 (d, $^3J_{HH}$ = 8.1 Hz, 1H, NH_α), 7.35 (t, $^3J_{HH}$ = 5.6 Hz, 1H, NH_ε), 7.62-7.68 (m, 2H, $H_{Ar\text{-}4,6}$), 8.38-8.43 (m, 1H, $H_{Ar\text{-}5}$), 8.44-8.47 (m, 1H, $H_{Ar\text{-}2}$).
MS (ESI), *m/z* calc for $C_{22}H_{30}N_6O_6$ 474.22
found 497.0 $[M+Na]^+$, 375.0 $[M\text{-}Boc+H]^+$, 473.1 $[M-H]^-$

N^6-(((3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine trifluoroacetate (mTetK)



BocmTetK (6.8 g, 14.3 mmol, 1 eq) was dissolved in 60 mL TFA/DCM (1:1) with a drop of H_2O and stirred at room temperature for two hours. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (5 g, 75 %) as a pink solid. The amino acid was further purified by RP-HPLC (C18, 250 x 21.2 mm, 10 mL/min, 1-80 % ACN in 16 min).

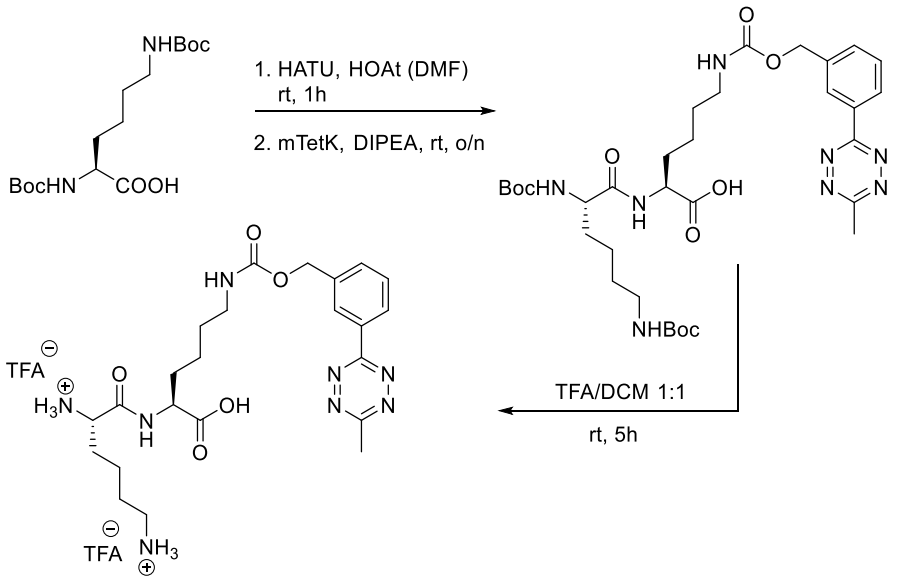
^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.27-1.53 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.67-1.84 (m, 2H, $CH_{2\text{-}\beta}$), 2.93-3.09 (m, 5H, CH_3, $CH_{2\text{-}\varepsilon}$), 3.82-3.93 (m, 1H, $CH_{2\text{-}\alpha}$), 5.17 (s, 2H, CH_2), 7.34 (t, $^3J_{HH}$ = 5.6 Hz, 1H, NH_ε), 7.62-7.68 (m, 2H, $H_{Ar\text{-}4,6}$), 8.22 (s, 3H, NH_α), 8.38-8.43 (m, 1H, $H_{Ar\text{-}5}$), 8.44-8.47 (m, 1H, $H_{Ar\text{-}2}$).

^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 20.9 (1C, CH_3), 21.7 (1C, C_γ), 28.9 (1C, C_δ), 29.7 (1C, C_β), 39.8, (1C, C_ε), 51.6 (1C, C_α), 64.8 (1C, CH_2), 126.4 (1C, $C_{Ar\text{-}2}$), 126.8 (1C, $C_{Ar\text{-}5}$), 129.6 (1C, $C_{Ar\text{-}6}$), 131.6 (1C, $C_{Ar\text{-}4}$), 132.1 (1C, $C_{Ar\text{-}1}$), 138.8 (1C, $C_{Ar\text{-}3}$), 156.1 (1C, NCOO), 163.2 (1C, $C_{Tet\text{-}6}$), 167.3 (1C, $C_{Tet\text{-}3}$), 171.1 (1C, COOH).
RP HPLC (C18, 250 x 21.2 mm, 10 mL/min, 1-80 % ACN in 16 min**)**: t_R = 13-16 min
MS (ESI), *m/z* calc for $C_{17}H_{22}N_6O_4$ 374.21, found 375.2 $[M+H]^+$, 373.1 $[M-H]^-$

1.5.2 Synthesis of K-mTetK

1.5.2.1 Synthesis route in solution

N^2-(L-lysyl)-N^6-(((3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine trifluoroacetate (K-mTetK)



N^2,N^6-bis(*tert*-butoxycarbonyl)-L-lysine (1.4 g, 3.25 mmol, 1 eq) was dissolved in 5 mL DMF, then HATU (1.1 g. 2.9 mmol, 0.9 eq) and HOAt (0.4 g. 2.9 mmol, 0.9 eq) were added and the reaction stirred at room temperature for 1 h, before addition of DIPEA (2.4 mL, 16 mmol, 5 eq) and mTetK (1.7 g, 3.6 mmol, 1.1 eq). After stirring overnight the reaction was concentrated under reduced pressure, dissolved in EtOAc and washed with 10 % citric acid. The crude residue was purified by flash chromatography (0-10 % MeOH/DCM + 0.1 % AcOH) to yield the product as a pink solid, which was directly dissolved in 10 mL TFA/DCM (1:1) with a few drops of water and stirred at room temperature for 5 h. The reaction was concentrated, the product dissolved in MeOH, then precipitated from cold Et_2O and dried to obtain the product (0.27 g, 11 % over two steps) as pink solid.

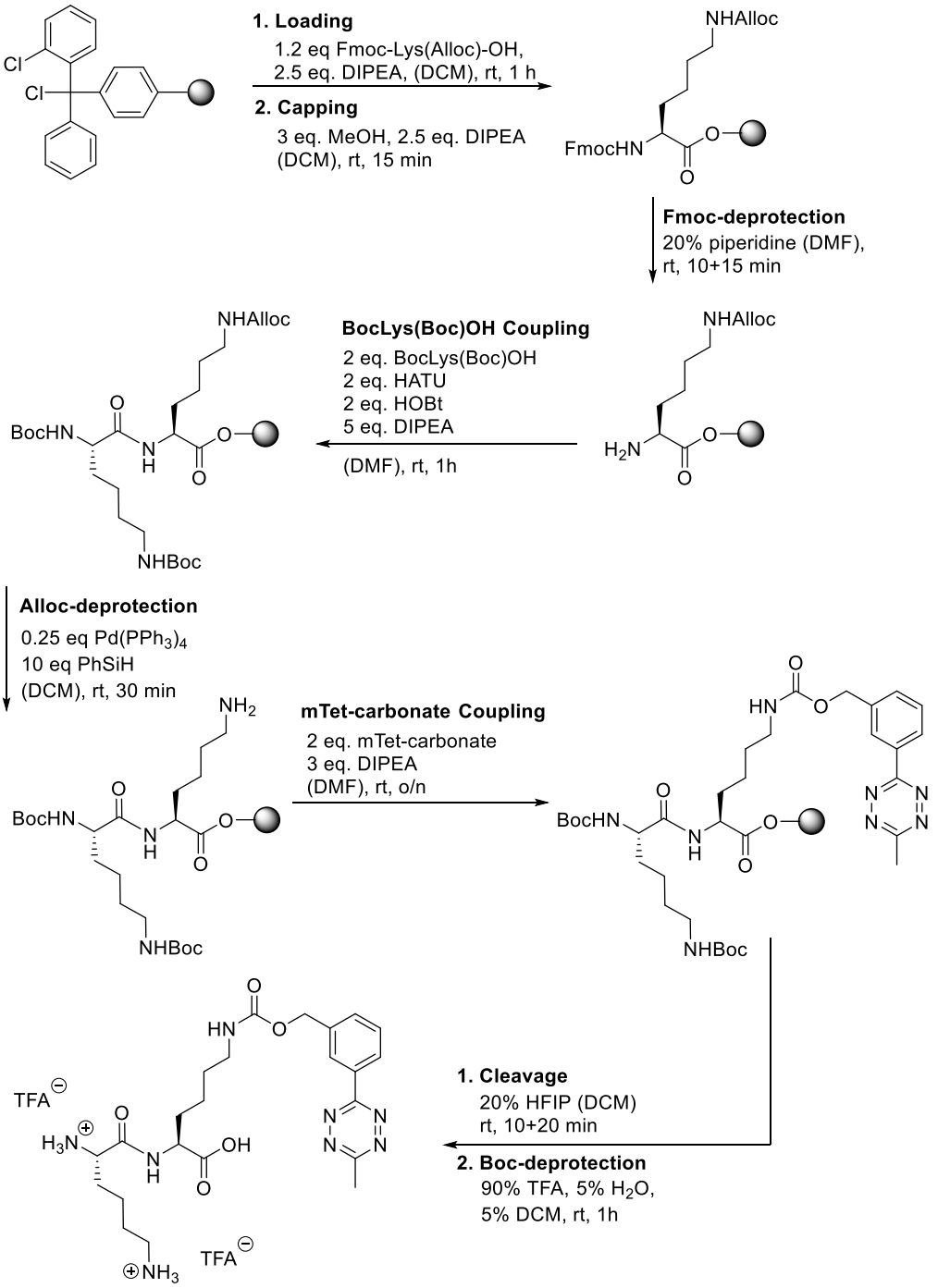
<u>Protected product $(Boc)_2$K-mTetK:</u>
MS (ESI), *m/z* calc for $C_{33}H_{50}N_8O_9$ 702.37, found 603.1 $[M\text{-Boc}+H]^+$, 701.3 $[M-H]^-$

<u>Deprotected product K-mTetK double TFA salt:</u>
^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.22-1.64 (m, 8H, $CH_{2\text{-}\gamma,\delta}$-K, mTetK), 1.61-1.79 (m, 4H, $CH_{2\text{-}\beta}$-K, mTetK), 2.68-2.80 (m, 2H, $CH_{2\text{-}\varepsilon}$-K), 2.92-3.05 (m, 5H, CH_3, $CH_{2\text{-}\varepsilon}$-mTetK),

3.77-3.86 (m, 1H, CH_{α}-K), 4.15-4.26 (m, 1H, CH_{α}-mTetK), 5.17 (s, 2H, CH_2), 7.30-7.41 (m, NH_{ε} -mTetK), 7.62-7.68 (m, 2H, $H_{Ar\text{-}4,6}$), 8.44-8.47 (m, 1H, $H_{Ar\text{-}5}$), 8.62-8.66 (m, 1H, $H_{Ar\text{-}2}$), 8.76 (s, 1H, NH_{α}-Peptide).
MS (ESI), *m/z* calc for $C_{23}H_{34}N_8O_5$ 502.27, found 503.2 $[M\text{-}Boc+H]^+$, 501.3 $[M\text{-}H]^-$

1.5.2.2 Solid phase synthesis

N^2-(L-lysyl)-N^6-(((3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine trifluoroacetate (K-mTetK)



2-Chlorotrityl chloride resin (1 g, 1.6 mmol, 1 eq) was weighed in a 25 mL SPPS cartridge and resuspended in DCM for 10 minutes, then the excess DCM was drained. N^2-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-N^6-((allyloxy)carbonyl)-L-lysine (0.87 g, 1.9 mmol, 1.2 eq) and DIPEA (0.8 mL, 4.6 mmol, 2.8 eq) dissolved in 5 mL DCM were added and the reaction shaken for 45 minutes at room temperature. Then MeOH (0.65 mL, 16 mmol, 10 eq) and DIPEA (0.35 mL, 1.3 eq) were added and shaken for 10 minutes for endcapping. The resin was then washed with DCM (4x) and DMF (4x) before Fmoc deprotection with 20 % piperidine in DMF for 15 minutes. Afterwards the resin was washed with DMF (4x). N^2,N^6-bis(*tert*-butoxycarbonyl)-*L*-lysine (1.7 g, 3.2 mmol, 2 eq), HATU (1.2 g, 3.2 mmol, 2 eq), HOBt (0.43 g, 3.2 mmol, 2 eq) and DIPEA (1.4 mL, 8 mmol, 5 eq) were resuspended in 10 mL DMF and then added to the resin and shaken for 1 h at room temperature. Then the resin was drained and washed with DMF (4x) and DCM (4x), followed by Alloc deprotection with $Pd(PPh_3)_4$ (0.46 g, 0.4 mmol, 0.25 eq) and phenylsilane (2.2 mL, 16 mmol, 10 eq) in DCM for 20 to 30 minutes until the formation of CO_2 gas was complete. The resin was then washed each with 0.5 % sodium *N*,*N*-diethyldithiocarbamate in DMF (3x), 0.5 % DIPEA in DMF (3x) and DMF (3x). 3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate (1.2 g, 3.2 mmol, 2 eq) and DIPEA (0.5 mL, 4.8 mmol, 3 eq) were dissolved in 10 mL DMF, added to the resin and shaken overnight at room temperature. Then, the resin was drained and washed with DMF (4x) and DCM (4x) before splitting the dipeptide from the resin with 2x 20 % HFIP in DCM (10 mL) for 20 minutes. The solvent was then removed under reduced pressure to obtain the product (0.66 g, 59 %) as a pink solid.

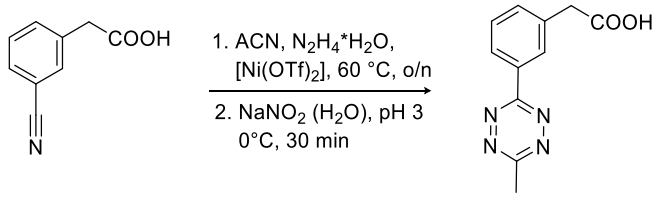
MS (ESI), *m/z* calc for $C_{33}H_{50}N_8O_9$ 702.37, found 603.1 $[M\text{-}Boc+H]^+$, 701.3 $[M\text{-}H]^-$

N^2-(N^2,N^6-bis(tert-butoxycarbonyl)-L-lysyl)-N^6-(((3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine (0.66 g, mmol, 1 eq) was dissolved in a solution of 6 mL TFA with DCM (5 %) and water (5 %) and stirred for 30 minutes. The reaction was concentrated under reduced pressure, the product dissolved in a minimal amount of MeOH, then precipitated from cold Et_2O and subsequently dried to obtain the product as a pink solid, which was further purified by HPLC (0.25 g, 37 %).

RP HPLC (C18, 250 x 21.2 mm, 8 mL/min, 20-90 % ACN in 17 min**)**: t_R = 12.0-14.2 min
MS (ESI), *m/z* calc for $C_{23}H_{34}N_8O_5$ 502.27, found 503.2 $[M\text{-}Boc+H]^+$, 501.3 $[M\text{-}H]^-$

1.5.3 Synthesis of mTet2K

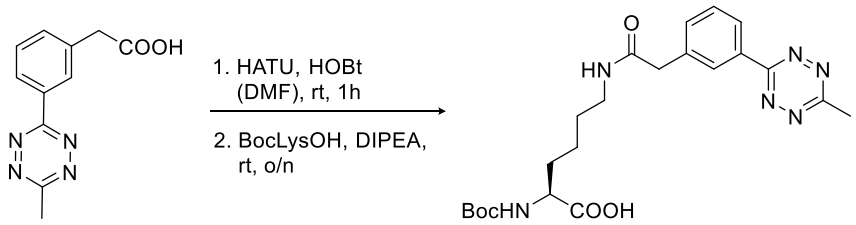
3-(3-(Carboxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine



3-(Carboxymethyl)benzonitrile (0.5 g, 3.1 mmol, 1 eq) and $Ni(OTf)_2$ (0.5 g, 1.6 mmol, 0.5 eq) were weighed in a flask, then ACN (1.6 mL, 31 mmol, 10 eq) and hydrazine hydrate (3.8 mL, 78 mmol, 25 eq) were added and the reaction heated to 60 °C for 24-36 h. $NaNO_2$ (1.1 g, 16 mmol, 5 eq) dissolved in H_2O (5 mL) was added and the recation mixture cooled in an ice bath before slowly acidifying to pH 3 using a 6 M HCl solution. The reaction was stirred for another 30 minutes on ice until the formation of nitrous fumes was complete, before it was extracted with EtOAc (5x 50 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude residue was purified by flash chromatography (0-50 % EtOAc/pentane + 0.1 % AcOH), which resulted the product (0.22 g, 31 %) as a pink solid.

^{1}H-NMR (500 MHz, DMSO-d_6): δ = 3.00 (s, 3H, CH_3), 3.77 (s, 2H, CH_2), 7.56-7.64 (m, 2H, $H_{Ar\text{-}4,6}$), 8.34-8.37 (m, 1H, $H_{Ar\text{-}5}$), 8.48-8.40 (m, 1H, $H_{Ar\text{-}2}$).

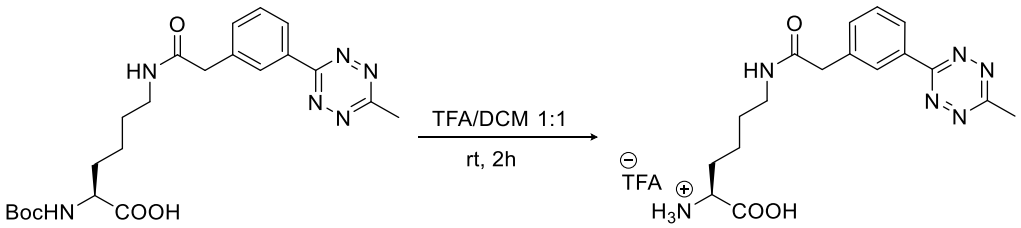
***N*^{2}-(*tert*-butoxycarbonyl)-*N*^{6}-(2-(3-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)acetyl)-L-lysine (BocmTet2K)**



3-(3-(Carboxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine (0.22 g, 0.96 mmol, 1 eq) was dissolved in 5 mL DMF, then HATU (0.33 g, 0.86 mmol, 0.9 eq) and HOBt (0.13 g, 0.86 mmol, 0.9 eq) were added and stirred at room temperature for 1 h, before addition of Boc-L-lysine (0.28 g, 1.2 mmol, 1.2 eq) and DIPEA (0.84 mL, 4.8 mmol, 5 eq). The resulting reaction mixture was stirred at room temperature overnight. The solvent was evaporated under reduced pressure, the residue dissolved in EtOAc (50 mL) and washed with a 20 % citric acid solution (3x 20 mL). The crude product was purified by flash chromatography (0-4 % MeOH/DCM + 0.1 % AcOH) to obtain the product (0.25 g, 57 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.22-1.44 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.37 (s, 9H, Boc), 1.47-1.67 (m, 2H, $CH_{2\text{-}\beta}$), 3.00 (s, 3H, CH_3), 2.98-3.08 (m, 2H, $CH_{2\text{-}\varepsilon}$), 3.55 (s, 2H, CH_2), 3.75-3.86 (m, 1H, $CH_{2\text{-}\alpha}$), 6.99 (d, $^3J_{HH}$ = 8.0 Hz, 1H, NH_α), 7.52-7.63 (m, 2H, $H_{Ar\text{-}4,6}$), 8.15 (t, $^3J_{HH}$ = 5.6 Hz, 1H, NH_ε), 8.28-8.36 (m, 1H, $H_{Ar\text{-}5}$), 8.38-8.42 (m, 1H, $H_{Ar\text{-}2}$).

N^6-(2-(3-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)acetyl)-L-lysine trifluoroacetate (mTet2K)

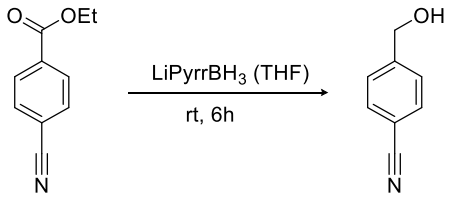


BocmTet2K (0.25 g, 3 mmol, 1 eq) was dissolved in 20 mL TFA/DCM (1:1) with a drop of H_2O and stirred at room temperature for two hours. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (0.17 g, 68 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.27-1.49 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.67-1.84 (m, 2H, $CH_{2\text{-}\beta}$), 2.98 (s, 3H, CH_3), 3.00-3.11 (m, 2H, $CH_{2\text{-}\varepsilon}$), 3.56 (s, 2H, CH_2), 3.82-3.93 (m, 1H, $CH_{2\text{-}\alpha}$), 7.53-7.58 (m, 2H, $H_{Ar\text{-}4,6}$), 8.18 (s, 3H, NH_α), 8.30-8.35 (m, 1H, $H_{Ar\text{-}5}$), 8.38-8.41 (m, 1H, $H_{Ar\text{-}2}$).
^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 20.9 (1C, CH_3), 22.1 (1C, C_γ), 28.9 (1C, C_δ), 30.0 (1C, C_β), 38.5, (1C, C_ε), 42.5 (1C, CH_2), 52.2 (1C, C_α), 125.8 (1C, $C_{Ar\text{-}5}$), 128.2 (1C, $C_{Ar\text{-}2}$), 129.5 (1C, $C_{Ar\text{-}6}$), 132.1 (1C, $C_{Ar\text{-}1}$), 133.4 (1C, $C_{Ar\text{-}4}$), 138.1 (1C, $C_{Ar\text{-}3}$), 163.6 (1C, $C_{Tet\text{-}6}$), 167.4 (1C, $C_{Tet\text{-}3}$), 170.0 (1C, HNCO), 171.4 (1C, COOH).

1.5.4 Synthesis of pTetK

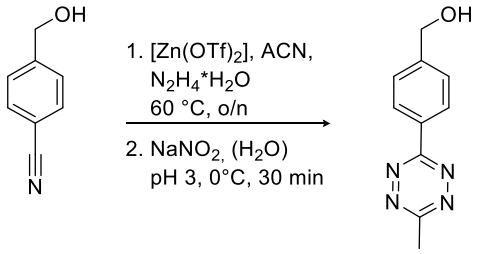
4-(Hydroxymethyl)benzonitrile[307]



A 1 M solution of lithium pyrrolidinoborohydride in THF (34.2 mL, 34.2 mmol, 1.2 eq) was added to ethyl 4-cyanobenzoate (5 g, 28.5 mmol, 1 eq) dissolved in THF (25 mL) and the reaction mixture stirred at room temperature for 6 h. The reaction was quenched with 3 M HCl (48 mL) solution and then extracted with Et_2O (3x 30 mL). The solvent was removed under reduced pressure to yield the product (3.4 g, 75 %) as a white solid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.21-2.29 (m, 1H, OH), 4.77 (d, $^3J_{HH}$ = 5.5 Hz, 2H, CH_2), 7.47 (d, $^3J_{HH}$ = 8.3 Hz, 2H, $H_{Ar\text{-}2,2'}$), 7.64 (d, $^3J_{HH}$ = 8.3 Hz, 2H, $H_{Ar\text{-}3,3'}$).
MS (ESI), *m/z* calc for C_8H_7NO 133.05, found 134.1 $[M+H]^+$

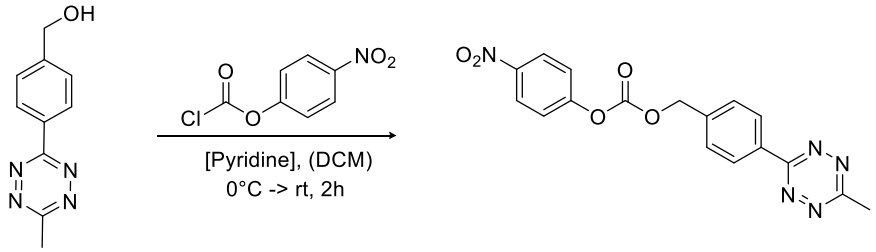
3-(4-(Hydroxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine[187]



4-(Hydroxymethyl)benzonitrile (2.95 g, 22.1 mmol, 1 eq) and $Zn(OTf)_2$ (4 g, 11 mmol, 0.5 eq) were weighed in a flask, then ACN (11.5 mL, 0.22 mol, 10 eq) and hydrazine hydrate (26.9 mL, 0.55 mol, 25 eq) were added and the reaction heated to 60 °C for 24-36 h. 6 M HCl solution was used to slowly acidify the reaction mixture to pH 3 on ice after addition of a solution of $NaNO_2$ (7.6 g, 0.11 mol, 5 eq) dissolved in H_2O (20 mL). After stirring for 30 minutes, the reaction was extracted with EtOAc (5x 150 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude residue was purified by flash chromatography (0-50 % EtOAc/pentane), which yielded the product (1.73 g, 38 %) as a pink solid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.10 (s, 3H, CH_3), 4.84 (, 2H, CH_2), 7.59 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $H_{Ar\text{-}3,3'}$), 8.59 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $H_{Ar\text{-}2,2'}$).
MS (ESI), *m/z* calc for $C_{10}H_{10}N_4O$ 202.09, found 202.7 $[M+H]^+$

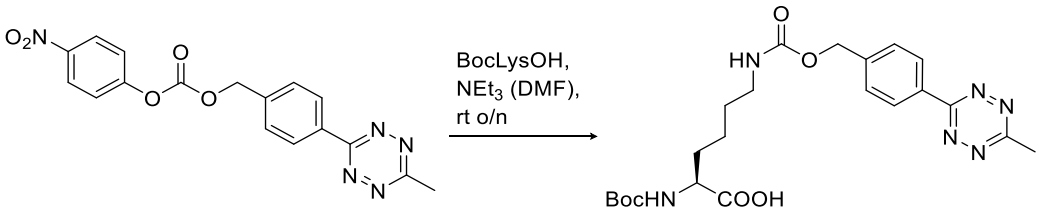
4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate



4-Nitrophenyl chloroformate (2.1 g, 10.3 mmol, 1.2 eq) was added to a solution of 3-(4-(Hydroxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine (1.73 g, 8.6 mmol, 1 eq) in DCM (40 mL), followed by the addition of pyridine (0.69 mL, 8.6 mmol 1 eq) and then stirred at room temperature for 2 hours. After removal of the solvent under reduced pressure, flash chromatography was used to purify the crude mixture (0-30 % EtOAc/pentane), which resulted the product (3 g, 96 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 3.12 (s, 3H, CH_3), 5.41 (s, 2H, CH_2), 7.41 (d, $^3J_{HH}$ = 9.2 Hz, 2H, $H_{Ar\text{-}metaNO2}$), 7.67 (d, $^3J_{HH}$ = 8.3 Hz, 1H, $H_{Ar\text{-}2,2'}$), 8.29 (d, $^3J_{HH}$ = 9.2 Hz, 2H, $H_{Ar\text{-}orthoNO2}$), 8.65 (d, $^3J_{HH}$ = 8.3 Hz, 1H, $H_{Ar\text{-}3,3'}$).

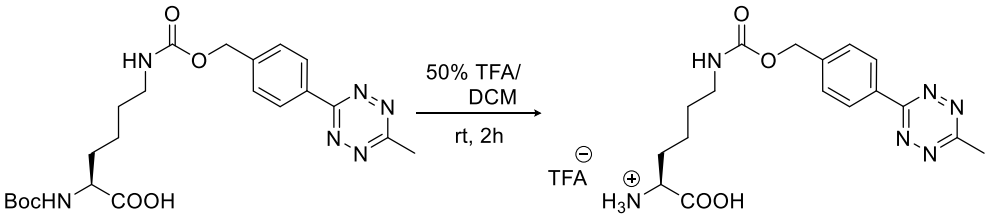
N^2-(*tert*-butoxycarbonyl)-N^6-(((4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine



4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate (1.3 g, 3.4 mmol, 1 eq) was added to a solution of Boc-L-lysine (1 g, 4 mmol, 1.2 eq) and NEt_3 (0.56 mL, 4 mmol, 1.2 eq) in DMF (11 mL), which was then stirred at room temperature overnight. Afterwards the solvent was evaporated under reduced pressure and the crude residue purified by flash chromatography (0-2 % MeOH/DCM + 0.1 % AcOH) to result the product as a pink solid (1.44 g, 89 %).

^{1}H-NMR (500 MHz, DMSO-d_6): δ = 1.22-1.44 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.37 (s, 9H, Boc), 1.47-1.69 (m, 2H, $CH_{2\text{-}\beta}$), 2.96-3.03 (m, 2H, $CH_{2\text{-}\varepsilon}$), 3.00 (s, 3H, CH_3), 3.79-3.86 (m, 1H, $CH_{2\text{-}\alpha}$), 5.15 (s, 2H, CH_2), 7.03 (d, $^3J_{HH}$ = 8. Hz, 1H, NH_α), 7.35 (t, $^3J_{HH}$ = 5.6 Hz, 1H, NH_ε), 7.61 (d, $^3J_{HH}$ = 8.3 Hz, 1H, $H_{Ar\text{-}2,2'}$), 8.47 (d, $^3J_{HH}$ = 8.3 Hz, 1H, $H_{Ar\text{-}3,3'}$).
MS (ESI), *m/z* calc for $C_{22}H_{30}N_6O_6$ 474.22, found 375.1 $[M\text{-Boc+H}]^+$, 497.2 $[M+Na]^+$, 473.1 $[M-H]^-$,

N^6-(((4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-*L*-lysine trifluoroacetate (pTetK)



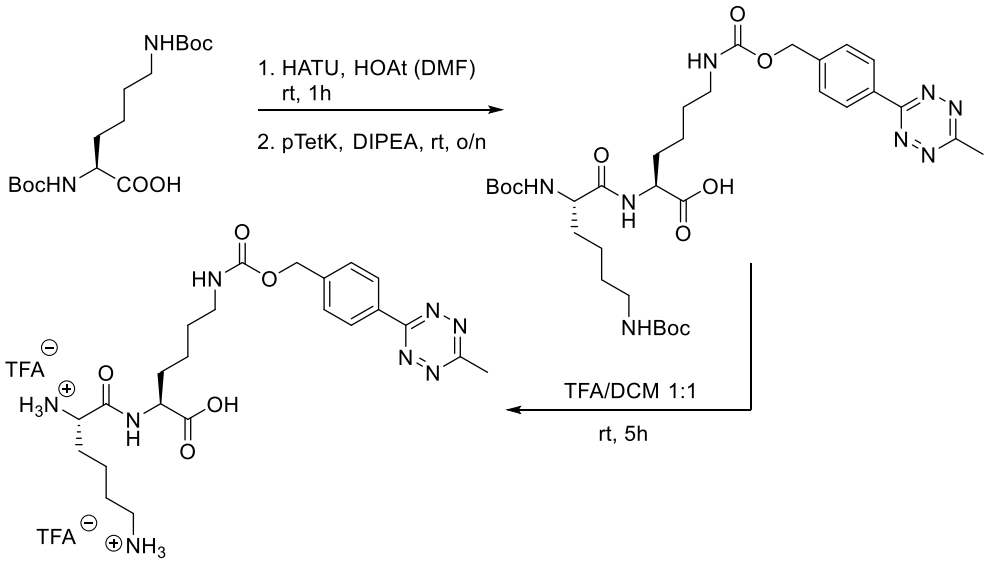
BocpTetK (1.44 g, 3 mmol, 1 eq) was dissolved in 20 mL TFA/DCM (1:1) with a drop of H_2O and stirred at room temperature for two hours. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (1 g, 73 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.27-1.53 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.67-1.84 (m, 2H, $CH_{2\text{-}\beta}$), 2.98 (s, 3H, CH_3), 2.93-3.09 (m, 2H, $CH_{2\text{-}\varepsilon}$), 3.82-3.93 (m, 1H, $CH_{2\text{-}\alpha}$), 5.14 (s, 2H, CH_2), 7.31 (s, 1H, NH_ε), 7.60 (d, $^3J_{HH}$ = 8.3 Hz, 1H, $H_{Ar\text{-}2,2'}$), 8.19 (s, 3H, NH_α), 8.46 (d, $^3J_{HH}$ = 8.3 Hz, 1H, $H_{Ar\text{-}3,3'}$).
^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 20.9 (1C, CH_3), 21.9 (1C, C_γ), 29.1 (1C, C_δ), 29.9 (1C, C_β), 30.2, (1C, C_ε), 52.2 (1C, C_α), 64.8 (1C, CH_2), 127.7 (2C, $C_{Ar\text{-}2,2'}$), 128.4 (2C, $C_{Ar\text{-}3,3'}$), 131.5 (1C, $C_{Ar\text{-}6}$), 142.2 (1C, $C_{Ar\text{-}1}$), 156.3 (1C, NCOO), 163.4 (1C, $C_{Tet\text{-}6}$), 167.3 (1C, $C_{Tet\text{-}3}$), 171.3 (1C, COOH).
MS (ESI), *m/z* calc for $C_{17}H_{22}N_6O_4$ 374.21, found 375.2 $[M+H]^+$, 373.1 $[M-H]^-$

1.5.5 Synthesis of K-pTetK

1.5.5.1 Synthesis in solution

N^2-(L-lysyl)-N^6-(((4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine



N^2,N^6-bis(*tert*-butoxycarbonyl)-L-lysine (0.42 g, 1.2 mmol, 1 eq) was dissolved in 5 mL DMF, then HATU (0.41 g. 1.1 mmol, 0.9 eq) and HOAt (0.15 g. 1.1 mmol, 0.9 eq) added and the reaction stirred at room temperature for 1 h before addition of DIPEA (0.6 mL, 6 mmol, 5 eq) and pTetK (0.65 g, 1.3 mmol, 1.1 eq). After further stirring overnight the reaction was concentrated under reduced pressure, dissolved in EtOAc and washed with 10 % citric acid. The crude residue was purified by flash chromatography (0-10 % MeOH/DCM + 0.1 % AcOH) to yield the product as a pink solid. This was dissolved in 10 mL TFA/DCM (1:1) with a few drops of water and stirred at room temperature for 2 h. The reaction was concentrated in vacuo, the product dissolved in a minimal amount of MeOH, precipitated from cold Et_2O and dried to obtain the product (0.16 g, 18 % over two steps) as a pink solid.

MS (ESI), *m/z* calc for $C_{33}H_{50}N_8O_9$ 702.37, found 603.1 $[M\text{-Boc+H}]^+$, 701.3 $[M\text{-H}]^-$

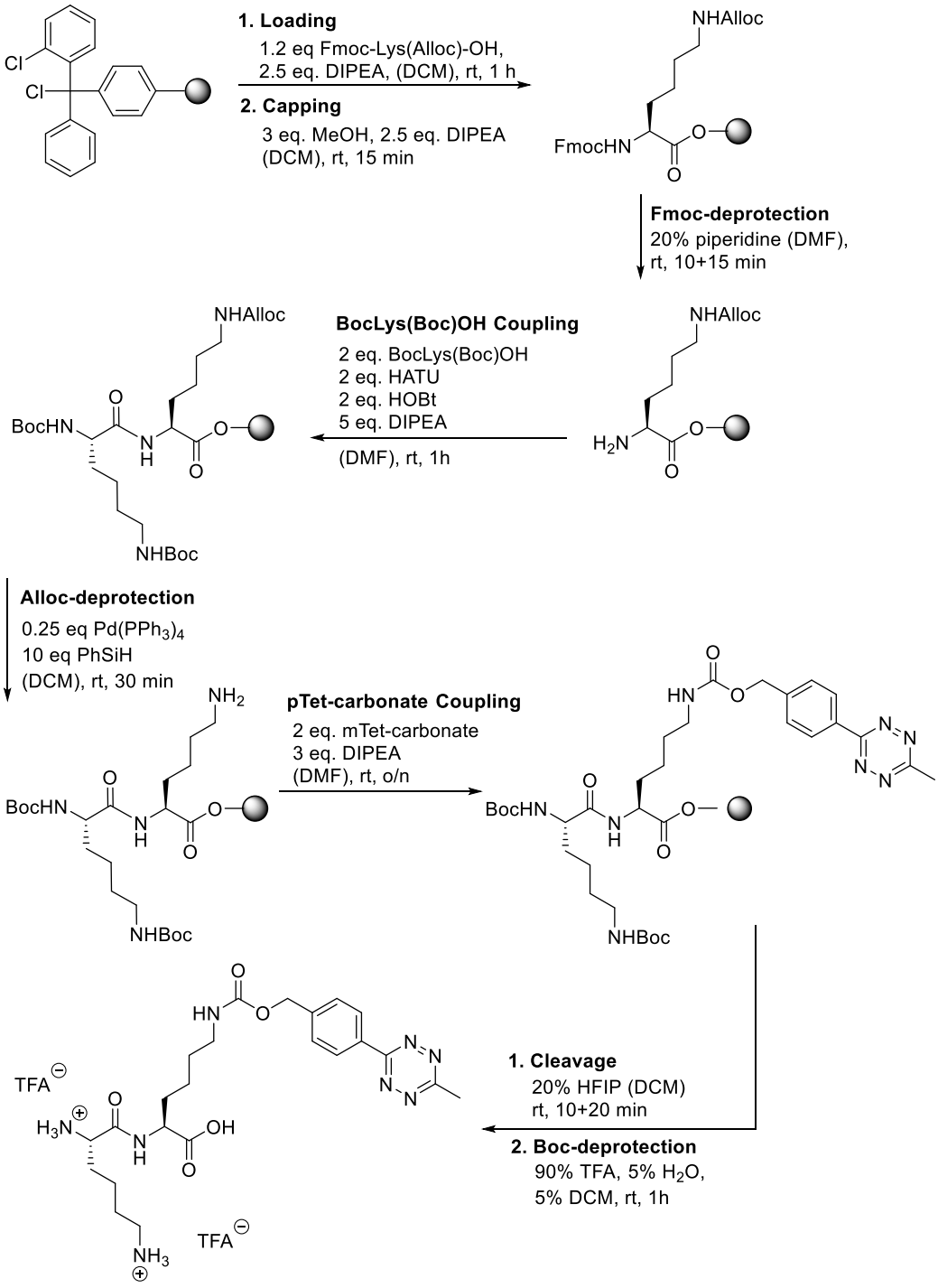
^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.29-1.64 (m, 8H, CH_{2}-$_{\gamma,\delta}$-K, pTetK), 1.68-1.89 (m, 4H, CH_{2}-$_{\beta}$-K, pTetK), 2.70-2.79 (m, 2H, CH_{2}-$_{\varepsilon}$-K), 2.96-3.05 (m, 5H, CH_3, CH_{2}-$_{\varepsilon}$-mTetK), 3.77-3.86 (m, 1H, CH_{α}-K), 4.15-4.26 (m, 1H, CH_{α}-pTetK), 5.15 (s, 2H, CH_2), 7.34-7.42 (m, NH_{ε}-pTetK), 7.61 (d, $^3J_{HH}$ = 8.5 Hz, 2H, $H_{Ar\text{-}3,3'}$), 8.46 (d, $^3J_{HH}$ = 8.5 Hz, 2H, $H_{Ar\text{-}2,2'}$), 8.67 (s, 1H, NH_{α}-Peptide).

^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 20.6 (1C, CH_3), 20.9 (1C, C_{γ}-K), 22.7 (1C, C_{γ}-pTetK) 26.5 (1C, C_{δ}-K), 29.0 (1C, C_{δ}-pTetK), 30.4 (1C, C_{β}-K), 30.6 (1C, C_{β}-pTetK), 38.3 (1C, C_{ε}-K) 39.9, (1C, C_{ε}-pTetK), 51.6 (1C, C_{α}-K), 52.5 (1C, C_{α}-pTetK), 64.5 (1C, CH_2), 127.1 (2C, $C_{Ar\text{-}2,2'}$), 128.3 (2C, $C_{Ar\text{-}3,3'}$), 129.6 (1C, $C_{Ar\text{-}6}$), 131.2 (1C, $C_{Ar\text{-}4}$), 141.9 (1C, $C_{Ar\text{-}1}$), 156.1 (1C, NCOO), 162.8 (1C, $C_{Tet\text{-}6}$), 167.1 (1C, $C_{Tet\text{-}3}$), 168.9 (1C, HNCO), 173.0 (1C, COOH).

MS (ESI), *m/z* calc for $C_{23}H_{34}N_8O_5$ 502.27, found 503.2 $[M\text{-Boc+H}]^+$, 501.3 $[M\text{-H}]^-$

1.5.5.2 Solid phase synthesis

N^2-(L-lysyl)-N^6-(((4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine



2-Chlorotrityl chloride resin (1 g, 1.6 mmol, 1 eq) was weighed in a 25 mL SPPS cartridge and resuspended in DCM for 10 minutes, then the excess DCM was drained. N^2-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-N^6-((allyloxy)carbonyl)-L-lysine (0.87 g, 1.9 mmol, 1.2 eq) and DIPEA (0.8 mL, 4.6 mmol, 2.8 eq) dissolved in 5 mL DCM were added and the reaction shaken for 45 minutes at room temperature. Then MeOH (0.65 mL, 16 mmol, 10 eq) and DIPEA (0.35 mL, 1.3 eq) were added and shaken for 10 minutes for endcapping. The resin was then washed with DCM (4x) and DMF (4x) before Fmoc deprotection with 20 %

piperidine in DMF for 15 minutes. Afterwards the resin was washed with DMF (4x). *N*2,*N*6-bis(*tert*-butoxycarbonyl)-L-lysine (1.7 g, 3.2 mmol, 2 eq), HATU (1.2 g, 3.2 mmol, 2 eq), HOBt (0.43 g, 3.2 mmol, 2 eq) and DIPEA (1.4 mL, 8 mmol, 5 eq) were resuspended in 10 mL DMF and then added to the resin and shaken for 1 h at room temperature. Then the resin was drained and washed with DMF (4x) and DCM (4x), followed by Alloc deprotection with $Pd(PPh_3)_4$ (0.46 g, 0.4 mmol, 0.25 eq) and phenylsilane (2.2 mL, 16 mmol, 10 eq) in DCM for 20 to 30 minutes until the formation of CO_2 gas was complete. The resin was then washed each with 0.5 % sodium *N*,*N*-diethyldithiocarbamate in DMF (3x), 0.5 % DIPEA in DMF (3x) and DMF (3x). 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate (1.2 g, 3.2 mmol, 2 eq) and DIPEA (0.5 mL, 4.8 mmol, 3 eq) were dissolved in 10 mL DMF, added to the resin and shaken overnight at room temperature. Then, the resin was drained and washed with DMF (4x) and DCM (4x) before splitting the dipeptide from the resin with 2x 20 % HFIP in DCM (10 mL) for 20 minutes. The solvent was then removed under reduced pressure to obtain the product (0.9 g, 81 %) as a pink solid.

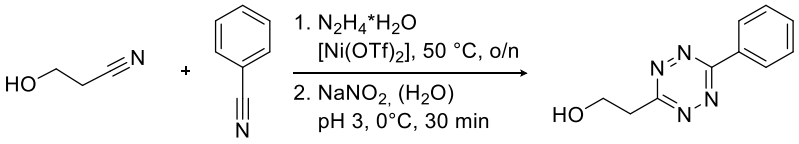
MS (ESI), *m/z* calc for $C_{33}H_{50}N_8O_9$ 702.37, found 603.1 $[M\text{-}Boc+H]^+$, 701.3 $[M\text{-}H]^-$

N^2-(N^2,N^6-bis(tert-butoxycarbonyl)-L-lysyl)-N^6-(((3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)-oxy)carbonyl)-L-lysine (0.9 g, mmol, 1 eq) was dissolved in a solution of 6 mL TFA with DCM (5 %) and water (5 %) and stirred for 30 minutes. The reaction was concentrated under reduced pressure, the product dissolved in a minimal amount of MeOH, then precipitated from cold Et_2O and subsequently dried to obtain the product as a pink solid, which was further purified by HPLC (0.61 g, 65 %).

RP HPLC (C18, 250 x 21.2 mm, 8 mL/min, 20-90 % ACN in 17 min**)**: t_R = 11.8-13.6 min
MS (ESI), *m/z* calc for $C_{23}H_{34}N_8O_5$ 502.27, found 503.2 $[M\text{-}Boc+H]^+$, 501.3 $[M\text{-}H]^-$

1.5.6 Synthesis of PhTetK

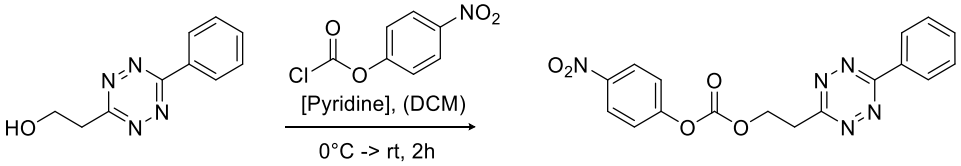
3-(2-Hydroxyethyl)-6-phenyl-1,2,4,5-tetrazine[308]



Hydroxypropionitrile (0.4 g, 5.7 mmol, 1 eq) was added to a flask, containing $Ni(OTf)_2$ (1 g, 2.8 mmol, 0.5 eq), followed by benzonitrile (1.1 mL, 11.3 mol, 2 eq) and hydrazine hydrate (6.8 mL, 0.14 mol, 25 eq). The resulting reaction mixture was stirred at 60 °C for 36 h. Then $NaNO_2$ (24.3 g, 0.35 mol, 5 eq) dissolved in a minimal amount of water was added and the mixture cooled on ice. 6 M HCl solution was slowly added to acidify the reaction mixture to pH 3. It was stirred for 30 minutes on ice until the formation of nitrous fumes was completed, then the water phase was extracted with EtOAc (5x 50 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 before removing the solvent under reduced pressure. The crude residue was purified by flash chromatography (0-60 % EtOAc/pentane), which resulted the product (0.49 g, 43 %) as a pink solid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.64 (t, $^3J_{HH}$ = 5.8 Hz, 2H, CH_2), 4.32 (t, $^3J_{HH}$ = 5.8 Hz, 2H, CH_2OH), 7.56-7-68 (m, 3H, $H_{Ar\text{-}3,3',4}$), 8.59-8.62 (m, 2H, $H_{Ar\text{-}2,2'}$).

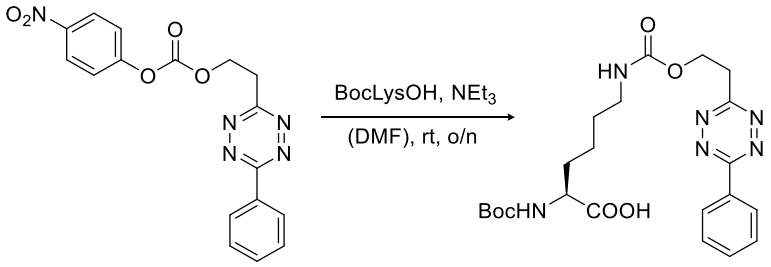
2-(6-phenyl-1,2,4,5-tetrazin-3-yl)ethyl-(4-nitrophenyl)-carbonate



3-(2-Hydroxyethyl)-6-phenyl-1,2,4,5-tetrazine (0.49 g, 2.4 mmol, 1 eq) was dissolved in DCM (10 mL), then 4-nitrophenyl chloroformate (0.58 g, 2.9 mmol, 2.2 eq) was added to the solution followed by the addition of pyridine (0.19 mL, 2.4 mmol 1 eq). The resulting solution was stirred at room temperature for two hours. After removal of the solvent under reduced pressure flash chromatography was used to purify the crude mixture (0-50 % EtOAc/pentane), which resulted the product (0.85 g, 96 %) as a red liquid.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 3.85 (t, $^3J_{HH}$ = 6.2 Hz, 2H, CH_2), 4.97 (t, $^3J_{HH}$ = 6.2 Hz, 2H, CH_2O), 7.36 (d, $^3J_{HH}$ = 9.2 Hz, 2H, $H_{Ar\text{-}metaNO2}$), 7.56-7-68 (m, 3H, $CH_{Ar\text{-}3,3',4}$), 8.26 (d, $^3J_{HH}$ = 9.2 Hz, 2H, $H_{Ar\text{-}orthoNO2}$), 8.59-8.64 (m, 2H, $CH_{Ar\text{-}2,2'}$).

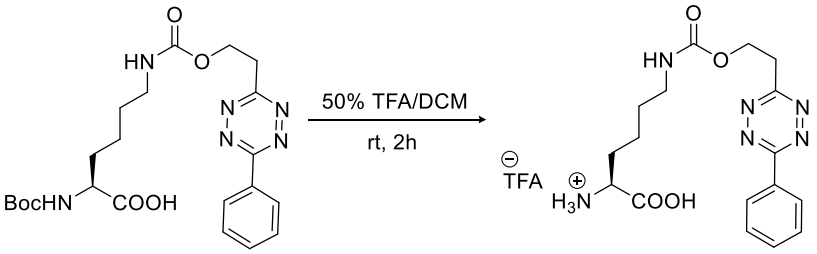
N^2-(*tert*-butoxycarbonyl)-N^6-((2-(6-phenyl-1,2,4,5-tetrazin-3-yl)ethoxy)-carbonyl)-L-lysine (BocPhTetK)



Boc-L-lysine (0.68 g, 2.8 mmol, 1.2 eq) was dissolved in DMF (10 mL), then NEt_3 (0.38 mL, 2.8 mmol, 1.2 eq) and 2-(6-phenyl-1,2,4,5-tetrazin-3-yl)ethyl-(4-nitrophenyl)-carbonate (0.85 g, 2.3 mmol, 1 eq) were added to the solution, which was stirred at room temperature overnight. Afterwards the solvent was evaporated under reduced pressure and the crude residue purified by flash chromatography (0-10 % MeOH/DCM + 0.1 % AcOH) to result the product (0.15 g, 14 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.19-1.36 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.37 (s, 8H, Boc), 1.46-1.64 (m, 2H, $CH_{2\text{-}\beta}$), 2.84-2.90 (m, 2H, $CH_{2\text{-}\varepsilon}$), 3.63 (t, $^3J_{HH}$ = 6.3 Hz, 2H, CH_2), 3.76-3.8480 (m, 1H, $CH_{\text{-}\alpha}$), 4.54 (t, $^3J_{HH}$ = 6.3 Hz, 2H, CH_2O), 6.98 (d, $^3J_{HH}$ = 8.2 Hz, 1H, $NH_{\text{-}\alpha}$), 7.10 (t, $^3J_{HH}$ = 6.0 Hz, 1H, $NH_{\text{-}\varepsilon}$), 7.56-7-68 (m, 3H, $CH_{Ar\text{-}3,3',4}$), 7.99-8.10 (m, 2H, $CH_{Ar\text{-}2,2'}$).
MS (ESI), *m/z* calc for $C_{22}H_{30}N_6O_6$ 474.21, found 375.2 $[M\text{-Boc+H}]^+$, 497.2 $[M+Na]^+$, 473.2 $[M-H]^-$.

N^6-((2-(6-phenyl-1,2,4,5-tetrazin-3-yl)ethoxy)-carbonyl)-*L*-lysine trifluoroacetate (PhTetK)

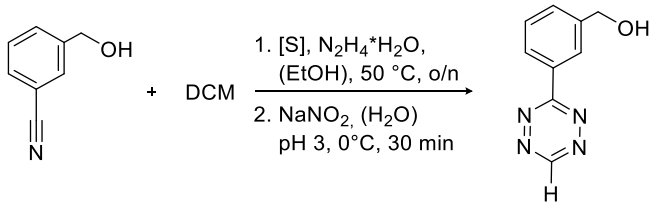


BocPhTetK (0.14 g, 3 mmol, 1 eq) was dissolved in 20 mL TFA/DCM (1:1) with a drop of H_2O and stirred at room temperature for two hours. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (0.86 g, 60 %) as pink solid, which was further purified by RP-HPLC (C18, 250 x 21.2 mm, 10 mL/min, 1-50 % ACN in 16 min).

RP HPLC (C18, 250 x 21.2 mm, 10 mL/min, 1-50 % ACN in 16 min**)**: t_R = 17-18 min
MS (ESI), *m/z* calc for $C_{17}H_{22}N_6O_4$ 374.17, found 375.2 $[M+H]^+$

1.5.7 Synthesis of HmTetK

3-(3-(Hydroxymethyl)phenyl)-1,2,4,5-tetrazine[309]

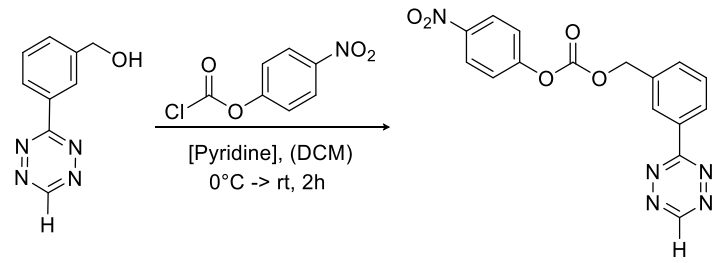


3-(Hydroxymethyl)benzonitrile (0.5 g, 3.8 mmol, 1 eq) and sulfur (0.24 g, 7.5 mmol, 2 eq) were suspended in EtOH, then DCM (0.25 mL, 3.8 mmol, 1 eq) and hydrazine hydrate (1.5 mL, 30 mmol, 8 eq) were added and the reaction heated to 50 °C for 16-20 h. $NaNO_2$ (1.3 g, 19 mmol, 5 eq) dissolved in H_2O (5 mL) was added and the reaction mixture cooled in an ice bath before slowly acidifying to pH 3 using a 6 M HCl solution. After stirring for 30 minutes until the formation of nitrous fumes was complete, the reaction was extracted with EtOAc (5x 30 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude residue was purified by flash chromatography (0-50 % EtOAc/pentane), which resulted the product (0.26 g, 36 %) as a pink solid, mixed 1:1 with 3-(Hydroxymethyl)benzonitrile starting material.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 4.75 (s, 2H, CH_2), 7.58-7.64 (m, 2H, $H_{Ar-4,6}$), 8.53-8.58 (m, 1H, H_{Ar-5}), 8.60-8.65 (m, 1H, H_{Ar-2}), 10.22 (s, 1H, H_{Tet3}).
MS (ESI), *m/z* calc for $C_9H_8N_4O$ 188.07, found 189.1 $[M+H]^+$

<u>Starting material:</u>
^{1}H-NMR (300 MHz, $CDCl_3$): δ = 4.65 (s, 2H, CH_2), 7.42-7.50 (m, 1H, H_{Ar-5}), 7.55-7.59 (m, 1H, H_{Ar-2}), 7.65-7.70 (m, 2H, $H_{Ar-4,6}$).
MS (ESI), *m/z* calc for C_7H_8NO 133.05, found 134.1 $[M+H]^+$

3-(1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate

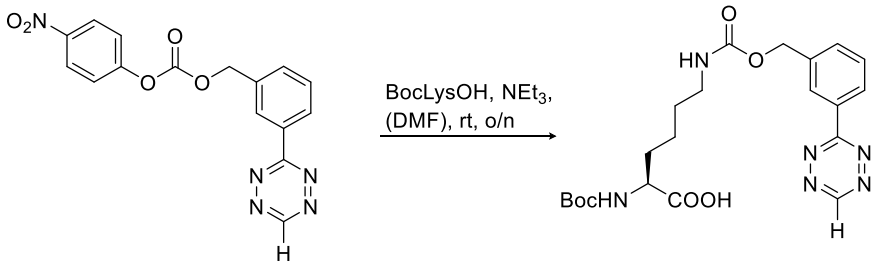


The mixture of starting material and 3-(3-(Hydroxymethyl)phenyl)-1,2,4,5-tetrazine from the previous step (2 g, 11 mmol, 1 eq) were dissolved in DCM (20 mL) and cooled to 0 °C. 4-Nitrophenyl chloroformate (4 g, 20 mmol, 1.8 eq) was added to the solution, followed by the addition of pyridine (1.3 mL, 16 mmol 1.5 eq). The resulting mixture was stirred at room temperature for 2 hours. After removal of the solvent under reduced pressure flash chromatography was used to purify the crude mixture (0-60 % EtOAc/pentane), which resulted the product (2.8 g, 71 %) as a pink solid mixed with activated starting material.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 5.44 (s, 2H, CH_2), 7.42 (d, $^3J_{HH}$ = 9.3 Hz, 2H, $H_{Ar\text{-}metaNO2}$), 7.64-7.71 (m, 2H, $H_{Ar\text{-}4,6}$), 8.29 (d, $^3J_{HH}$ = 9.3 Hz, 2H, $H_{Ar\text{-}orthoNO2}$), 8.64-8.70 (m, 1H, $H_{Ar\text{-}5}$), 8.72-8.75 (m, 1H, $H_{Ar\text{-}2}$), 10.26 (s, 1H, H_{Tet3}).
MS (ESI), *m/z* calc for $C_{16}H_{11}N_5O_5$ 353.08, found 376.1 $[M+Na]^+$

Activated starting material:
^{1}H-NMR (300 MHz, $CDCl_3$): δ = 5.32 (s, 2H, CH_2), 7.39-7.42 (m, 1H, $H_{Ar\text{-}5}$), 7.39 (d, $^3J_{HH}$ = 9.3 Hz, 2H, $H_{Ar\text{-}metaNO2}$), 7.51-7.59 (m, 1H, $H_{Ar\text{-}2}$), 7.72-7.77 (m, 2H, $H_{Ar\text{-}4,6}$), 8.29 (d, $^3J_{HH}$ = 9.3 Hz, 2H, $H_{Ar\text{-}orthoNO2}$).

***N*^{2}-(*tert*-butoxycarbonyl)-*N*^{6}-(((3-(1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine**

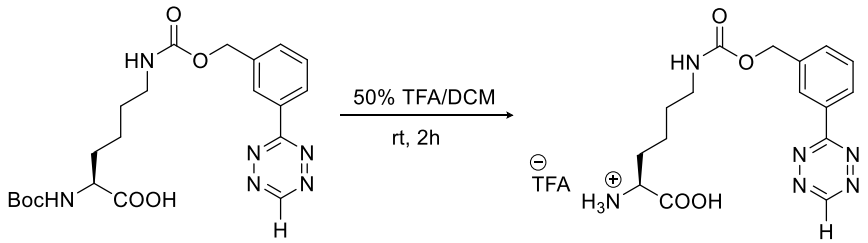


3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate (0.4 g, 1.1 mmol, 1 eq) was added to a solution of Boc-L-lysine (0.33 g, 1.34 mmol, 1.2 eq) and NEt_3 (0.23 mL, 1.34 mmol, 1.2 eq) in DMF (10 mL), which was stirred at room temperature overnight. Afterwards the solvent was evaporated under reduced pressure and the crude residue purified by flash chromatography (0-2 % MeOH/DCM + 0.1 % AcOH) to result the product as a pink solid (0.38 g, 74 %) mixed with side product.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 1.39-1.48 (m, 4H, $CH_{2\text{-}\gamma}$), 1.44 (s, 9H, Boc), 1.50-1.61 (m, 2H, $CH_{2\text{-}\delta}$), 1.67-1.93 (m, 2H, $CH_{2\text{-}\beta}$), 3.15-3.29 (m, 2H, $CH_{2\text{-}\varepsilon}$), 4.25-4.36 (m, 1H, $CH_{2\text{-}\alpha}$), 4.98 (s, 1H, NH_ε), 5.16 (s, 1H, NH_α), 5.24 (s, 2H, CH_2), 7.43-7.50 (m, 1H, $H_{Ar\text{-}5}$), 7.54-7.60 (m, 1H, $H_{Ar\text{-}4,6}$), 8.55-8.64 (m, 1H, $H_{Ar\text{-}2}$).
MS (ESI), *m/z* calc for $C_{21}H_{28}N_6O_6$ 460.21
found 361.2 $[M\text{-Boc}+H]^+$, 483.2 $[M+Na]^+$, 459.2 $[M-H]^-$

Side product: *N*^{2}-(*tert*-butoxycarbonyl)-*N*^{6}-(((3-cyanobenzyl)oxy)carbonyl)-L-lysine
MS (ESI), *m/z* calc for $C_{20}H_{27}N_3O_6$ 405.2, found 306.1 $[M\text{-Boc}+H]^+$, 404.1 $[M-H]^-$

N^6-(((3-(1,2,4,5-tetrazin-3-yl)benzyl)oxy)carbonyl)-L-lysine trifluoroacetate (HmTetK)



BocHmTetK (0.12 g, 0.24 mmol, 1 eq) was dissolved in 6 mL TFA/DCM (1:1) with a drop of H_2O and stirred at room temperature for two hours. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (89 mg, 75 %) as a pink solid mixed with side product. The amino acid was further purified by RP-HPLC (C18, 250 x 21.2 mm, 10 mL/min, 10-30 % ACN in 75 min), which decreased the amount of side product, but did not completely separate from it.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.24-1.53 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.67-1.86 (m, 2H, $CH_{2\text{-}\beta}$), 2.95-3.07 (m, 2H, $CH_{2\text{-}\varepsilon}$), 3.82-3.93 (m, 1H, $CH_{2\text{-}\alpha}$), 5.18 (s, 2H, CH_2), 7.35 (t, $^3J_{HH}$ = 5.7 Hz, 1H, $NH_{\text{-}\varepsilon}$), 7.65-7.61 (m, 2H, $H_{Ar\text{-}4,6}$), 8.22 (s, 3H, NH_α), 8.42-8.47 (m, 1H, $H_{Ar\text{-}5}$), 8.49-8.52 (m, 1H, $H_{Ar\text{-}2}$), 10.61 (s, 1H, $H_{Tet\text{-}3}$).
^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 21.7 (1C, C_γ), 28.9 (1C, C_δ), 29.7 (1C, C_β), 40.0, (1C, C_ε), 51.9 (1C, C_α), 64.7 (1C, CH_2), 126.7 (1C, $C_{Ar\text{-}2}$), 127.1 (1C, $C_{Ar\text{-}5}$), 129.6 (1C, $C_{Ar\text{-}6}$), 132.0 (1C, $C_{Ar\text{-}4}$), 132.1 (1C, $C_{Ar\text{-}1}$), 138.8 (1C, $C_{Ar\text{-}3}$), 156.1 (1C, NCOO), 158.1 (1C, CH_{Tet}), 165.4 (1C, $C_{Tet\text{-}6}$), 171.1 (1C, COOH).

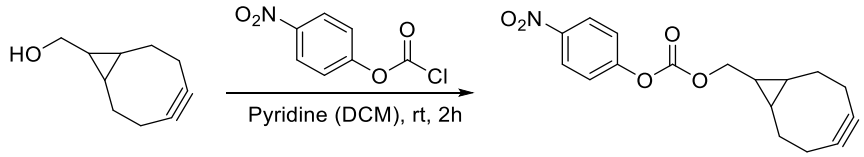
RP HPLC (C18, 250 x 21.2 mm, 10 mL/min, 10-30 % ACN in 75 min**)**: t_R = 42.8-43.8 min
MS (ESI), *m/z* calc for $C_{16}H_{20}N_6O_4$ 360.15, found 361.2 $[M+H]^+$, 359.2 $[M-H]^-$

<u>Side product: N^6-(((3-cyanobenzyl)oxy)carbonyl)-L-lysine trifluoroacetate (3-CNPhK)</u>
MS (ESI), *m/z* calc for $C_{15}H_{19}N_3O4$ 305.14, found 306.1 $[M+H]^+$, 304.1 $[M-H]^-$

1.5.8 Synthesis of Fluorophores

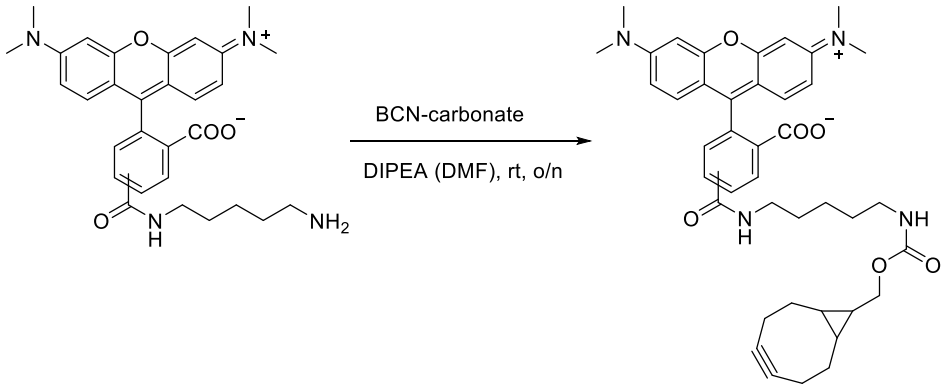
1.5.8.1 Synthesis of BCN-TAMRA

Bicyclo[6.1.0]non-4-yn-9-ylmethyl (4-nitrophenyl) carbonate (BCN-carbonate)



Bicyclo[6.1.0]non-4-yn-9-ylmethanol (0.1 g, 0.65 mmol, 1 eq) was dissolved in 2 mL DCM, then 4-nitrophenyl chloroformate (0.16 g, 0.8 mmol, 1.2 eq) and pyridine (54 µL, 0.65 mmol, 1 eq) was added and the reaction stirred at room temperature for 2 h, before it the solvent was removed under reduced pressure. The crude residue was purified by flash chromatography to result the product (0.15 g, 70 %) as a yellowish solid.

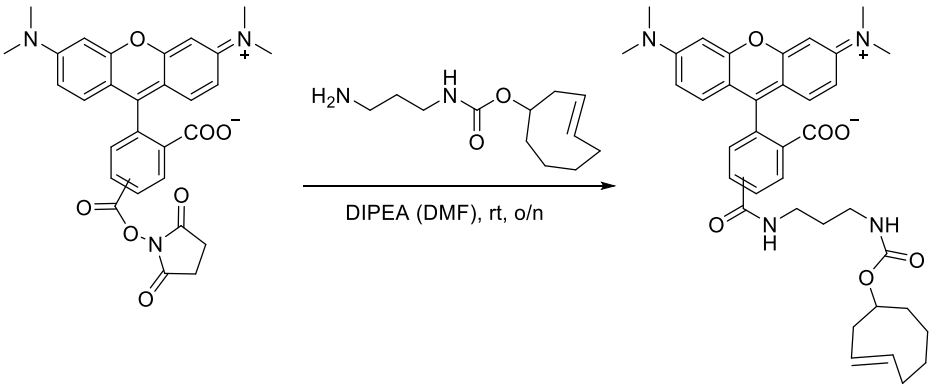
5-(and-6)-((5-(((bicyclo[6.1.0]non-4-yn-9-ylmethoxy)carbonyl)amino)pentyl)carbamoyl)-2-(6-(dimethylamino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)benzoate (BCN-TAMRA)



5-(and -6)-TAMRA-cadaverine isomer mixture (5 mg, 9.7 µmol, 1 eq) was dissolved in 2 mL DMF, followed by addition of DIPEA (10 µL, 58 µmol, 6 eq) and BCN-carbonate (9.2 mg, 29 µmol, 3 eq). The resulting reaction was stirred at room temperature overnight, then the solvent was evaporated under reduced pressure and the crude product purified by RP-HPLC (C18, 250 x 10 mm, 4 mL/min, 20-90 % ACN in 15 min) to obtain the product (4.7 mg, 70 %) as a purple solid, which was dissolved in DMSO to make a 2 mM stock solution, which was frozen at -20 °C.

RP HPLC (C18, 250 x 10 mm, 4 mL/min, 20-90 % ACN in 15 min**):** t_R = 11.5, 12.5 min
MS (ESI), *m/z* calc for $C_{41}H_{46}N_4O_6$ 690.34, found 691.3 $[M+H]^+$

1.5.8.2 Synthesis of (*E*)-5-(and-6)-((3-(((cyclooct-3-en-1-yloxy)carbonyl)amino)propyl)-carbamoyl)-2-(6-(dimethyl-amino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)benzoate (TCO-TAMRA)



5-(and -6)-TAMRA-succinimidyl ester isomer mixture (3 mg, 5.7 µmol, 1 eq) was added to a solution of (*E*)-cyclooct-3-en-1-yl (3-aminopropyl)carbamate (3.9 mg, 17 µmol, 3 eq) and DIPEA (6 µL, 37 µmol, 6 eq) in 2 mL DMF. The resulting reaction was stirred at room temperature overnight, then the solvent was evaporated under reduced pressure and the crude product purified by RP-HPLC (C18, 250 x 10 mm, 4 mL/min, 20-90 % ACN in 15 min) to obtain the product (1.9 mg, 52 %) as a purple solid, which was dissolved in DMSO to make a 2 mM stock solution, which was frozen at -20 °C.

RP HPLC (C18, 250 x 10 mm, 4 mL/min, 20-90 % ACN in 15 min**):** t_R = 11.2, 11.4 min
MS (ESI), *m/z* calc for $C_{37}H_{42}N_4O_6$ 638.31, found 639.3 $[M+H]^+$, 637.3 $[M-H]^-$

CHAPTER 2

Site-directed spin labeling via tetrazine ligation to probe protein structure

Chapter 2: Site-directed spin labeling via tetrazine ligation to probe protein structure

2.1 Aim

The aim of this project was to develop a site-directed spin labeling approach based on genetic code expansion together with our collaboration partners from the Sattler group (Department of Chemistry, TUM). We based our approach on site-specific incorporation of phenylalanine-based 1,2,4,5-tetrazine containing uAAs, which are more rigid compared to lysine based uAAs. Tetrazine-modified proteins should be amenable to chemoselective iEDDA cycloaddition with nitroxide spin labels bearing strained alkene/alkyne moieties after protein expression and purification, which should overcome the challenge of spin label reduction during cytosolic expression.

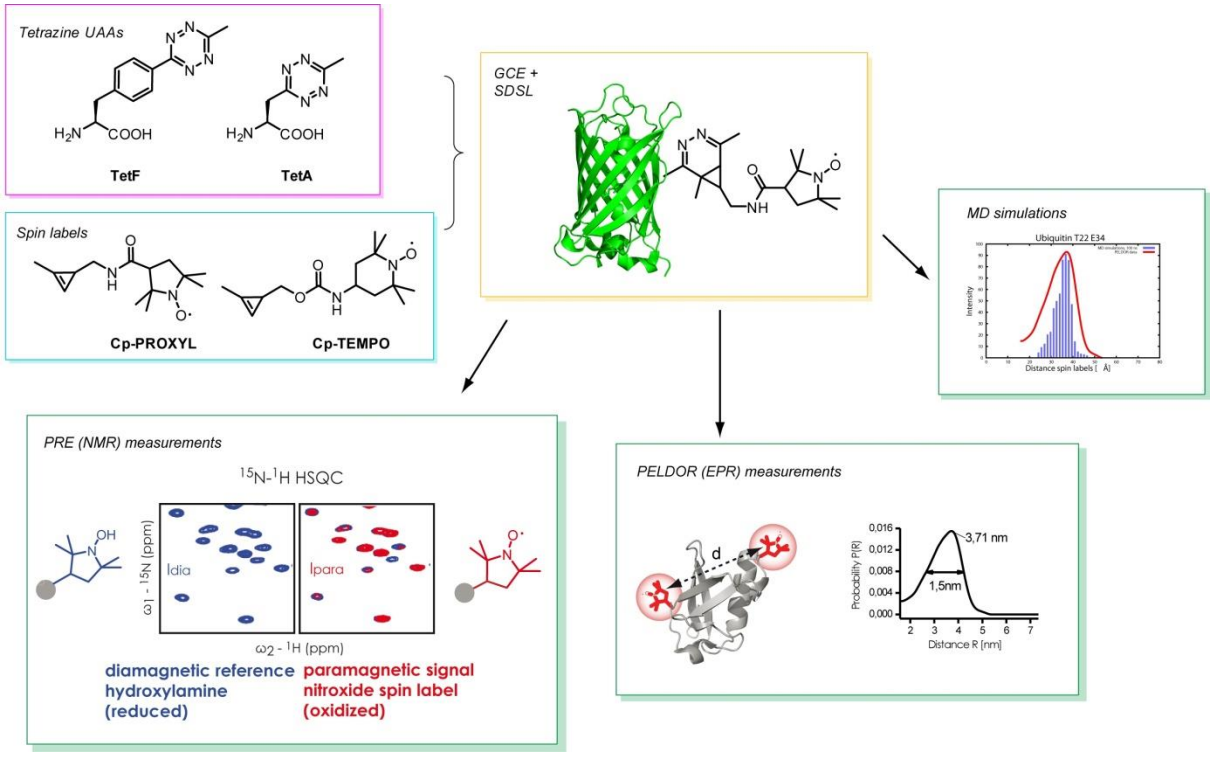


Figure 2.1: Site-directed spin labeling (SDSL) approach.
Tetrazine amino acids with a rather rigid side-chain are planned to be incorporated into proteins via GCE and the spin labels bearing a Cp moiety introduced via bioorthogonal iEDDA cycloaddition. PRE and PELDOR measurements will be used to gather structural information as well as to measure intra-protein distances, which will be validated using MD simulations.

As published uAA for iEDDA cycloaddition are based on the rather flexible and long lysine side chain,[60b, 310] here we aimed at synthesizing small and rather rigid tetrazine amino acids TetF and TetA. The synthesis of corresponding cyclopropene (Cp)-based nitroxide spin labels is based on published synthetic routes. The uAA CpK and corresponding tetrazine-based nitroxide spin labels will be utilized as comparison for a less rigid spin label system. Site-specific incorporation of the novel tetrazine uAAs into proteins will be tested using mutated tRNA synthetase/tRNA pairs from *M. barkeri* as well as *M. jannaschii*. Spin labeling and characterization of the resulting spin labeled proteins will be performed on model proteins sfGFP and ubiquitin by determining distances by pulsed electron electron resonance

(PELDOR). Molecular dynamics simulations performed will be used to verify of our SDSL approach.

Validation of the method is planned by applying it to the protein Loqs-PD, which contains several natural cysteine residues making a site-directed spin labeling approach using MTSSL (Figure 2.2a), which is commonly utilized, not applicable. In our approach, combining genetic code expansion of rigid tetrazine-amino acids and fast bioorthogonal reactions, amber stop codons will be placed at different positions in the two dsRBDs and in the linker region of Loqs-PD to site-specifically decorate Loqs-PD with spin labels. PRE-NMR experiments will be performed to determine domain orientation of Loqs-PD. PELDOR measurements are carried out to determine intra-protein distances in presence and absence of dsRNA.

2.2 General background

2.2.1 Loquacious-PD (Loqs-PD): processing of endogenous siRNA in *Drosophila melanogaster*

RNA interference (RNAi) is a highly conserved regulatory mechanism in multicellular organisms for the translational control of protein expression or viral defense by silencing mRNA via complementary interference with small RNA molecules.[311] Malfunction of the regulatory control can lead to cardiovascular diseases, neurological disorders and various types of cancer.[312] RNAi involves three different types of RNA: Piwi-interacting RNA (piRNA), whose mechanism of interference is not yet understood; microRNA (miRNA) and short interfering RNA (siRNA) that have their origin in longer double-stranded RNA (dsRNA) precursors, which are processed to shorter fragments.[313] Structurally, these two classes differ in the base pairing of the double strands; miRNA is mismatched, while siRNA shows perfect base pairing. Their mechanism however is the same: double strand RNA binding domains (dsRBDs) of double strand RNA binding proteins (dsRBPs) recognize the short dsRNAs and direct them to the Dicer enzyme. This endoribonuclease cuts the dsRNAs into smaller fragments of 21 to 25 basepairs, before they are loaded onto an Argonaute protein forming the RISC (RNA interference silencing complex) loading complex (RLC). Here, the passenger strand is degraded and the targeted mRNA silenced by generation of the RISC with the guide strand (interfering strand).[313]

In *Drosophila melanogaster* different dsRBPs are responsible for the recognition and guidance of siRNA to the Dicer enzyme. R2D2 guides siRNA from exogenous sources like viral infection, while Loquacious isoform d (Loqs-PD) performs that role for endogenous siRNA.[313a] Loqs-PD possesses two dsRBDs, with an αββββα topology each, which bind their substrate with similar affinities, with a slight preference for the thermodynamically more stable end of the siRNA. Upon binding to siRNA the two dsRBDs can slide along the RNA. The two dsRBDs are connected via a highly flexible linker of 45 amino acids that ensures proper arrangement of the domains.[314] The orientation of the domains towards each other and the distance between them remain to be elucidated.

2.2.2 Site-directed spin labeling (SDSL)

The site-specific introduction of unpaired electrons into proteins and their subsequent analysis in an external magnetic field helps understanding structure, dynamics and conformational changes of proteins.[315] Electron paramagnetic resonance (EPR) spectroscopy allows to determine dynamics due to motion, solvent accessibility, as well as intra- and intermolecular distances via the site-directed introduction of one or more spin labels, [315-316] while paramagnetic relaxation enhancement (PRE) NMR experiments detect the presence of unpaired electrons, due to faster relaxation of protons in a vicinity of up to 20 Å encompassing the spin label.[317] This can be used to study changes in structure of proteins.

The most common method to place spin labels in a site-directed manner into proteins (Figure 2.2) consists in targeting cysteines (natural occurring or introduced by mutagenesis) with cysteine-reactive nitroxide radicals like 1-oxyl-2,2-5,5-tetramethylpyrroline-3-methyl-methane-thiosulfonate (MTSSL, Figure 2.2a). As cysteines may occur multiple times in a

proteins targeting these can lead to labeling with multiple spin labels, abolishing the specificity of the approach. Knockout of these cysteine might lead to disruption of structure and loss of function.[315] Therefore several other methods have been reported in recent years that combine genetic code expansion with bioorthogonal reactions to install spin labels in a site-specific manner.[60a, 318] One method published in 2009 is based on the incorporation of p-acetylphenylalanine, which is modified with hydroxylamine-containing spin labels (Figure 2.2c). Aniline catalysis allows the reaction to take place at pH 6, but due to the slow reaction rates a high excess of label is necessary.[60a] Another approach reported in 2014 relies on the direct incorporation of a lysine-based spin label into the protein, where the nitroxide radical is tethered to the ε-amino group via an amide or a carbamate (Figure 2.2b). This resulted in partial reduction of the spin label during protein expression due to the reducing environment in *E. coli*. The high rotational freedom of the lysine linker gave also rise to line broadening, which complicated analysis of the resulting data.[60b, 319] Newer reports decorate encoded *p*-iodophenylalanine via Suzuki-Miyaura cross-coupling reaction with nitroxide spin labels or use CuAAC reaction between azides and alkynes for the spin labeling of proteins *in vitro* (Figure 2.2c).[320] The conditions of CuAAC based spin labeling are problematic in itself, as the use of excess of ascorbate and prolonged reaction times lead to partial reduction of the nitroxide radical. Drescher and coworkers tested different phenylalanine based uAAs bearing azide or alkyne moieties, differentiating in the linkage between the aromatic ring and the reactive group. An allyl linkage displayed better labeling yields than direct attachment of the alkyne to the aromatic ring (98 vs. 57 %), but in turn lead to line broadening due to higher flexibility of the linker. Furthermore, the labeled protein had to be treated with EDTA to remove Cu(I) before EPR measurements.[320b] The same group also applied a Suzuki-Miyaura cross-coupling reaction for SDSL (Figure 2.2c). The resulting biphenyl linkage displays less rotational freedom than the MTSSL label, but the reaction either demands a higher excess of labeling reagent and catalyst or a prolonged reaction time overnight at 37 °C, which is not applicable to a wide range of proteins. Labeling also does not seem to work to completion, as some of the incorporated *p*-iodophenylalanine is deiodinated and the spin counting on the double-labeled thioredoxin variant indicates 18 % spin labeled protein.[321] Very recently, Drescher and coworkers also developed an photo-inducible spin labeling approach (Figure 2.2d). For this, the group incorporated SCOK and TCOK into proteins, followed by subsequent labeling via iEDDAC with a photo-caged nitroxide derivative of 3,6-dipyridyl-1,2,4,5-tetrazine. Light irradiation at 302 nm induced cleavage of the photocage to free hydroxylamine that spontaneously oxidized to the nitroxide radical.[310] An advantage of this approach is the photo-inducibility of the spin label, which renders this stable until deprotected using light. However, the deprotected spin label was observed to completely degrade in *E. coli* lysate within 80 min and spin concentration was not fully recovered upon irradiation. The uAAs used to tether the spin label via iEDDAC are flexible lysine based uAAs with a higher rotational freedom, which could lead to line broadening.

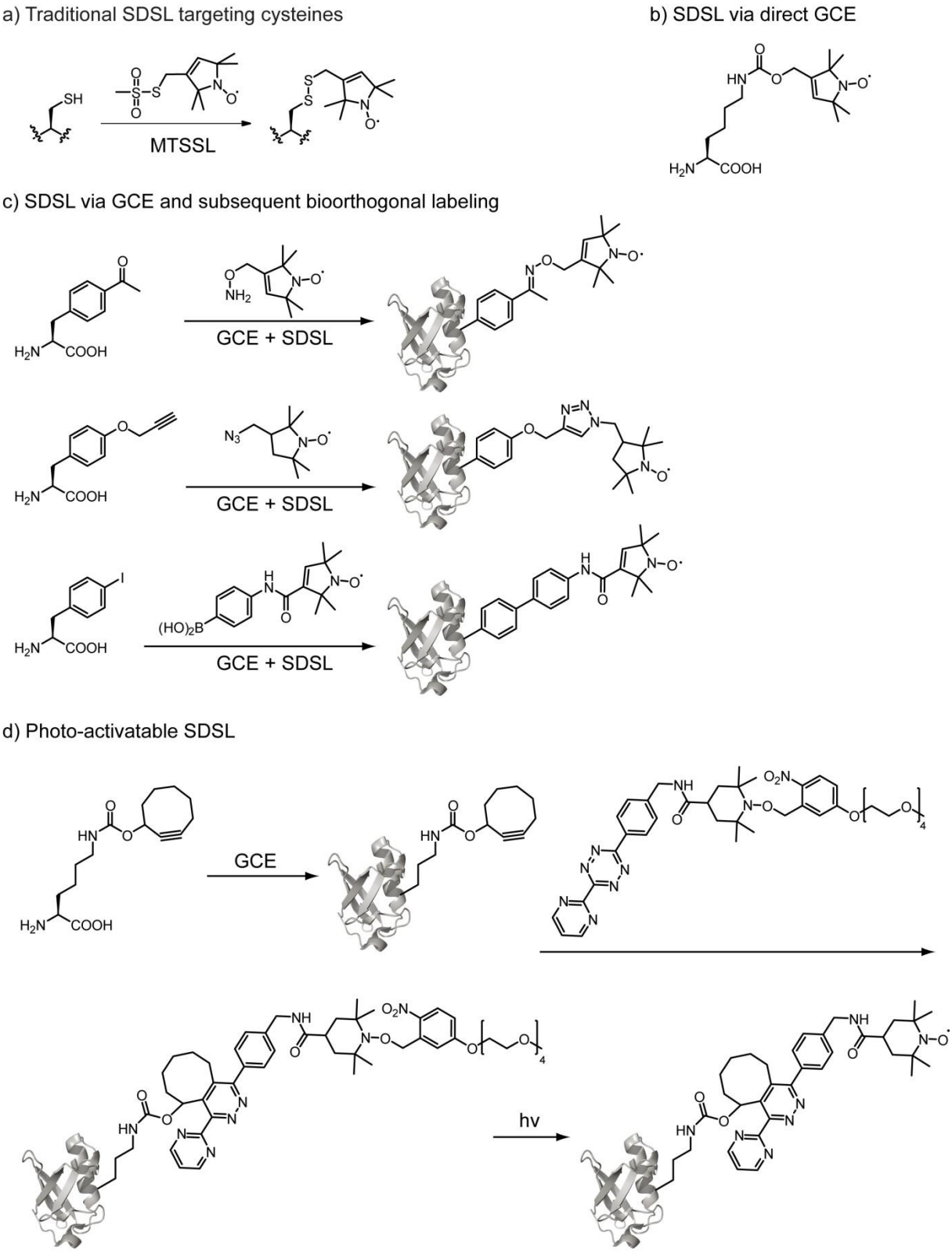


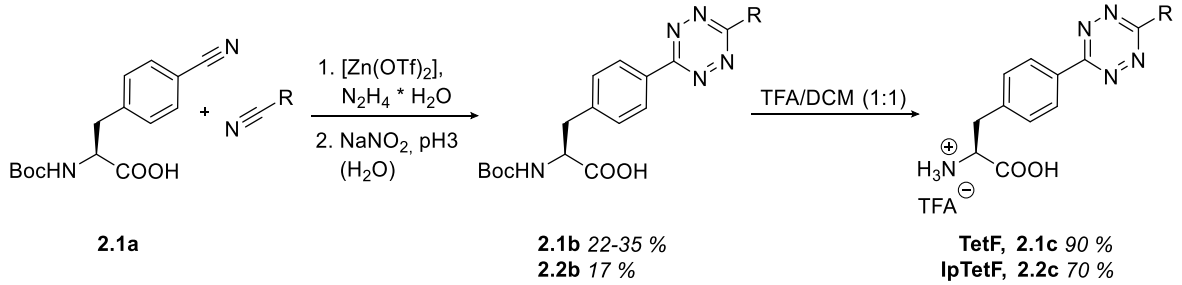
Figure 2.2: Various reported methods for site-directed spin labeling.
a) Labeling of cysteines with sulfhydryl-reactive nitroxide spin label MTSSL b) Direct incorporation of nitroxide bearing uAA c) Decoration of different incorporated uAA with spin labels via bioorthogonal labeling. Top: Oxime ligation, middle: CuAAC, bottom: Suzuki-Miyaura cross-coupling reaction.

2.3 Results and Discussion

2.3.1 Synthesis routes for amino acids and spin labels

2.3.1.1 Synthesis routes for TetF

As a first step taken in this project several unnatural amino acids bearing bioorthogonally reactive tetrazine moieties were developed. The design of the 1,2,4,5-tetrazine analogs of biphenyl-alanine was based on the publications of Devaraj *et al.*[322] and Blizzard *et al.*[151b] The synthetic approach (Scheme 2.1) used, condenses Boc-(4-cyano)-L-phenylalanine with acetonitrile or isobutyronitrile and hydrazine hydrate using a Lewis acid catalyst to form the asymmetric 3,6-substituted 1,4-dihydrotetrazine amino acids. Oxidation to the tetrazine occurred by using sodium nitrate in an acidic environment. The free amino acid was obtained by Boc deprotection with trifluoroacetic acid in DCM.



Scheme 2.1: Synthetic strategy for the generation of tetrazine uAAs based on nitrile containing starting materials. R = Methyl, Isopropyl

The yields reportet by Devaraj *et al.* could be improved using a syringe pump for the addition of the second nitrile to the reaction mixture containing Boc-(4-CN)-Phe-OH (**2.1a**). The resulting lower local concentration in the reaction mixture reduced the production of the symmetric tetrazine and improved the yields of the tetrazine amino acids bearing methyl (TetF, **2.1c**) or isopropyl (IpTF, **2.2c**) substituents by 10 % to 35 %. Blizzard and coworkers obtained the amino acid TetF in 60 % yield, most likely due to the use of anhydrous hydrazine, compared to this approach, which is not available in Europe due to its toxicity and explosion potential.

To improve the overall yield of 30 % over two steps, alternative synthetic strategies were developed. Instead of generating the tetrazine moiety directly from the amino acid starting block, a new C-C bond was formed between amino acid precursor and a tetrazine. At first, a *β*-CH arylation approach using L-alanine with hydroxylamine[323] or 2-methylthioaniline[324] directing groups on the C-terminus was chosen. There, $Pd(OAc)_2$ as the catalyst, coordinates onto heteroatoms of the directing group and can then activate one of the CH-bonds of the alanine side chain by liberation of an acetate ligand as acetic acid. Subsequently, oxidative addition of aryl iodide is facilitated by complexation with silver acetate, followed by reductive elimination and release of the monoarylated product (Figure 2.3).[324]

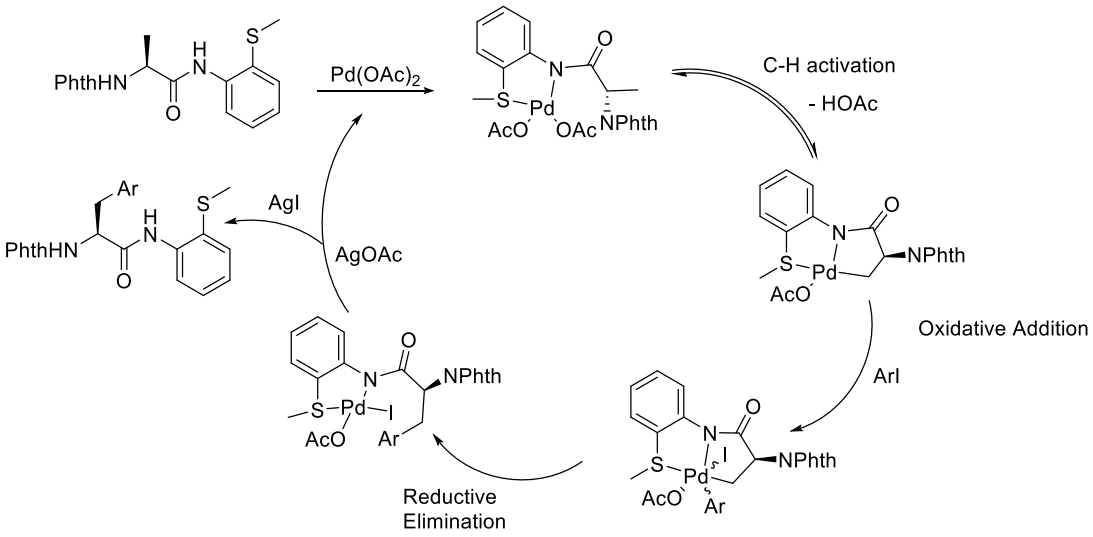
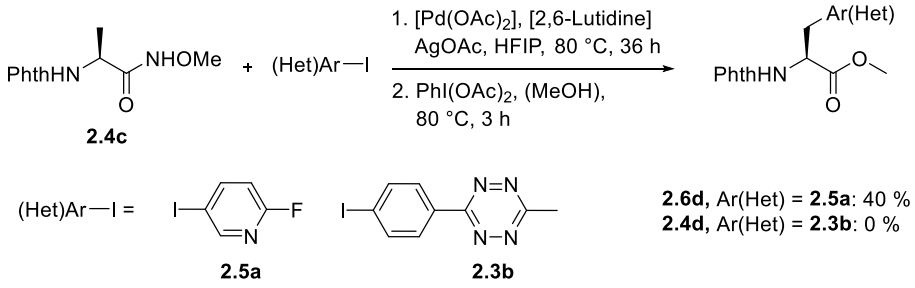


Figure 2.3: Mechanism of a *β*-CH arylation using 2-methylthioaniline directing group

For performing *β*-CH arylation reactions starting material 3-(4-iodophenyl)-6-methyl-1,2,4,5-tetrazine (**2.3b**) was prepared after the instructions by Yang *et al.*[187] with a yield of 58 %, which was according to literature.[325] The reaction partners were synthesized in two steps each in approx. 90 % according to literature.[323-324, 326] Here, L-alanine (**2.4a**) was subsequently protected at the *N*- terminus with phthalimide (**2.4b**) and the *C*-terminus with the corresponding directing group, either methoxyamine (**2.4c**) or 2-(methylthio)aniline (**2.5a**).

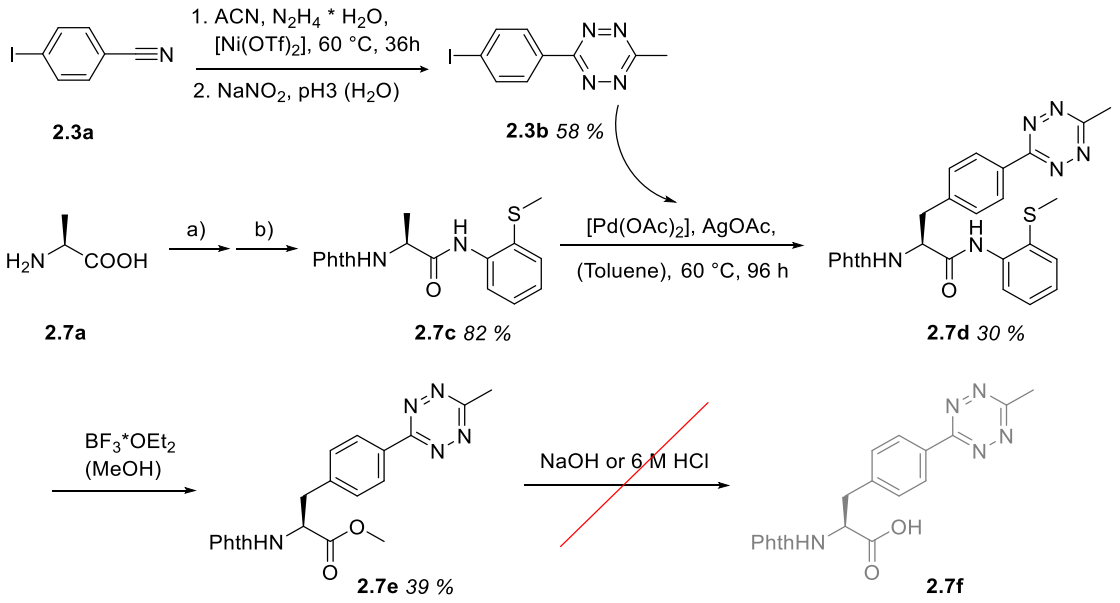
The *β*-CH arylation reaction directed by *C*-terminal methoxyamide had to be carried out in a glass tube, which was sealed by melting the glass, since the reaction temperature was higher than the boiling point of the used solvent. The synthetic scheme is displayed in Scheme 2.2. At first, the reaction was performed with readily available 2-fluoro-5-iodopyridine (**2.5a**) as heteroaryl iodide to test the reaction conditions. Optimization then gave the mono-5-pyridinylated product (**2.6d**) in 40 % yield. The same reaction conditions were applied for the synthesis using 3-(4-iodophenyl)-6-methyl-1,2,4,5-tetrazine (**2.3b**) as aryl iodide, however product formation could not be observed. This might be due to a strong coordination of the tetrazine to the palladium and therefore poisoning of the catalyst.[323]



Scheme 2.2: Synthesis scheme for methoxyamine directed *β*-CH Arylation

To test, if a different directing group could facilitate the reaction with 3-(4-iodophenyl)-6-methyl-1,2,4,5-tetrazine (**2.3b**), 2-(methylthio)aniline was introduced on the C-terminus of phthaloyl-L-alanine. Scheme 2.3 shows the synthetic route using the 2-(methylthio)aniline directing group. After 96 h product formation was checked by LC-MS analysis and resulted in isolation of 30 % monoarylated product, which was transformed into a methyl ester by boron

trifluoride diethyl etherate in 39 % yield. The 2-(methylthio)aniline directing group seems to be able to facilitate product formation, but catalyst poisoning might still be responsible for low yields. Hydrolysis of the methyl ester in basic conditions led to decomposition of the tetrazine by loss of gaseous nitrogen, which was also observed by refluxing in half concentrated hydrochloric acid.



a) *N*-ethoxycarbonylphthalimid, Na_2CO_3 (H_2O); b) 1. Oxalyl chloride, [DMF] (DCM), 2. 2-(methylthio)aniline, 2,6-lutidine (DCM/H_2O)

Scheme 2.3: Synthesis of protected TetF by 2-methylthioaniline directed *β*-CH Arylation

Due to problems with deprotection and an overall yield of 6 % after 5 steps, Negishi cross-coupling reactions[327] were investigated as a synthetic alternative. There, a new C-C bond is formed between an organohalide and an organozinc halide under palladium or nickel catalysis. The catalytic cycle (Figure 2.4) starts with an oxidative addition of the organohalide to palladium, followed by a transfer of the organozinc halide from zinc to palladium by transmetallation and release of zinc halide. In a reductive elimination step the product is formed and the catalyst regenerated.[327] The organozinc species is generated *in situ* by iodine facilitated zinc insertion into the C-halide bond of the starting material, since the obtained organozinc halides are moisture and air sensitive.[328]

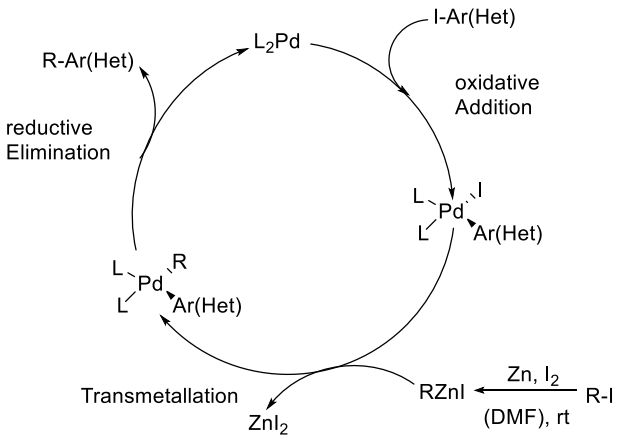
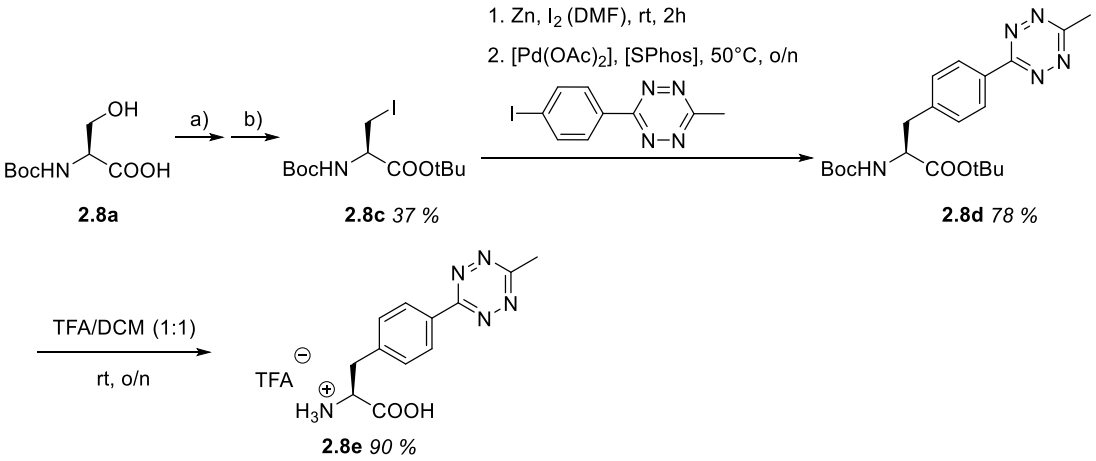


Figure 2.4: Mechanism of Negishi cross-coupling reaction[327]

Based on these requirements a synthetic route for TetF comprising 5 steps was developed (Scheme 2.4). Boc and OtBu protected *β*-iodo-L-alanine (**2.8c**) was synthesized starting from Boc protected L-serine (**2.8a**), in two steps, first by esterification with *tert*-butyl 2,2,2-trichloroacetimidate and subsequent Appel type reaction using iodine as the halogen source giving the product in 37 % yield over two steps in accordance with literature.[329]



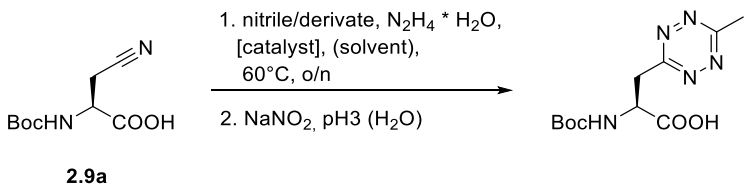
Scheme 2.4: Synthesis of TetF by Negishi cross-coupling reaction

The *in situ* generation of the organozinc reagent from Boc-*β*-iodo-L-alanine-OtBu (**2.8b**) with zinc and iodine in DMF was monitored by TLC and complete after 2 h at room temperature. DMF was used as a solvent since it suppresses the coordination of the Boc carbonyl moiety to zinc in comparison to THF and therefore enhances zinc insertion into the C-I bond.[329a] The subsequent cross-coupling reaction yielded the protected TetF amino acid (**2.8d**) in 78 %, followed by almost quantitative (90 %) deprotection with TFA in DCM overnight. TetF (**2.8e**) was successfully synthesized via Negishi cross-coupling reaction in 15 % overall yield over 5 steps. However, this alternative synthesis route was not able to improve the overall yield of 30 % obtained by condensation of the nitrile compounds and was therefore not repeated, it gave however valuable insights into synthesis of TetA.

2.3.1.2 Synthesis routes for TetA

2.3.1.2.1 Nitrile and nitrile derivatives as starting materials

To synthesize the methyltetrazine analog TetA of phenylalanine in a first attempt Boc-*β*-cyano-L-alanine (**2.9a**) was condensed with nitriles or nitrile derivates and hydrazine hydrate as shown in the following Scheme 2.5.



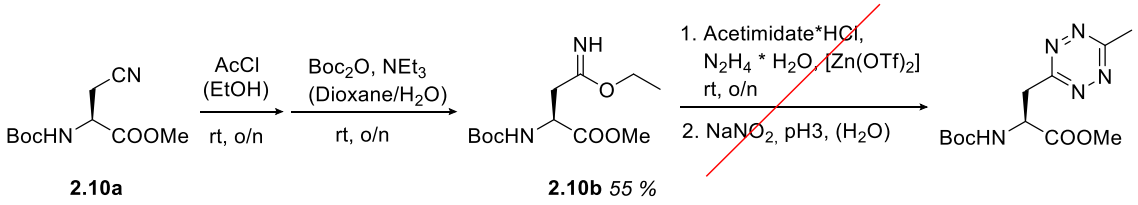
Scheme 2.5: Synthetic route for TetA via nitrile condensation with hydrazine

Standard reaction conditions developed by Yang *et al.*[187] were adapted for the formation of TetA facilitated by transition metal salts. These include the use of 10 eq nitrile or nitrile derivate, 0.5 eq catalyst in 25 eq hydrazine, which also functions as solvent for the reaction stirred at 60 °C for 24 h. Addition of $Zn(OTf)_2$ as catalyst to the reaction did not result in any product formation. Therefore, different reaction conditions (Table 2.1) were screened and evaluated. The application of $ZnCl_2$ did yield product, but in very low amounts. Elemental sulfur, initially reported as catalyst for the formation of symmetric tetrazines with aromatic substituents,[174a, 175] but later applied in optimized reaction conditions to generate asymmetrically substituted tetrazines (Table 2.1),[174b] was also tested, if it could facilitate tetrazine formation between Boc-*β*-cyano-L-alanine (**2.9a**) and nitriles or nitrile derivatives. Unfortunately, no product formation could be observed in any of the reaction mixtures where sulfur was added. Only increasing Boc-*β*-cyano-L-alanine (**2.9a**) compared to the nitrile did lead to slightly better yields, which was not sufficient enough to synthesize TetA in gram scale. The reactivity of acetonitrile and isobutyronitrile in tetrazine condensation reactions is similar. Due to these encountered difficulties, alternative synthetic approaches were developed.

Table 2.1: Tested reaction conditions for TetA via nitrile condensation with Boc-*β*-cyano-L-alanine and hydrazine

Nitrile/nitrile derivate	Ratio aa/nitrile	catalyst	solvent	Yield [%]
Acetonitrile	1:10	$Zn(OTf)_2$	-	-
Acetonitrile	1:10	$ZnCl_2$	-	3
Isobutyronitrile	5:1	$Zn(OTf)_2$	-	10
Acetonitrile	1:10	Sulfur (4 eq)	-	-
Acetonitrile	1:10	Sulfur (4 eq)	-	-
Acetamidine HCl	1:10	Sulfur (4 eq)	EtOH	-

Test reactions performed with different phenyl or methyl bearing imidate and amidine salts, first described by A. Pinner[330], lead to formation of asymmetrically substituted 3-methyl-6-phenyl-1,2,4,5-tetrazine at room temperature, while reactions with the corresponding nitriles were carried out at elevated temperatures of 50-70 °C. This seems to indicate a higher reactivity of the imidate and amidine salts in condensation reactions with hydrazine compared to nitriles, possibly explaining the low yields obtained with alkyl nitrile starting materials. Therefore, an alternative approach for the formation of TetA was based on transforming the cyano moiety of Boc-*β*-cyano-L-alanine into an imidate to increase its reactivity (Scheme 2.6).



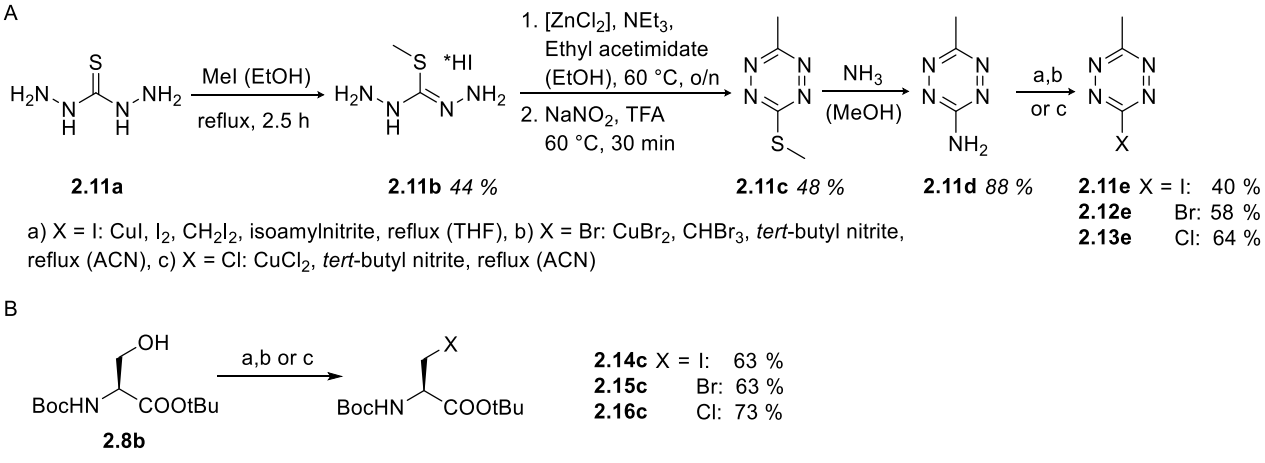
2.10a **2.10b** *55 %*

Scheme 2.6: Conditions for TetA formation from ethyl imidate starting materials

Formation of the Pinner salt was achieved with acetyl chloride in ethanol, generating HCl *in situ*.[331] This leads to protonation of the nitrile, which is then attacked by the alcohol to obtain the ethyl imidoester hydrochloride salt. Since the acidic environment led to deprotection of the Boc group, the Boc group had to be reintroduced subsequently resulting in formation of the *N*- and *C*-terminally protected methyl-*ß*-(ethoxy(imino)methyl)-L-alaninate (**2.10b**) in 55 % yield over two steps. Unfortunately, asymmetric tetrazine formation of the wanted TetA amino acid by condensation of imidates with hydrazine was not observed, only formation of the favorable, symmetrically substituted 3,6-dimethyl-1,2,4,5-tetrazine was confirmed.

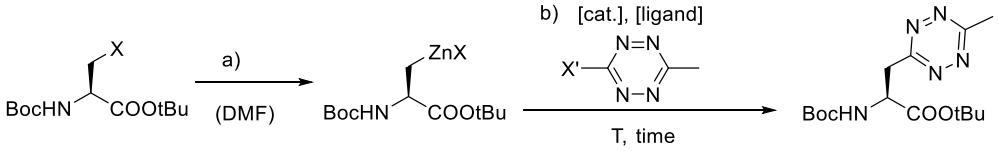
2.3.1.2.2 Negishi cross-coupling reactions

Since TetA was not synthetically accessible by condensation of nitriles or nitrile derivates with hydrazine hydrate, the Negishi cross-coupling reaction applied during synthesis of TetF, was considered as an alternative approach for the synthesis of TetA. The fact that palladium facilitated cross-coupling reactions have been performed before using tetrazines[185-186] and that the triazine analog of TetA by Horner *et al.*[332] was synthesized using the Negishi cross-coupling, supported those synthetic plans. Firstly, the organohalides, in this case 3-halo-6-methyl-1,2,4,5-tetrazines (Figure 2.7) were synthesized in 4 steps from thiocarbohydrazide in overall yields of 7.5-12 % based on literature procedures.[180a, 333] Methylation of the sulfur atom of thiocarbohydrazide was followed by zinc(II) facilitated condensation with ethyl acetimidate and subsequent oxidation to the 3-methylthio-6-methyl-1,2,4,5-tetrazine (**2.11**c). Reaction with methanolic ammonia solution replaced the 3-methylthio substituent with amine by nucleophilic aromatic substitution. Halo-deamination was achieved by diazotization with isoamyl- or *tert*-butyl nitrite and Cu(I) or Cu(II) halogen salts in combination with diiodomethane or bromoform. The protected *β*-halo-L-alanine cross-coupling partners were synthesized through an Appel like reaction of protected L-serine (**2.8b**) with the corresponding halide sources and triphenylphosphine in 63-73 % yield.



Scheme 2.7: Synthesis of A) 3-halo-6-methyl-1,2,4,5-tetrazines and B) protected *β*-halo-L-alanines

The first Negishi cross-coupling reaction used in the synthesis towards of TetA was performed with iodine substituents on both starting materials and a catalyst-ligand combination of $Pd(OAc)_2$ and SPhos identical to the reaction conditions used for the synthesis of TetF. However, this approach resulted in only 20 % product formation.



Scheme 2.8: Reaction scheme for Negishi cross-coupling reactions for Tet. Conditions a) and b) are defined in **Table 2.2**

Therefore different Negishi cross-coupling reaction conditions (Scheme 2.8:a and b) were screened (Table 2.2). A prolonged reaction time from 5 h to overnight improved the yield to 37 %, while changing the temperature or increasing the amount of 3-iodo-6-methyl-1,2,4,5-tetrazine (**2.11e**) had no effect. Switching to a different catalyst was also not giving any improvement. Likewise, the use of 3-methylthio substituted 6-methyl-1,2,4,5-tetrazine (**2.11c**) with the standard catalyst-ligand combination of $Pd(OAc)_2$ and SPhos at different temperatures, as well as its substitution with $Ni(acac)_2$ and DPEPhos according to a procedure from Melzig *et al.*[334] did not yield any cross-coupled product. The use of 3-chloro- and 3-bromo-6-methyl-1,2,4,5-tetrazine (**2.13e or 2.12e**) together with *N*- and *C*-terminally protected *β*-iodo-L-alanine (**2.14c**) allowed the isolation of protected TetA, but in decreased yields, especially in the case of the chloro-substituted compound, which gave less than 10 % product. Also, execution of the reaction with protected *β*-chloro or *β*-bromo-L-alanine (**2.16c or 2.15c**) did not yield any product. In this case, the insertion of zinc into the C-halogen bond could not be achieved with protected *β*-chloro-L-alanine (**2.16c**), no matter what activation conditions used.[328] For protected *β*-bromo-L-alanine (**2.15c**) TLC indicated activation with Zn at 80 °C, however product formation was not observed. These observations are in accordance with the known decrease in reactivity from I > Br > Cl.[335]

Table 2.2: Tested reaction conditions for TetA cross-coupling; bold type indicates the reaction conditions with the best yield

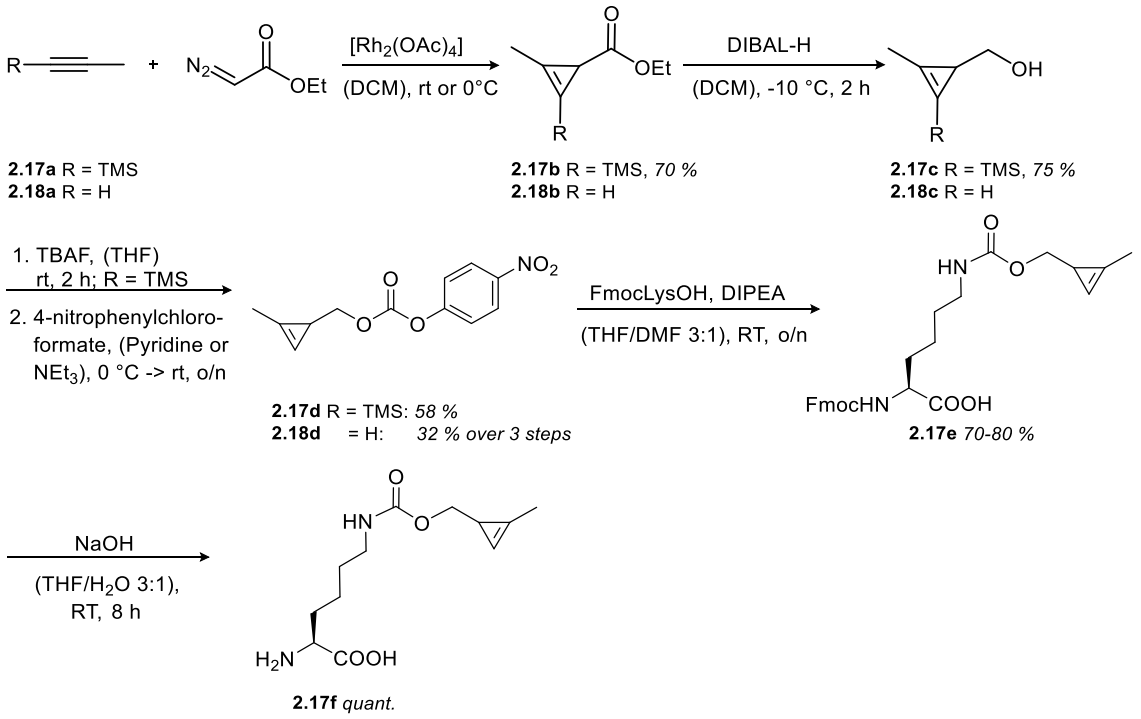
X	activation (a)	X‘	[cat] (b)	[ligand] (b)	T [°C]	time	Yield [%]
I	3 eq Zn, I_2, rt, 2 h	1.3 eq I	$Pd(OAc)_2$	SPhos	50	5 h	20
I	3 eq Zn, I_2, rt, 2 h	1.3 eq I	$Pd(OAc)_2$	SPhos	50	o/n	30
I	3 eq Zn, I_2, rt, 2 h	1.8 eq I	$Pd(OAc)_2$	SPhos	60	o/n	22
I	3 eq Zn, I_2, rt, 2 h	2 eq I	$Pd(OAc)_2$	SPhos	50	o/n	26
I	3 eq Zn, I_2, rt, 2 h	1.3 eq I	$Pd(OAc)_2$	SPhos	50	o/n	37
I	1.5 eq Zn, I_2, rt, 2 h	1.3 eq I	$Pd(OAc)_2$	SPhos	75	o/n	32
I	3 eq Zn, I_2, rt, 2 h	1.3 eq I	$Pd_2(dba)_3$	$(o\text{-Tol})_3P$	50	o/n	10
I	1.5 eq Zn, I_2, rt, 2 h	SMe	$Pd(OAc)_2$	SPhos	25	o/n	-
I	1.5 eq Zn, I_2, rt, 2 h	SMe	$Pd(OAc)_2$	SPhos	50	o/n	-
I	1.5 eq Zn, I_2, rt, 2 h	SMe	$Ni(acac)_2$	DPEPhos	50	o/n	-
I	1.5 eq Zn, I_2, rt, 2 h	Cl	$Pd(OAc)_2$	SPhos		o/n	<<< 10

I	1.5 eq Zn, I_2, rt, 2 h	Br	$Pd(OAc)_2$	SPhos		o/n	15
I	**1.5 eq Zn, I_2, rt, 2 h**	**1.3 eq I**	**$Pd(OAc)_2$**	**Tris(2-furyl)P**	**50**	**o/n**	**60**
Cl	3 eq Zn, I_2, RT, 2 h		-	-	-	-	-
Cl	1.5 eq LiCl, 1.5 eq Zn, Dichloroethane, TMSCl, rt o/n -> 50 °C o/n		-	-	-	-	-
Br	3 eq Zn, I_2, RT, 2 h				-	-	-
Br	3 eq Zn, I_2, 80 °C, o/n	I	$Pd(OAc)_2$	SPhos		o/n	-

A new literature search on Negishi cross-coupling reactions resulted in a paper[336], where tris(2-furyl)phosphine was used as a ligand. It is usually applied as a ligand in Stille reactions, where it enhances the transmetallation[337] step by being a better π-acceptor towards Pd(0) and a less effective σ-donor towards Pd(0) and Pd(II) in comparison to triphenylphosphine. It was tested in combination with $Pd(OAc)_2$ instead of SPhos and lead to product formation in 60 % yield. This was an almost 2-fold improvement in yield compared to the use of SPhos as a ligand and lead to the successful synthesis of TetA in multigram scale.

2.3.1.3 Synthesis of CpK

To compare the newly developed amino acids TetF and TetA to established bioorthogonally reactive uAAs, with a more flexible linker and therefore higher rotational freedom, CpK (**2.17f**) was synthesized based on literature[158, 192, 196, 198b] starting from propyne gas or the less volatile trimethylsilylpropyne.



Scheme 2.9: Synthesis of CpK

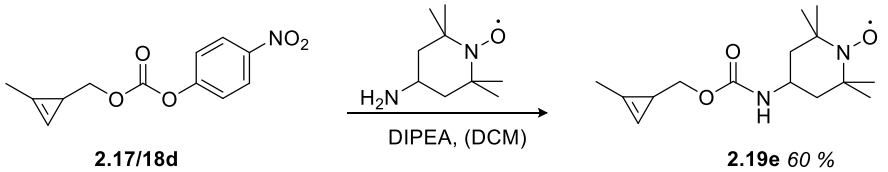
The first step builds the reactive cyclopropene moiety by rhodium-catalyzed [2+1] cycloaddition between a propyne moiety and ethyl 2-diazo acetate. The reaction could be optimized in the amount of catalyst used, since a 10-fold decrease also gave the product in comparably good yields. Reduction of the ester (**2.17/18b**) to the alcohol (**2.17/18c**) was performed with 2 equivalents of DIBAL-H at -10 °C to avoid reduction of the cyclopropene to cyclopropane. When trimethylsilyl-protected propyne (TMS-propyne) was used, the TMS moiety was necessary to be deprotected before activation of the alcohol (**2.17c**) with 4-nitrochloroformate, as a later deprotection approach was problematic. This could be achieved in good yields by using TBAF in THF. The deprotected product was purified by flash chromatography before activation, since this led to better yields of activated product, however it increased the volatility the resulting alcohol (**2.17d**). The coupling step to the ε-amino group of lysine and the subsequent deprotection were achieved in good yields according to literature. For use of propyne gas, yields could only be calculated from the activation step onwards as the previous products were very volatile components. Comparison of both synthesis routes proved the approach starting with propyne gas to be superior in yields; however the handling was more difficult than starting with TMS-propyne.

2.3.1.4 Synthesis routes for spin labels

After successful synthesis of the uAAs, which are planned to be genetically encoded into proteins for the site-specific chemoselective decoration with nitroxide spin labels, the corresponding reaction partners bearing the nitroxide moieties were synthesized. The synthetic procedures are discussed in more detail in the following chapters.

2.3.1.4.1 Cp-based spin labels

The route followed for the synthesis of CpK was used as a basis for the synthetic routes towards cyclopropene spin label conjugates. 4-amino-TEMPO was attached to the cyclopropene moiety via a carbamate linkage in 60 % yield to obtain Cp-TEMPO (**2.19e**) by nucleophilic attack of the primary amine on the carbonyl carbon of activated Cp carbonate and release of 4-nitrophenol (Scheme 2.10).

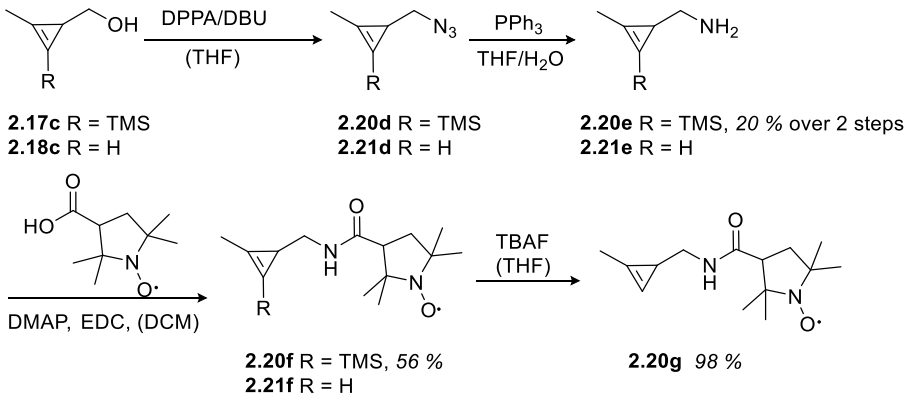


2.17/18d **2.19e** *60 %*

Scheme 2.10: Synthesis of Cp-TEMPO

In order to attach 3-carboxy-PROXYL onto the cyclopropene moiety the primary alcohol, either carrying a TMS group or proton in position 2 on the Cp ring depending on the synthesis route described in Scheme 2.9, was transformed into an azide, which was then reduced to the primary amine in 20 % yield over two steps. The coupling reaction was facilitated by optimized Steglich conditions in 56 % yield followed by a quantitative TMS deprotection with TBAF in THF (Scheme 2.11) in case of a TMS protecting group. Traces of residual TBAF even after purification of Cp-PROXYL (**2.20g**) might lead to problems in the labeling

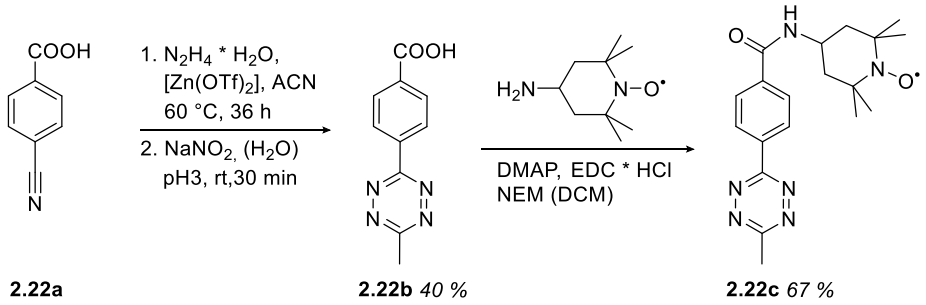
reaction with tetrazine moieties; therefore the route without a TMS substituent was preferred, even though the handling is more difficult due to volatility of the intermediate products.



Scheme 2.11: Synthesis of Cp-PROXYL. The last step is only necessary if R = TMS.

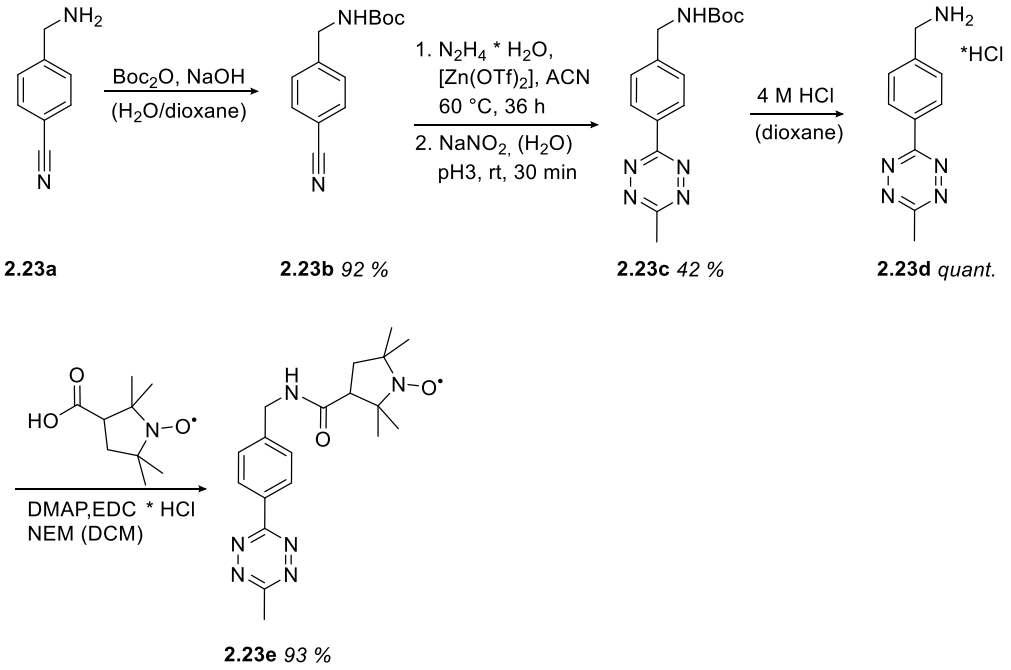
2.3.1.4.2 Tetrazine based spin labels

For spin labeling of CpK with iEDDAC reaction, nitroxide radicals TEMPO and PROXYL were attached onto tetrazines. 3-(4-Carboxyphenyl)-6-methyl-1,2,4,5-tetrazine (**2.22b**), obtained in 40 % yield by standard tetrazine reaction conditions, was coupled to 4-amino-TEMPO under optimized Steglich conditions to result Tet-TEMPO (**2.22c**) in 67 % yield (Scheme 2.12).



Scheme 2.12: Synthesis of Tet-TEMPO

Tet-PROXYL (**2.23e**) in turn was synthesized starting from 4-aminomethylbenzonitrile (**2.23a**), which was Boc-protected before formation of the tetrazine, followed by Boc deprotection and coupling to 3-carboxy-PROXYL under optimized Steglich conditions and obtaining the product in 36 % yield over 4 steps (Scheme 2.13).



Scheme 2.13: Synthesis of Tet-PROXYL

2.3.2 Biochemical experiments

2.3.2.1 Test expressions with PylRS mutants for TetF and TetA

After successful synthesis of the tetrazine bearing amino acids TetF and TetA site-specific incorporation into proteins via amber suppression was tested. First, a number of engineered PylRS/tRNA pairs from *M. barkeri* (D78, D171, D173, D197, see Table II.4), available in the lab, were chosen due to structural resemblance of their preferred substrates. Therefore, incorporation tests with sfGFP-N150TAG-His6 were performed in the presence of uAAs TetF and TetA. Cultures where no uAA was added served as a negative control. Only for synthetase D173 (Fig. 2a) full-length protein with molecular mass of approx. 27 kDA could be observed, but the incorporation seemed to be unspecific, since there was also a band visible, where no uAA was added. For all other synthetases (Figure 2.5) only truncated sfGFP (MW ~ 14 kDA) was produced, showing none of the natural amino acids was introduced instead of uAA. The incorporation of BocK by the wild type PylRS synthetase (D4)[70c], was used as a positive control for the production of full-length sfGFP (Figure 2.5a: lane 3, Figure 2.5b: lane 5).

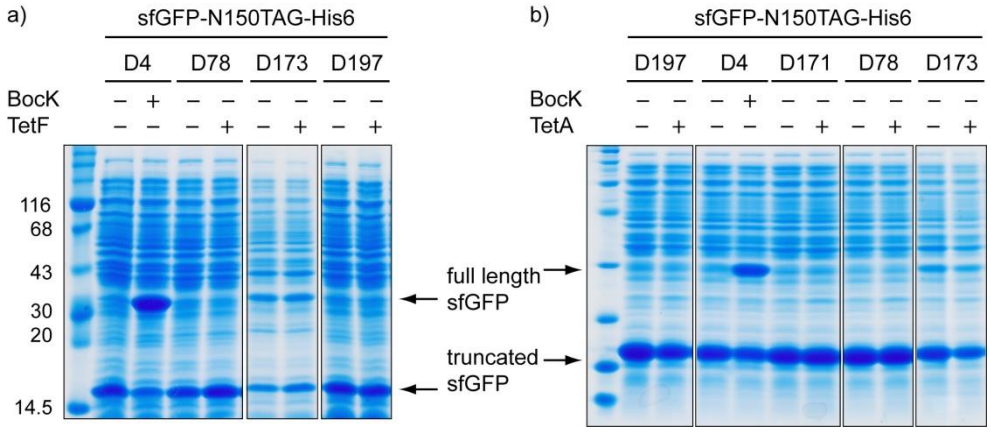
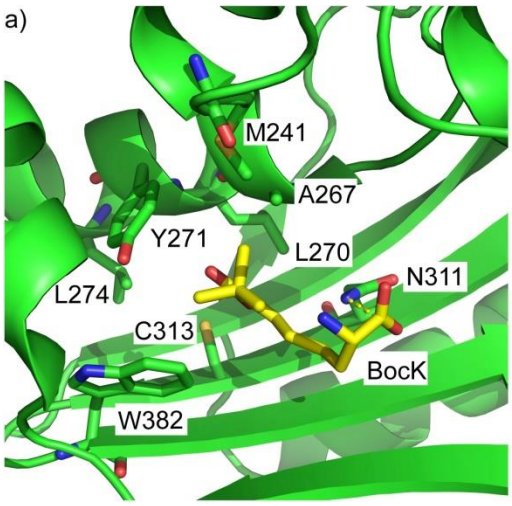


Figure 2.5: Coomassie stained SDS-PAGE of test expressions of sfGFP-N150TAG-His6 using different PylRS mutants with 2 mM a) TetF b) TetA.

Since none of the available mutant *Mb* PylRS/tRNA pairs were able to incorporate TetF or TetA into sfGFP new synthetase mutants were planned by rational, structure-guided design. Mutations were chosen based on the crystal structure of the *C*-terminal catalytic domain of the *M. mazei* PylRS in complex with BocK (Figure 2.6a). As the wt PylRS from *M. mazei* and *M. barkeri* only differ in 35 amino acids in the *C*-terminal lobe, but not in the residues building the active site, which are conserved, both enzymes are used interchangeably. The following mutations are described in the *Mb* nomenclature, which we commonly use in the lab. N311, also known as gatekeeper, is important for specificity of the wt PylRS by recognition of lysine derivatives via a hydrogen bond to the N^{ε}-carbonyl group of lysine.[338] Mutation to smaller amino acids were found to increase acceptance for aromatic amino acids,[339] however this leads to a decrease in specificity of the PylRS, which can be observed via an increase of phenylalanine misincorporation.[338b] Mutation of residue A264 to threonine compensates for the loss of the N311 hydrogen bond by formation of a hydrogen bond to the α-carboxyl moiety.[338c, 340] The selected mutations (Figure 2.6b) to change the active site for the accommodation of aromatic tetrazine analogs of phenylalanine were introduced by quickchange mutagenesis on the wt PylRS to generate the engineered mutants.



b)

	Mutations
SM1	N311A C313A
SM2	A267T N311A C313A
SM9	A267T L274A N311A C313A
SM10	A267T N311A C313A W382T
SM27	L270I Y271F L274G C313F Y349F

Figure 2.6: Structure-guided design of new PylRS mutants for tetrazine amino acids.
a) Crystal structure of PylRS wt active site (3VQY) in complex with BocK (yellow). The sidechain of BocK points into the binding pocket, which is spanned by amino acids L270, Y271, L274, N311, C313, Y349 (not shown) and W382 (*Mb* nomenclature) (b) Mutations chosen to enlarge the binding pocket for acceptance of tetrazine amino acids TetF and TetA.

Expression tests with newly engineered *Mb* synthetases were then performed as described before, but did also not result in successfull incorporation of tetrazine amino acids TetF and TetA (Figure 2.7). No full-length GFP production was detected for the new PylRS mutants having a N311A mutation, neither in the uAA samples nor the cultures without uAA. This suggests that the N311A mutation does not give rise to a loss in specificity here.

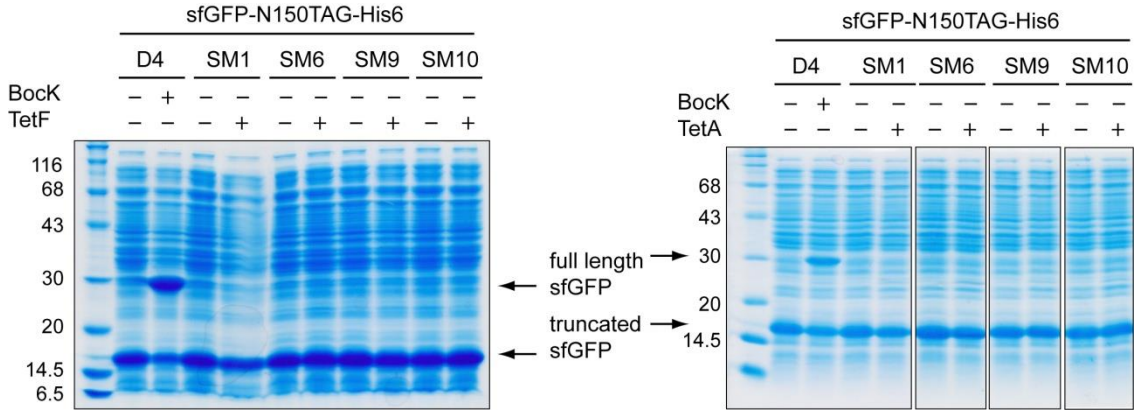


Figure 2.7: Coomassie stained SDS-PAGE of test expressions of sfGFP-N150TAG-His6 using newly generated PylRS mutants with 2 mM a) TetF b) TetA.

In the meantime, in the beginning of 2015, Mehl and coworkers published the development of an engineered tRNA synthetase/tRNA pair from *M. jannaschii* for the incorporation of TetF into proteins in *E. coli* via a directed evolution approach.[151b] This synthetase will be used for further experiments with uAA TetF, while it unfortunately does not accept TetA as a substrate (data not shown), which made new strategies to find a synthetase for TetA necessary.

2.3.2.2 Finding new synthetases via a directed evolution approach for TetA

Directed evolution of proteins is a powerful tool to find new enzymatic activities or to generate new substrate specificities. Therefore, this approach was chosen to find new PylRS mutants for the incorporation of TetA into proteins. The library AB2 selected for this task was created by a master student in the lab and contains the following randomization sites[341]:

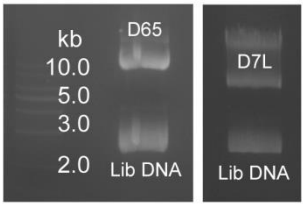
Library	Randomization sites
AB2	A267T (fixed) L270 Y271 N311 C313 Y349

The library comprises five randomized positions in the active site, resulting in a size of 3.3×10^7 members, which can be taken up by *E. coli* cells. The fixed A267T mutation enhances specificity for aromatic amino acids, when N311 is mutated, as discussed in 2.3.2.1.

The directed evolution approach covered one round of positive and negative selection each, followed by a fluorescent read-out with sfGFP. The positive selection plasmid D65 harbors a constitutively expressed chloramphenicol acetyltransferase (CAT) with one premature TAG codon, which leads to the survival of clones on chloramphenicol (CAM) containing positive selection plates, when the amber codon in CAT is effectively suppressed with the desired uAA or any natural amino acid. The initial transformation of the library AB2 into DH10B cells containing already the D65 plasmid resulted in 37 clones on the 10^{-7} plate of the dilution experiment. This denotes a 100-fold coverage of the library and was sufficient for the selection process. Positive selection with TetA yielded a distributed layer of rather small single colonies. Purified PylRS DNA (Figure 2.8a) from this step was subjected to a round of negative selection using L-arabinose induced barnase expression, containing two premature TAG codons (D7L plasmid) on LB agar without addition of uAA to remove false positives from the first positive selection round. Surviving members were collected, their DNA purified (Figure 2.8a) and then co-transformed with the read-out plasmid pPyltsfGFP-N150TAG CAT, which encodes sfGFP with a premature TAG codon for the fluorescence read-out and the orthogonal tRNA PylT. Additionally, it harbors one copy of the constitutively expressed chloramphenicol acetyltransferase with one premature TAG codon, which induces a second positive selection step in the read-out procedure upon addition of chloramphenicol.

Library AB2, TetA

a)



b) sfGFP fluorescence read out

	1	2	3	4	5	6	7	8	9	10	11	12
A	20,0	37,9	7,69	3,41	6,98	15,3	0,72	0,76	0,65	1,13	0,99	1,34
B	0,51	1,18	1,37	1,33	0,46	0,79	0,75	0,86	0,72	0,41	1,41	1,18
C	1,09	1,00	1,63	1,04	0,98	0,98	0,83	1,03	0,71	1,35	1,27	0,85
D	0,89	0,97	1,25	1,07	1,19	1,03	0,83	1,92	0,52	1,29	1,16	0,70
E	1,13	1,10	1,15	0,97	0,84	1,01	0,74	1,15	0,54	0,72	0,88	1,04
F	0,96	0,90	1,16	0,98	1,10	0,69	0,84	0,50	0,69	0,85	0,98	1,08
G	0,95	1,01	1,01	0,72	0,74	0,77	0,62	0,72	1,00	1,01	1,20	1,16
H	0,91	1,01	0,93	0,92	0,68	0,76	0,89	0,50	1,10	0,90	1,26	1,04

Figure 2.8: Directed evolution for TetA with library AB2.
a) Preparative agarose gel electrophoresis for the purification of library DNA after positive and negative selection. b) Fluorescence values of cultures with uAA normalized to the OD600 and fluorescence in absence of uAA. Green shaded values denote clones for the positive control with D4 and BocK.

The D4 plasmid encoding the wt PylRS was used as a positive control for the read-out together with BocK. From the autoinduction plate 90 slightly green colonies and 6 positive control clones were picked and grown in the absence and presence of TetA or BocK (positive control) respectively. Then, the OD_{600} and the fluorescence of the cultures were determined. The fluorescence of the cultures containing uAA was normalized to the OD_{600} and then divided by the normalized fluorescence values for cultures in the absence of uAA, which are displayed in Figure 2.8b. Only the positive control cultures show values from 3-40, while the other clones from the TetA read-out plate displayed values around or lower than one, indicating that TetA was not accepted and incorporated by any of the picked clones.

To find a synthetase recognizing TetA as a substrate further directed evolutions with other PylRS or even *Mj* TyrRS libraries would be necessary. Since a *Mj* TyrRS mutant was reported for the incorporation of TetF into proteins and TetF was assumed to have a higher reactivity, due to the stronger electron withdrawing properties of the phenyl ring compared to the alkyl substituent in TetA, approaches for finding synthetases for TetA were discontinued at this point.

2.3.2.3 Incorporation of uAAs TetF and CpK into sfGFP and characterization of the iEDDAC reaction with spin labels on protein

2.3.2.3.1 Expression of sfGFP-N150TAG with TetF and IpTetF using a *Mj* TyrRS mutant and CpK using the wt PylRS/tRNA pair

The Mehl lab kindly provided us with the plasmids for the incorporation of TetF by the *M. jannaschii* tyrosyl synthetase Tet-v.2.0RS. In this system, the genes for the tRNA synthetase and the *Mj* $tRNA_{CUA}$ are both encoded on one plasmid (pDULE), which is then co-transformed into *E. coli* DH10B cells with a pBAD plasmid harboring the gene for His-tagged sfGFP-N150TAG under an arabinose-inducible promoter. To test the incorporation with TetF and its derivative IpTetF, a test expression was performed (Figure 2.9). Since no uAA was available as a positive control for incorporation sfGFP wt was used as a reference for the production of full-length protein (approx. 27 kDa).The omission of UAA served as a negative control, as only truncated sfGFP (MW ~ 14 kDA) was produced in those samples.

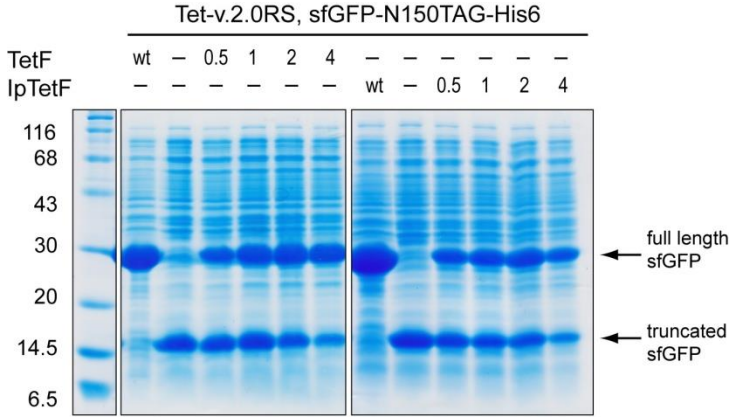


Figure 2.9: Coomassie stained SDS-PAGE of the test expressions of sfGFP-N150TAG-His6 with increasing concentrations of TetF and IpTF [mM]

Increasing concentrations of the uAAs were tested for the incorporation, which lead to the increase of full-length and the decrease of truncated GFP in a dose-dependent manner, respectively. Due to the slight preference of Tet-v2.0RS for TetF over IpTF (Figure 2.9), TetF was used for further experiments.

To confirm the incorporation of TetF into sfGFP by MS analysis, a 200 mL culture was expressed with 2 mM TetF (Figure 2.10a), since this was a good compromise between the amount of used amino acid and protein production (Figure 2.9 left side). Protein was purified using standard condition for Ni-NTA affinity purification, which yielded 70 mg/L of culture (Figure 2.10b). The purified sfGFP-N150TetF-His6 was observed to appear yellow instead of green, due to quenching of sfGFP fluorescence by the absorption of tetrazines around 520 nm. Mehl and coworkers also witnessed this deviation in appearance.[151b] Successful incorporation of TetF was confirmed by ESI-MS (Figure 2.10c).

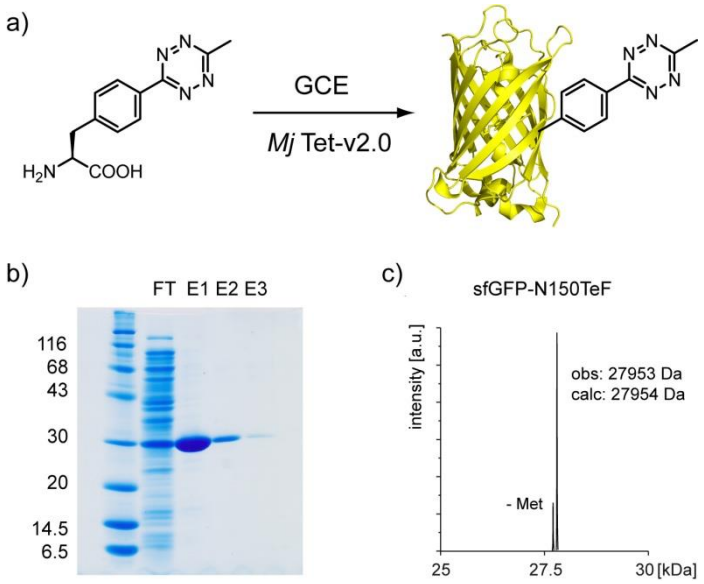


Figure 2.10: Characterization of the incorporation of TetF into sfGFP-N150TAG-His6 with Tet-v2.0
a) Schematic illustration of the incorporation. b) Coomassie stained SDS-Page of sfGFP-N150TetF-His6 purification with Ni-NTA beads (FT = flow through, E = elution fraction) c) ESI-MS characterization of sfGFP-N150TetF-His6.

To validate our general approach that a smaller, more rigid spin label would provide better data in EPR and PRE experiments, we chose the lysine-based uAA CpK as comparison. Due to its flexible lysine linker the resulting spin label displays more degrees of rotational freedom, which could negatively influence the PRE or PELDOR signal. After successful synthesis based on published reports[3] the incorporation of CpK into sfGFP was tested (Figure 2.11a), using the D4 plasmid encoding for the wt PylRS/tRNA pair from *M. barkeri*. Expression and purification of sfGFP-150CpK-His6 were performed using standard protocols (Figure 2.11b). The incorporation of CpK was confirmed by ESI-MS analysis (Figure 2.11c).

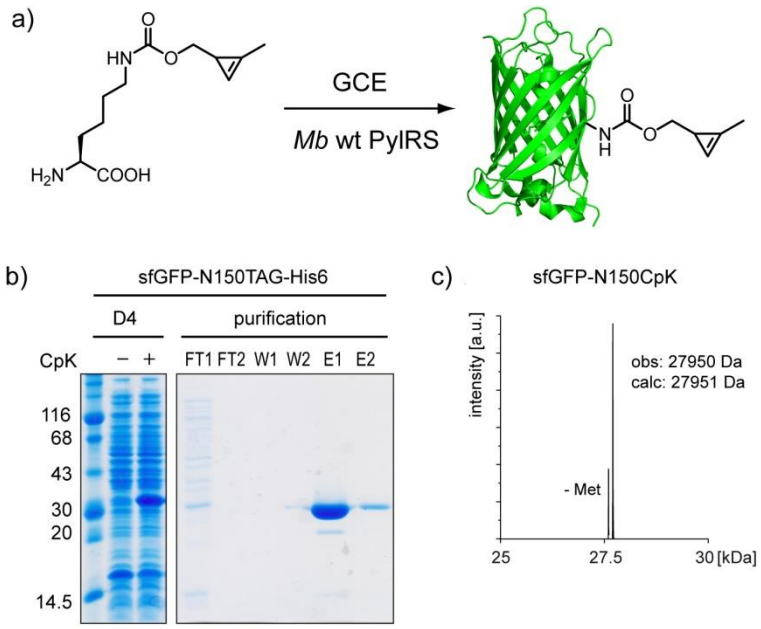


Figure 2.11: Characterization of the incorporation of CpK into sfGFP-N150TAG-His6 with wt PylRS
a) Schematic illustration of the incorporation of CpK into sfGFP-N150TAG-His6 b) Coomassie stained SDS-PAGE of expression (left) and purificaton (right) of sfGFP-N150CpK-His6 c) ESI-MS characterization of sfGFP-N150CpK-His6.

2.3.2.3.2 Characterization of iEDDAC reaction between TetF and different dienophiles

To characterize the iEDDAC reaction for SDSL, labeling experiments between Boc-protected TetF (1 mM) were conducted with various strained alkenes (norbornenol (Nor-OH) and CpK) and alkynes (BCN-OH) in a 1 to 5 ratio and analyzed after certain time points (0 min, 10 min, 1 hour, 4 hours and 24 hours) by HPLC (Figure 2.12a).

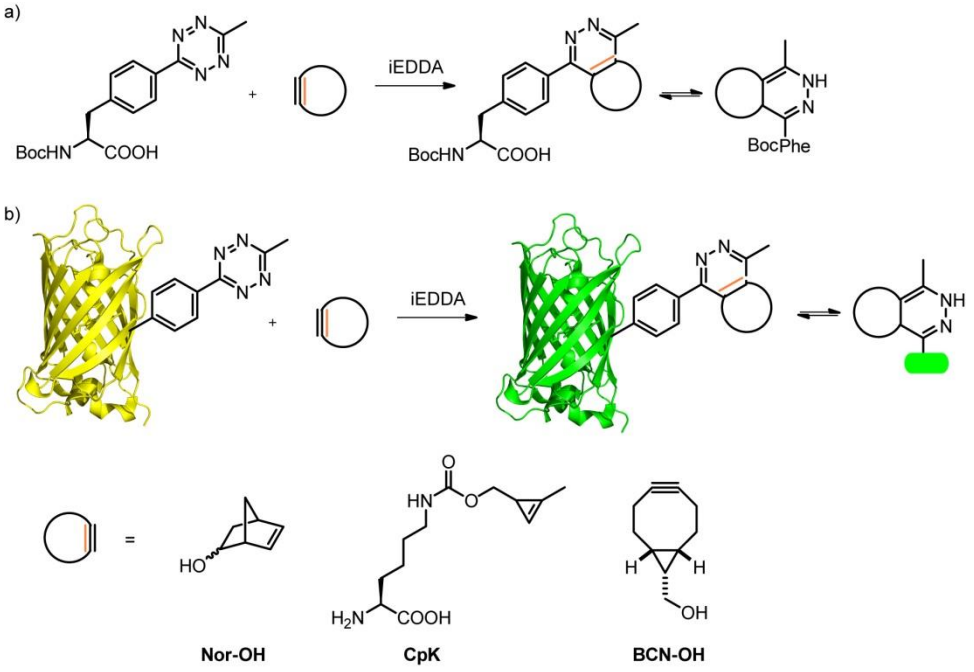


Figure 2.12: iEDDAC reaction between TetF and different strained dienophiles Nor-OH, CpK and BCN-OH
on a) small molecule and b) protein level. Formation of different dihydopyridazine tautomers is only possible in case of a strained alkene (black); strained alkynes (orange) form aromatic pyridazines (orange).

Nor-OH, which has the smallest ring strain, shows the slowest reaction with BocTetF; product formation seems to be complete after 4 hours (Figure 2.13a). CpK displays faster reaction kinetics and the reaction looks complete after only 1 hour (Figure 2.13b). The fastest iEDDAC reaction with BocTetF was observed with BCN-OH, where already at the t = 0 min sample product formation is complete (Figure 2.13c).

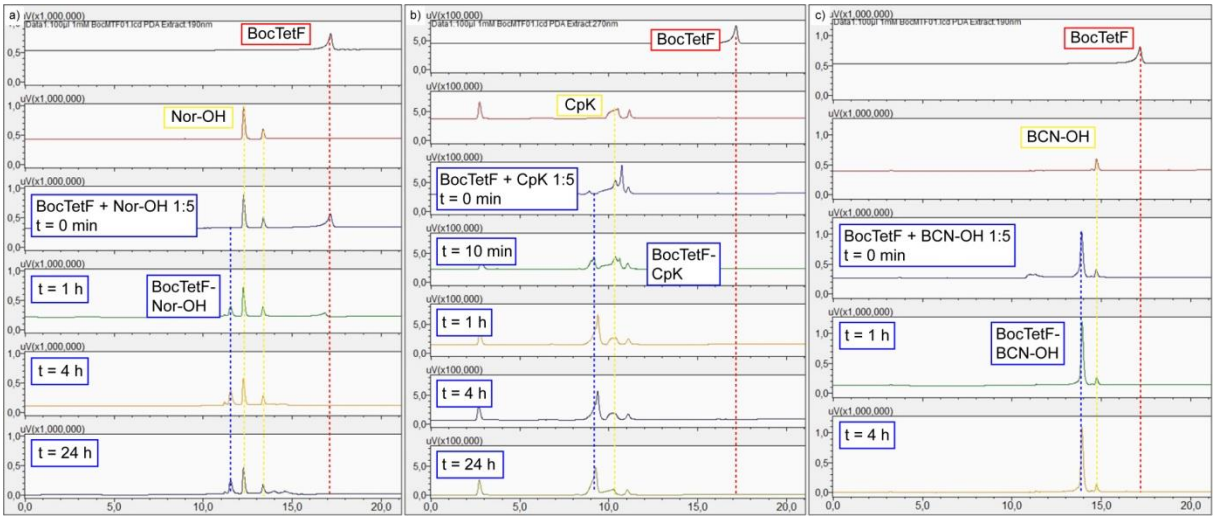


Figure 2.13: Characterization of the iEDDAC reaction by HPLC between BocTetF (1 mM) and 5 mM a) Nor-OH b) CpK and c) BCN-OH.

After characterization of the iEDDAC reaction with TetF on small molecule, the reaction was taken to protein level to verify that labeling there proceeds accordingly. Therefore, purified sfGFP-N150TetF-His6 (10 µM) was subjected to *in vitro* labeling experiments with 20 equivalents of the aforementioned strained alkenes and alkynes Nor-OH, CpK and BCN-OH at 25 °C (Figure 2.12b). Complete conversion to the labeled protein was observed for BCN-OH (Figure 2.14c) after only 15 min, while CpK labeling (Figure 2.14b) was complete after 2 hours. Labeling with Nor-OH resulted in a 40:60 mixture of unlabeled and labeled protein after 24 hours (Figure 2.14a). Reaction of sfGFP-N150TetF-His6 with dienophiles lead to restoration of fluorescence, due to loss of absorption of the tetrazine at 550 nm, which could be observed in a shift of color from yellow to green (cartoon representation in Figure 2.12b).

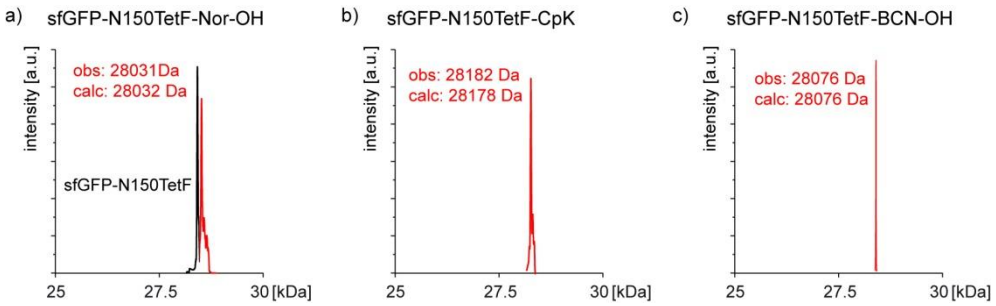


Figure 2.14: ESI-MS spectra of sfGFP-N150TetF-His6 labeling with a) Nor-OH, b) CpK and c) BCN-OH.

These labeling experiments with different strained alkenes and alkynes verifies our chosen approach to use the iEDDAC reaction between tetrazines and 3-methylcyclopropene derivatives for spin labeling. The Cp moiety combines its small size, which results in small

rigid linkages, with still suitable reactivity towards tetrazines, while other dienophiles are bigger in size (BCN-OH, Nor-OH) and/or display slower reaction rates (Nor-OH).

2.3.2.3.3 Characterization of SDSL on sfGFP-N150TetF-His6

After establishing that our approach using Cp moieties as reactive groups for bioorthogonal labeling showed suitable reactivity in iEDDAC, the next step was to test if the nitroxide moieties attached to the Cp ring would interfere with the labeling reaction on the protein. Therefore, *in vitro* labeling experiments (Figure 2.15) with 20 equivalents excess of spin labels Cp-PROXYL and Cp-TEMPO were performed with purified sfGFP-N150TetF-His6 (10 µM) at 25 °C. Samples were taken at certain time points and analyzed via LC-ESI-MS. The reaction with Cp-PROXYL showed complete labeling of sfGFP-N150TetF-His6 (Figure 2.15b) after 2 hours, while for Cp-TEMPO at this time point still 25 % unlabeled protein was observed. The difference in reactivity can be explained by the linkage of the nitroxide spin label to the Cp moiety. The TEMPO group is attached via a carbamate linkage, while Cp-PROXYL contains an amide bond to the Cp ring, which is a better electron donor than the carbamate, resulting in higher reactivity in iEDDAC reactions. Further experiments were therefore performed with Cp-PROXYL in a 20-fold excess compared to the protein.

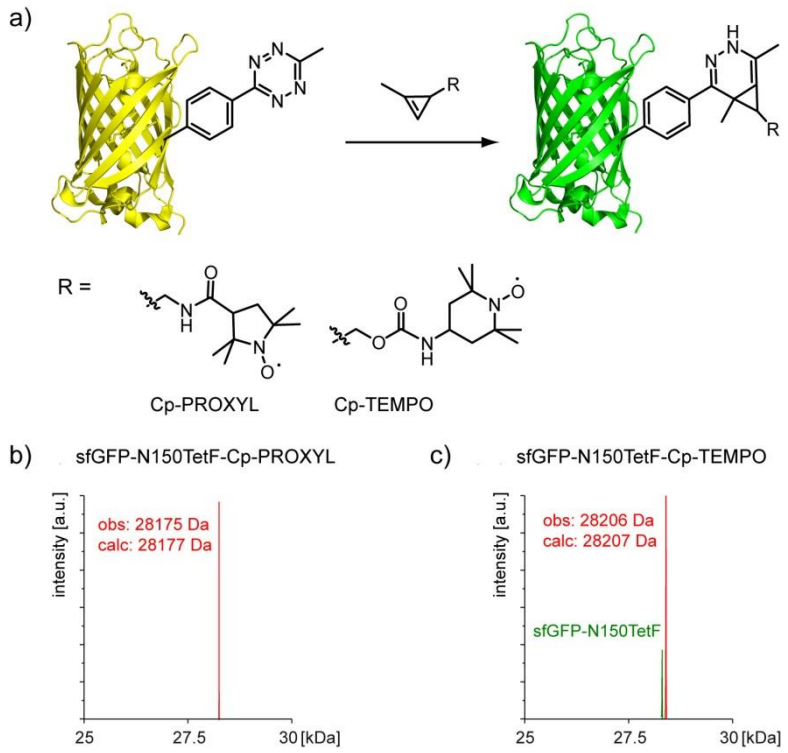


Figure 2.15: Labeling of sfGFP-N150TetF-His6 with Cp based spin labels Cp-PROXYL and Cp-TEMPO
a) Schematic illustration of the labeling reaction b) ESI-MS spectrum of labeling with Cp-PROXYL after 2h shows completely labeled protein. c) ESI-MS spectrum of labeling with Cp-TEMPO still shows 25 % unlabeled protein after 2 h.

To test if the tetrazine based spin labels also display differences in reactivity, *in vitro* labeling experiments (Figure 2.16) with purified sfGFP-N150CpK-His6 (10 µM) were performed. Tet-PROXYL and Tet-TEMPO were adds in a 20-fold excess and the resulting reaction mixtures analyzed at certain time points by LC-ESI-MS. Labeling of sfGFP-N150CpK-His6 with Tet-PROXYL was complete after 2 hours (Figure 2.16a), while Tet-TEMPO showed a ration of

40:60 of unlabeled to labeled protein at this point (Figure 2.16b), which was contrary to our expectations. In this case Tet-PROXYL was used for further experiments in a 20-fold excess compared to the protein.

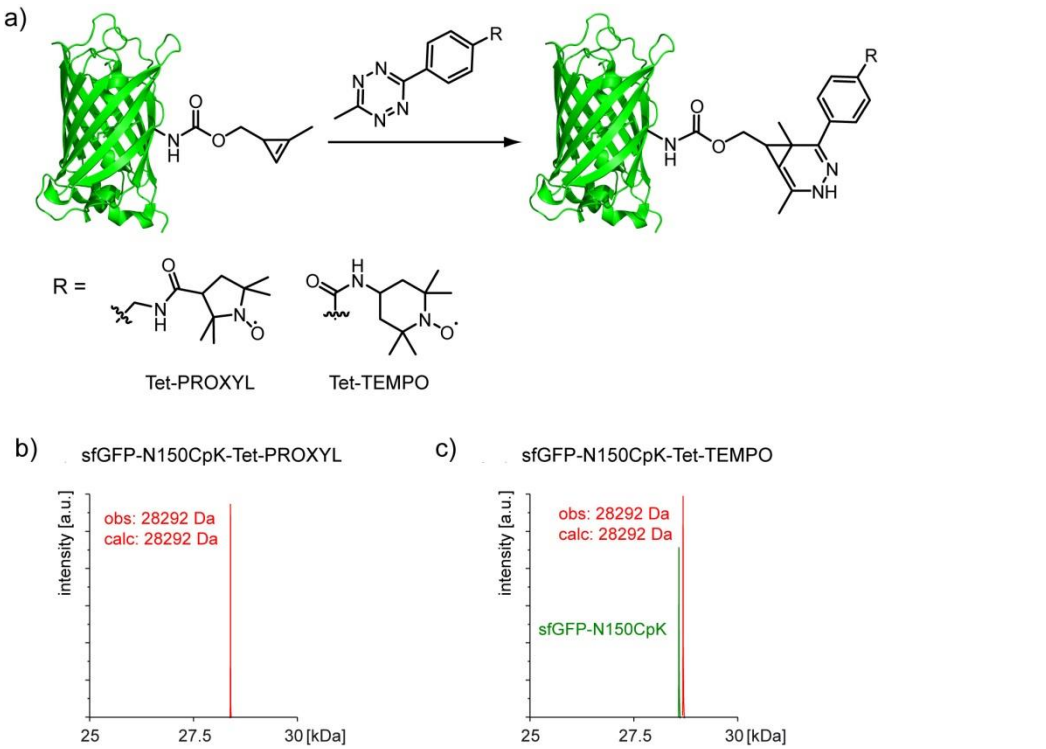
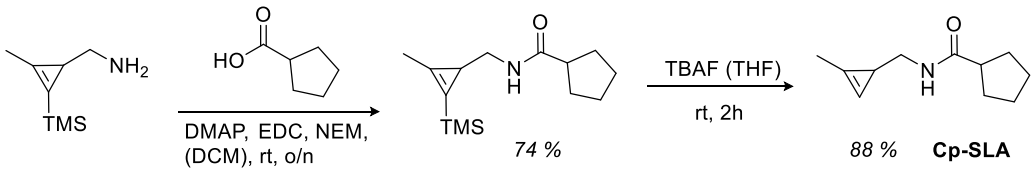


Figure 2.16: Labeling of sfGFP-N150CpK-His6 with tetrazine-based spin labels Tet-PROXYL and Tet-TEMPO.
a) Cartoon representation of the labeling reaction. b) ESI-MS spectrum of labeling with Tet-PROXYL after 2h. c) ESI-MS spectrum of labeling with Tet-TEMPO: only 60 % labeling was observed after 2h.

Our short labeling times for complete modification of the proteins with incorporated uAAs TetF and CpK with the corresponding spin labels showed that the nitroxide moieties do not interfere with the iEDDAC reaction on protein, confirming that our SDSL approach is suitable for the introduction of nitroxide spin labels into proteins.

2.3.2.3.4 Determination of on-protein reaction kinetics

After performing *in vitro* labeling experiments and analyzing samples at different time points, on-protein rate constants were determined in a more quantitative way. To achieve this, the fluorescence quenching properties of TetF in proximity to the sfGFP chromophore were exploited. As experimentally observed, iEDDAC reaction with dienophiles leads to restoration of sfGFP fluorescence. Following this exponential increase over time allows the determination of second order on-protein rate constants. As an active nitroxide spin label is not necessary for this step, a spin label analog (Cp-SLA, Scheme 2.14) bearing a cyclopentane ring instead of the PROXYL moiety was synthesized according to the procedure described for Cp-PROXYL (chapter 1.3.1.4.1).



Scheme 2.14: Synthesis of Cp-SLA

Rate constant k (Figure 2.17) between sfGFP-N150TetF-His6 (10 µM) and Cp-SLA was determined under pseudo first order conditions with an excess of Cp-SLA (50, 75 or 100-fold) in water measuring GFP fluorescence. Data processing resulted a rate constant of ~0.1 $M^{-1}s^{-1}$, which is slower compared to the reported reactions with sTCO or dTCO[151b], however these are the fastest dienophiles in iEDDAC reactions known to date. Kinetic measurements performed in our lab between mTetK and Cp-SLA (~1 $M^{-1}s^{-1}$) or Nor-OH (~0.06 $M^{-1}s^{-1}$) revealed a almost 20-fold faster reaction rate for Cp over Nor-OH (see chapter 1, Figure 1.17). The lower reactivity of TetF might be due to the higher rigidity and smaller size of TetF compared to more flexible lysine derivative mTetK, where the side chain is further away from the peptide backbone, resulting in a lower accessibility of TetF on the protein surface.

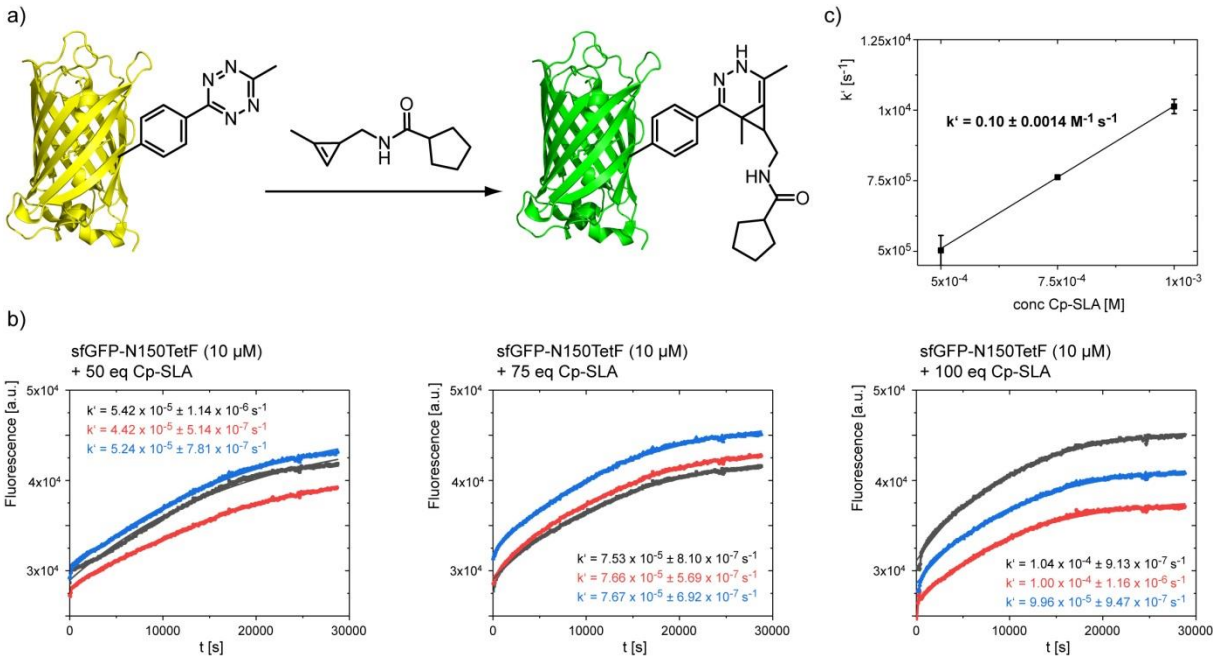


Figure 2.17: Determination of second-order rate constant k between sfGFP-N150TetF-His6 and Cp-SLA under pseudo-first order conditions at 25 °C.

a) Cartoon representation of the labeling reaction between sfGFP-N150TetF-His6 and Cp-SLA. b) sfGFP-N150TetF-His6 was individually treated with solutions of Cp-SLA to obtain final concentrations of 10 µM and 500 to 1000 µM respectively. GFP fluorescence was excited at 480 nm and followed at 527 nm over time. The k' values, determined by fitting a single exponential equation, were plotted against the concentration of Cp-SLA. Experiments were performed in biological triplicates. c) By fitting a linear equation, the rate constant k was obtained from the slope of the plot. All data processing was performed with OriginPro software (OriginLab Corporation, Northampton, USA).

2.3.2.3.5 Incorporation of TetF into sfGFP double mutant for subsequent spin labeling

After showing that our approach yields quantitatively spin labeled protein after short reaction times, characterization of the spin label was planned by determining the intra-protein distance between two spin labels placed in one protein using PELDOR measurements. Our first choice was sfGFP as model protein, which was used to establish SDSL conditions using genetic code expansion. To enable PELDOR measurements, a second amber stop codon needed to be introduced site-specifically into sfGFP-N150TAG-His6 to achieve double spin labeling. Position R109 was chosen for the introduction of a second TAG codon by quickchange mutagenesis creating the sfGFP-R109N150TAG-His6 mutant in the pBAD vector. Expression was performed with Tet-v2.0RS in the presence of 4 mM TetF instead of 2 mM, resulting in 20 mg/L of culture of purified sfGFP-R109N150TetF-His6 (Figure 2.18b). ESI-MS confirmed the successful incorporation of two molecules of TetF into sfGFP (Figure 2.18c). Subsequent labeling with a 40-fold excess of Cp-PROXYL yielded the quantitatively double-labeled protein (Figure 2.18a,c) and was therefore used as standard ratio (1:40) for double spin labeling.

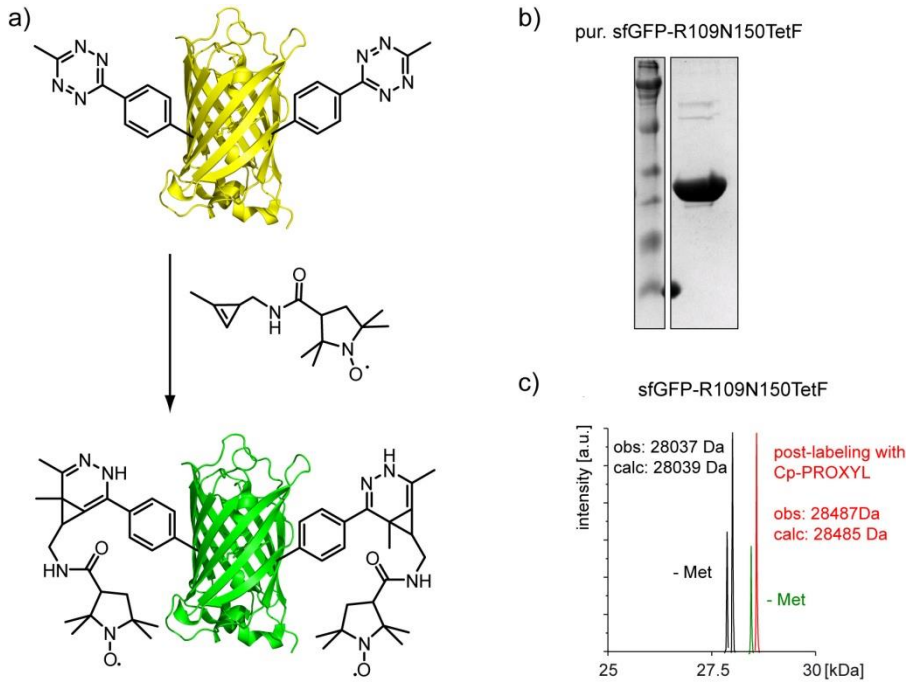


Figure 2.18: Incorporation of TetF into of sfGFP-R109N150TAG-His6 double mutant and subsequent labeling with Cp-PROXYL

a) Schematic illustration of the labeling reaction b) Coomassie stained SDS-PAGE of purified sfGFP-R109N150TetF-His6 c) ESI-MS spectrum of purified sfGFP-R109N150TetF-His6 before and after labeling with Cp-PROXYL, which shows completely labeled protein after 2 h.

PELDOR measurements to determine the distance between the spin labels at positions 109 and 150 in sfGFP were performed by our collaboration partners Stefan Gaussmann (Sattler group, Technical University Munich) and Burkhard Endeward (Prisner group, Goethe University Frankfurt am Main). Unfortunately, the sfGFP sample showed spin clustering, meaning the presence of more than the two expected spins, possibly due to the formation of sfGFP dimers, which is known to happen when using GFP at high protein concentrations.[342] Facing these problems, ubiquitin was chosen as an alternative model protein to characterize and benchmark our SDSL approach against traditional MTSSL labeling.

2.3.2.4 Ubiquitin as model protein for the characterization of our new iEDDAC SDSL approach

2.3.2.4.1 Expression and spin labeling of Ub mutants

As a first step, an expression system for the incorporation of TetF into ubiquitin was constructed. The synthetase Tet-v2.0RS was combined with a pET17 vector encoding a *C*-terminally His-tagged ubiquitin with a slightly modified *C*-terminus. Single and double TAG mutants at positions T22 and E34 were created by quickchange mutagenesis for PRE experiments or PELDOR measurements, respectively. The two TAG positions were chosen to be at the opposite ends of the α-helix of ubiquitin for pre-orientation of the spin labels, since we wanted to characterize the flexibility of the spin label and a disordered, flexible protein region would not allow to gather these data (Figure 2.19a). The expression of the three ubiquitin mutants Ub-T22TAG-His6, Ub-E34TAG-His6 and Ub-T22E34TAG-His6 was performed in *E. coli* BL21 (DE3) cells using an N5052 autoinduction medium with normal isotope abundance in the presence of 2 or 4 mM TetF for single or double mutants, respectively (Figure 2.19b).[343] Details of the expression optimization for amber suppression can be found in Stefan Gaussmann's Master's Thesis.[344] Standard purification via the C-terminal His-tag, followed by size exclusion chromatography (SEC) and subsequent ESI-MS analysis confirmed successful incorporation of TetF into the single or the double mutants of ubiquitin (Figure 2.19c). Labeling with excess of Cp-PROXYL overnight at 4 °C lead to complete modification of the ubiquitin mutants (Figure 2.19c).

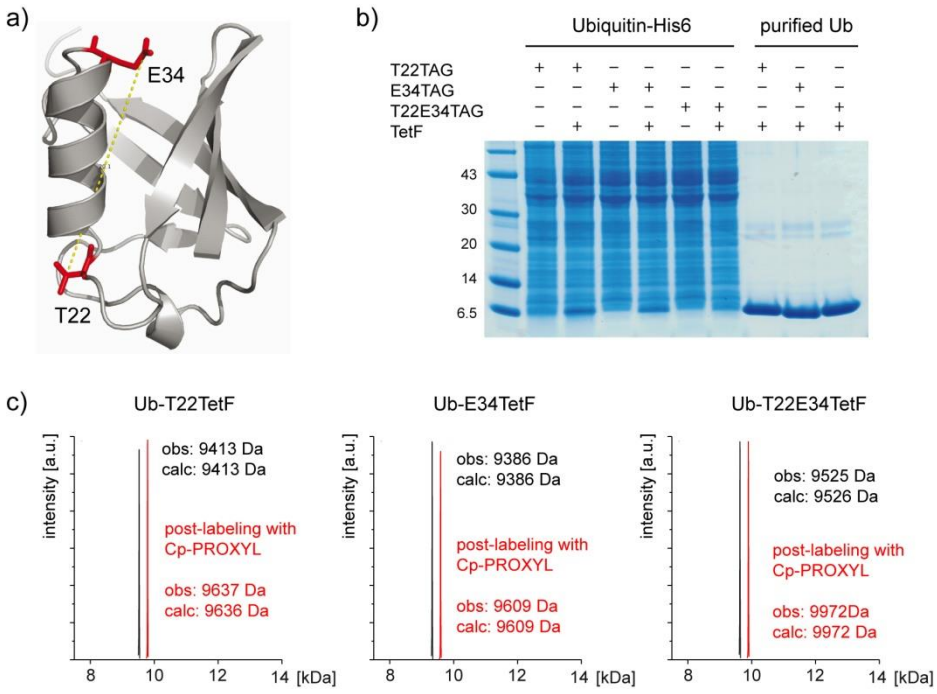


Figure 2.19: Ubiquitin (Ub) as model protein for SDSL.
a) Schematic illustration of the positions for SDSL in Ub b) Coomassie stained SDS-PAGE of the incorporation of TetF into Ub-T22TAG-His6, Ub-E34TAG-His6 (2 mM) and Ub-T22E34TAG-His6 (4 mM) and purified ubiquitin mutants c) ESI-MS characterization of purified ubiquitin mutants before and after labeling with 20 or 40 equivalents of Cp-PROXYL.

2.3.2.4.2 Performance of PELDOR measurements with Ub-T22E34TetF-His6

After successful expression and subsequent spin labeling of three ubiquitin mutants PELDOR measurements were performed by our collaboration partners from the Sattler and the Prisner group with the double mutant Ub-T22E34TetF-His6 (Figure 2.19a) to characterize our positioned spin labels.

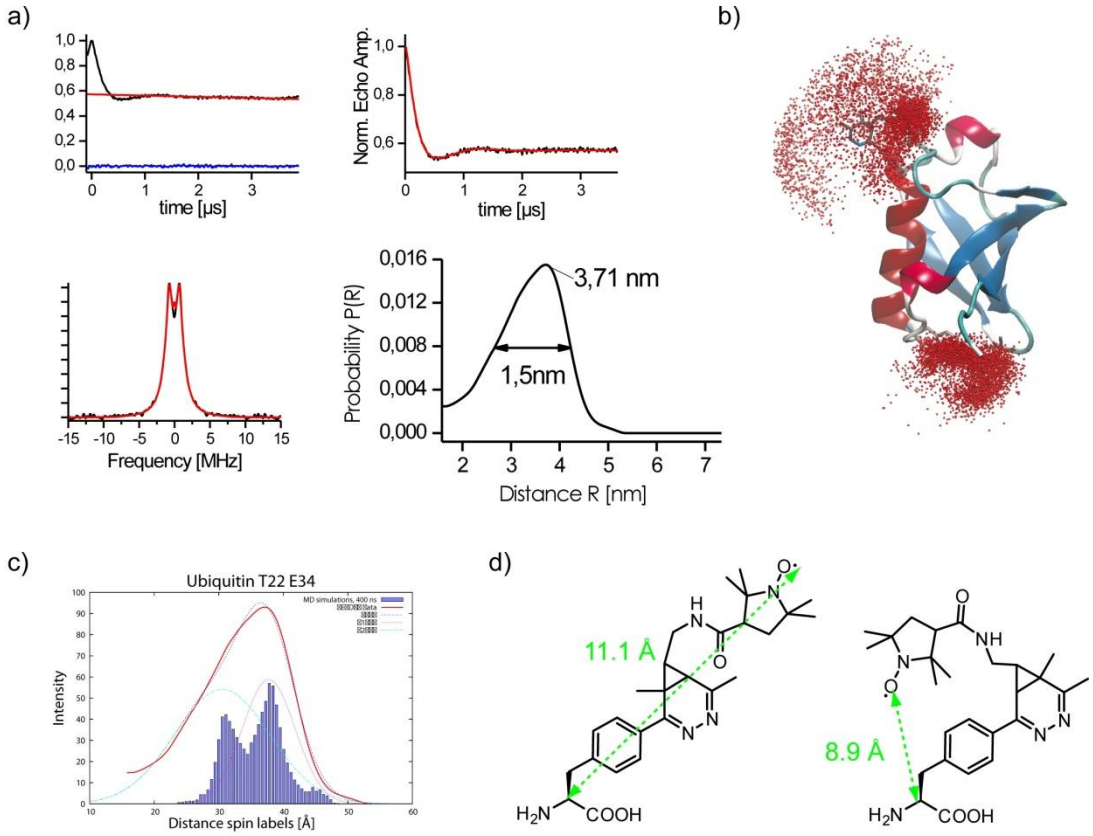


Figure 2.20: Characterization of SDSL approach on Ub-T22E34TetF-Cp-PROXYL.
a) PELDOR measurement. Top left: PELDOR time trace. Top right: PELDOR time trace normalized by background subtraction. Bottom left: Fourier transformed time trace. Bottom right: distance distribution between the spin labels in position T22 and E34 in Ub. b) MD simulation of Ub-T22E34TetF-Cp-PROXYL c) Overlay of experimental PELDOR data with MD simulation d) Possible stereoisomers after labeling reaction with schematic illustration of the distances between Cα atom and the radical.

The experiments determined an intra-protein distance distribution between the spin labels in positions T22 and E34 of 3.71 nm with a peak width of 1.5 nm (Figure 2.20a), which is broader than the distribution gathered from MTSSL labeling of Ub-T22E34C, which shows a narrower and shorter distribution of ~3 nm with a peak width of ~0.4 nm.[345] This might be due to two fact that the reaction between TetF and the 3-methylcyclopropene moiety can form two stereoisomers upon ligation, which differ in the position of the methyl group (4 or 5) on the generated dihydopyridazine (Figure 2.20d). After energy minimization with MM2 in Chem3D of these two stereoisomers, the minimal distances between the Cα atom and the radical constitute to 8.8 Å (11.7 Å without minimization) for one and 11.1 Å (13 Å without minimization) for the other isomer. A mixture of these two conformations might account for the broader distance distribution, since MTSSL labeled cysteine adopts only one conformation with a maximal distance of around 9 Å after energy minimization with MM2 in Chem3D (9.7 Å without minimization). MD simulations (Figure 2.20b), performed by our collaborators from the Kaila group, which include the information of the two stereoisomers fit

well to the experimental data, as shown in the overlay with the PELDOR distance distribution (Figure 2.20c).

2.3.2.4.3 Performance of PRE NMR measurements with Ub-T22TetF-His6 and Ub-E34TetF-His6

To characterize the paramagnetic relaxation enhancement of the spin label on surrounding nuclei ^{1}H T1 and T2 relaxation experiments were performed on the spin labeled Ub-T22TetF-His6 and Ub-E34TetF-His6 mutants, which were therefore expressed in ^{15}N and ^{13}C labeled M9 minimal medium. Purification and labeling was carried out as described before. ^{1}H relaxation experiments were measured with the active spin label first, which was then reduced to hydroxylamine by addition of five times excess of ascorbic acid in NMR buffer with adjusted pH. The second data set was measured with the reduced spin label. Data analysis is currently under way by our collaboration partners from the Sattler group.

2.3.2.5 Application of SDSL approach on Loqs-PD

After characterization of our SDSL approach based on GCE in combination with iEDDAC between tetrazines and 3-methylcyclopropenen moieties on the model system ubiquitin, validation of the method was planned on the protein Loqs-PD, which plays a role in the down-stream processing of siRNA. The structure of its two dsRBDs without the linker was published by our collaboration partners in the Sattler group[314], but domain orientation and intra-protein distances remained to be elucidated (Figure 2.21). Loqs-PD contains several cysteines in its sequences rendering MTSSL unsuitable; therefore our alternative spin labeling strategy comes into play.

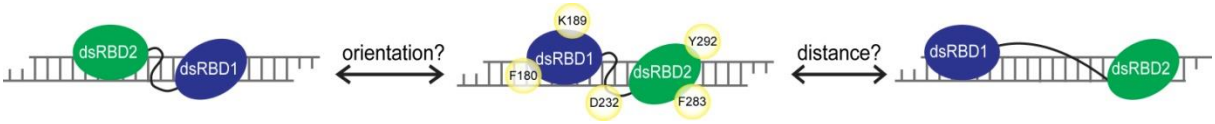


Figure 2.21: Cartoon illustration of Loqs-PD

The question of domain orientation will be addressed by performing PRE experiments, while the distances between the two dsRBDs will be elucidated using PELDOR measurements. For PRE experiments five different positions (F180, K189, D232, F283 and Y292) in the sequence of Loqs-PD were chosen to introduce amber codons, while for PELDOR measurements a Loqs double mutant (F180Y292) was planned. PRE experiments and data analysis was carried out by our collaboration partner Stefan Gaussmann from the Sattler group and PELDOR measurements by Stefan Gaussmann together with Burkard Endeward from the Prisner group. Results in the following chapters will be showed in an exemplified version.

2.3.2.5.1 Expression of Loqs-PD$_{\Delta NC}$ mutants

Amber codons were introduced at five different positions (F180, K189, D232, F283 and Y292, Figure 2.22b) distributed over the two dsRBDs and the linker region by quickchange mutagenesis into the *N*- and *C*-terminally truncated Loqs-PD construct (Loqs-PD$_{\Delta NC}$) encoded in a modified pETM13 vector.[344] The construct comprises a SUMO tag for solubility of the protein during expression and a *C*-terminal His6-tag with a TEV cleavage site for purification (Figure 2.20a).

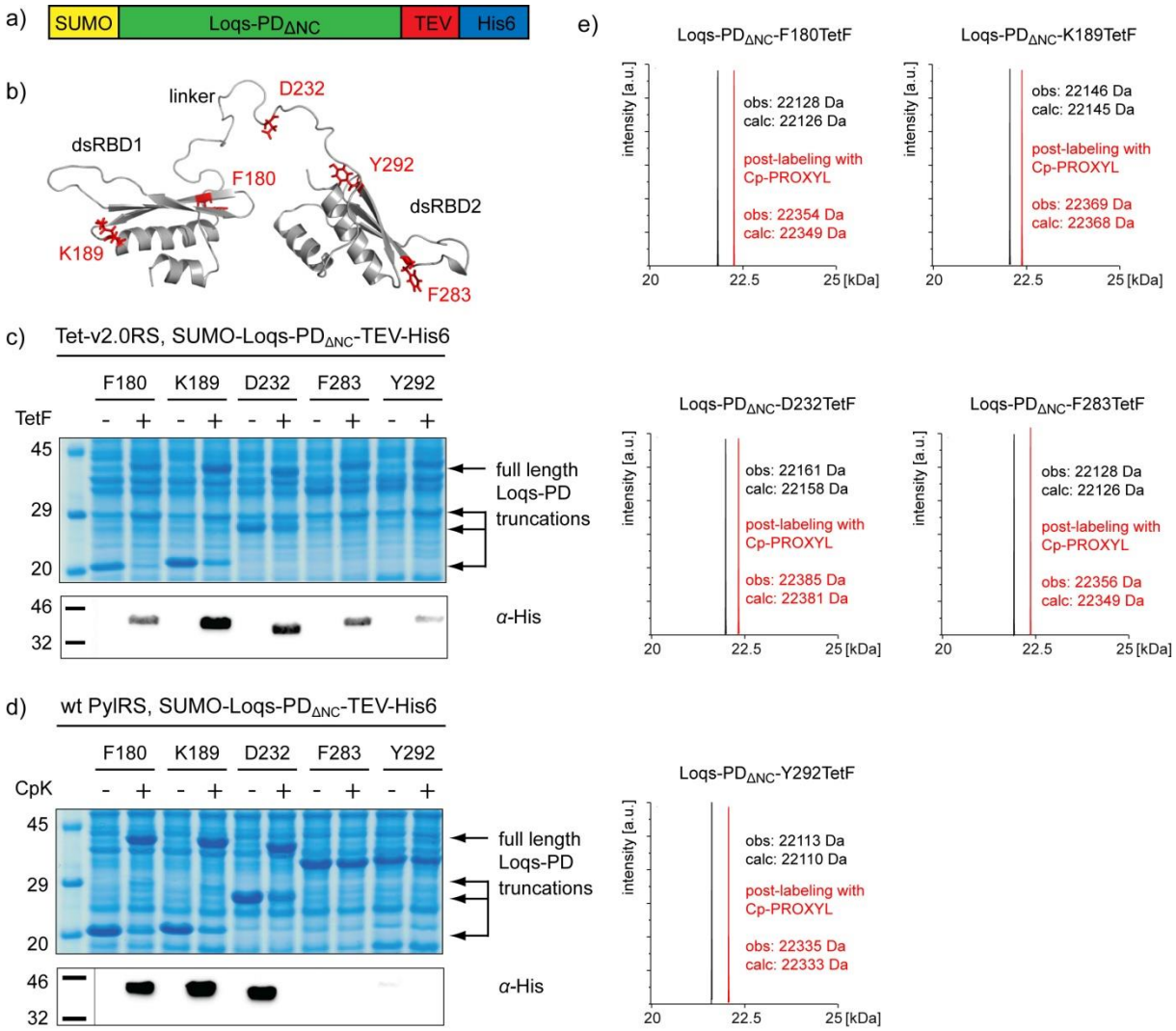


Figure 2.22: Incorporation of TetF and CpK into Loqs-PD$_{\Delta NC}$ TAG mutants.
a) Construct for expression b) Positions of TAG codons in Loqs-PD$_{\Delta NC}$ structure. c) Coomassie stained SDS-PAGE of expression of Loqs-PD$_{\Delta NC}$ mutants with TetF and d) CpK e) ESI-MS spectra of Loqs-PD$_{\Delta NC}$ mutants bearing TetF before and after labeling with Cp-PROXYL.

Expression of the different Loqs-PD$_{\Delta NC}$-TAG mutants was performed at 18 °C in BL21 (DE3) *E. coli* cells in N5052 autoinduction medium with 2 mM uAA in combination with Tet-v2.0RS for TetF incorporation or the wt PylRS for incorporation of CpK. All five Loqs-PD$_{\Delta NC}$ mutants showed good expression for the incorporation of TetF, which can be observed by SDS-PAGE or in the anti-His western blot (Figure 2.22c) and was confirmed by ESI-MS (Figure 2.22e). Incorporation of CpK showed production of full-length proteins for Loqs-PD$_{\Delta NC}$ F180TAG, K189TAG and D232TAG mutants, while mutants F283TAG and

Y292TAG resulted mostly truncated protein and full-length protein was not visible for these mutants by SDS-PAGE or anti-His western blot (Figure 2.22d), even thought purification did yield very low amounts of protein. This might be due to a disruption of the structure by replacing the aromatic residues in F283 and Y292 with the lysine derivative CpK at these positions, while TetF, being an aromatic uAA might here be a suitable substitute for the natural phenylalanine or tyrosine residues. Other positions do not seem to be prone to structural disruption by replacement of the natural amino acid with TetF or CpK, since the expression levels were not decreased (Figure 2.22d) in these cases, suggesting this has to be evaluated for each position and cannot generally be assumed. Replacement by TetF, however, seemed to be less problematic than with CpK. Details for expression and purification of the Loqs-PD$_{\Delta NC}$ mutants are described in the Master's Thesis of Stefan Gaussmann.[344] ESI-MS analysis confirmed incorporation of the uAAs into the different TAG mutants of Loqs-PD$_{\Delta NC}$, as well as complete modification upon addition of a 20-fold excess of the corresponding spin labels at 4 °C overnight (Figure 2.22e).

2.3.2.5.2 Determination of structural integrity of the Loqs-PD mutants

Since the incorporation of CpK into Loqs-PD$_{\Delta NC}$ mutants F283TAG and Y292TAG was problematic, chemical shift perturbations (CSP) were determined for the incorporation of TetF and CpK into each position. For this, ^{1}H,^{15}N-HSQC spectra of the TetF- and CpK-bearing mutants were measured and the position of the peaks compared to the spectrum of wt Loqs-PD, for example shown for Loqs-PD$_{\Delta NC}$ mutant F180 in Figure 2.23.

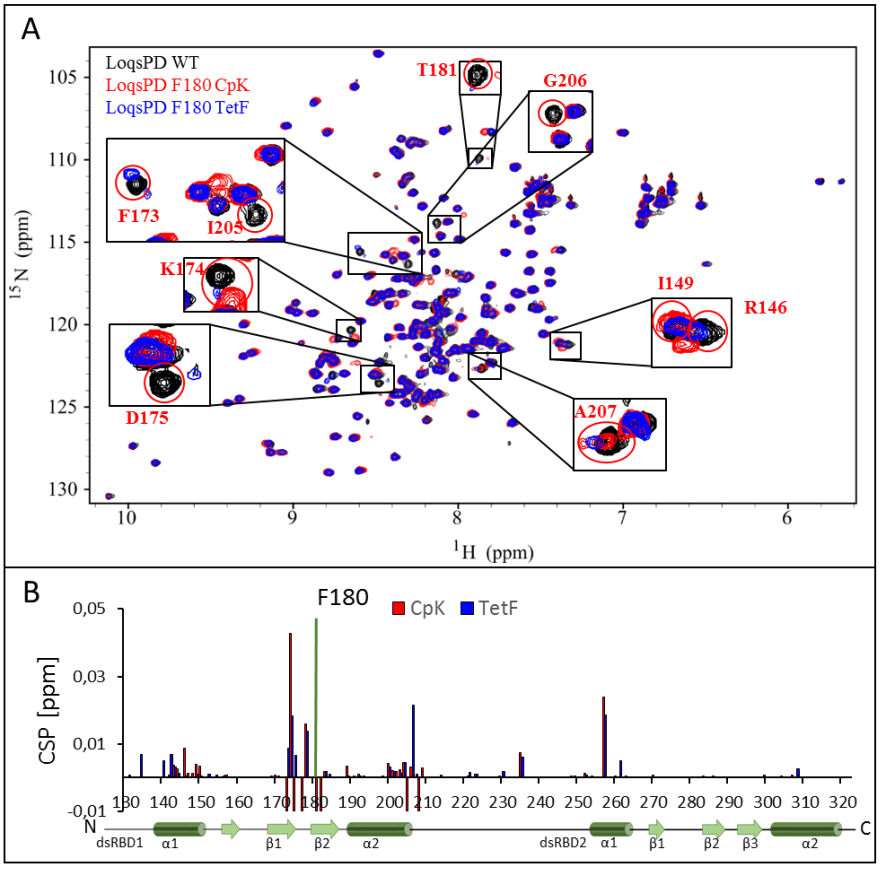


Figure 2.23: CSP of LoqsPD$_{\Delta NC}$ F180 mutants.
A: Overlay of WT (black), F180 CpK (red), and F180 TetF (blue) ^{1}H,^{15}N-HSQC spectra. B: CSPs plotted against the sequence.[344] - Reproduced by permission of Stefan Gaussmann

CSP were highest in structural proximity to the position of the incorporated uAA, in case of F180 at positions R146, I149, K174 and A207. CpK showed slightly higher perturbations than TetF, which might be due to the closer structural similarity of TetF and phenylalanine. In general, all mutants showed low CSPs suggesting the structure was not disturbed by incorporation of the uAA. The strongest perturbations were observed in the structural neighborhood of the uAA and a stronger shift, when the wt residue was substituted by a dissimilar uAA, for example K189 with TetF and F180 or F283 with CpK.[344]

2.3.2.5.3 PRE measurements of spin labeled Loqs-PD$_{\Delta NC}$ mutants

Since the CSP showed structural integrity of the Loqs-PD$_{\Delta NC}$ mutants with incorporated uAA TetF and CpK, PRE effects were determined on the free Loqs-PD$_{\Delta NC}$ mutants without RNA to evaluate the spin labels (Figure 2.24).

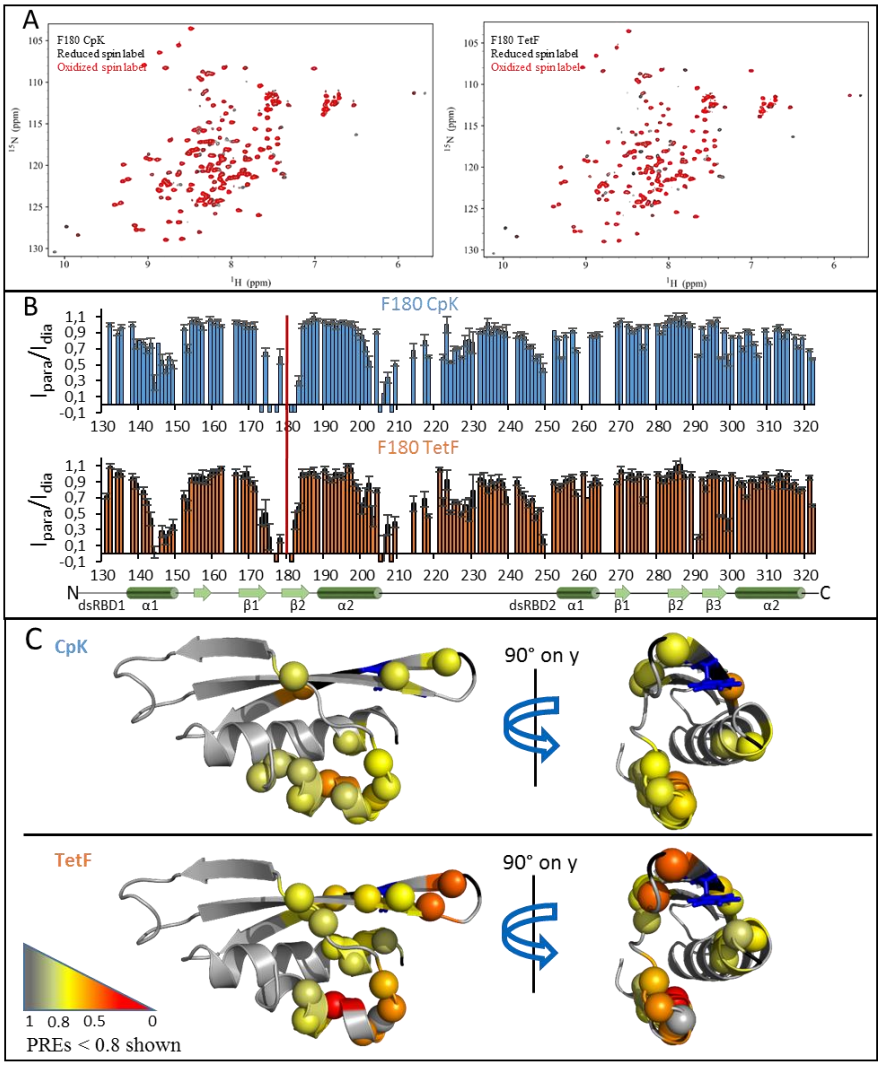


Figure 2.24: Analysis of PRE on Loqs-PD$_{\Delta NC}$ mutants F180TetF and F180CpK.
A: Overlay of ^{1}H,^{15}N-HSQC spectra of reduced spin label (black) and oxidized spin label (red). Spectra were recorded with 24 scans, 140 increments and a delay D1 of 3 seconds. PRE effects are detectable by comparison of peaks in the black (diamagnetic) and the red (paramagnetic) spectrum.[344] - Reproduced by permission of Stefan Gaussmann.

$^{1}H,^{15}N$-HSQC spectra were measured with the active spin label first. Then a five-fold excess of ascorbic acid in NMR buffer with adjusted pH was added to reduce the spin label to hydroxylamine, followed by measuring the second data set with the reduced spin label. $^{1}H,^{15}N$-HSQC spectra of active (paramagnetic signal) and reduced (diamagnetic) spin label were overlayed and the peak intensities compared. If a nuclei is in close proximity to the active (oxidized) spin label it relaxes faster and its peak intensity decreases or vanishes completely in the HSQC spectra compared to the inactive label. PREs allow gathering structural and distance information in the surrounding of the spin label. Figure 2.24a shows such an overlay of $^{1}H,^{15}N$-HSQC spectra for Loqs-PD$_{\Delta NC}$ mutants F180CpK (left) and F180TetF (right). The $^{1}H,^{15}N$-HSQC spectrum measured with reduced spin labels is depicted in black, while the $^{1}H,^{15}N$-HSQC spectrum measured oxidized spin labels in red. When the ratio of the peak intensities between paramagnetic and diamagnetic signal are plotted against the sequence of the protein, the PREs are more easily visible, as demonstrated in Figure 2.24b. Both spin labeled amino acids in position F180 (CpK: blue, TetF: orange), located in dsRBD1, have an impact on the same areas, which are residues 144-153, 170-183 and 205-210, as well as additional single residues in dsRBD2, however TetF shows stronger PRE effects than CpK. This was also observed for the other Loqs-PD$_{\Delta NC}$ mutants (K189, D232, F283 and Y292), which might be due to the higher rigidity of TetF, with five less rotational axes compared to CpK (Figure 2.25).[344]

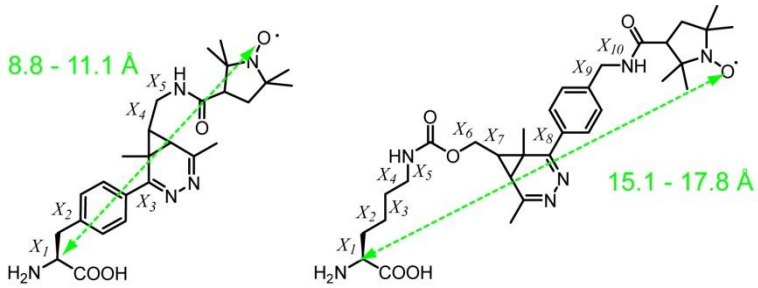


Figure 2.25: Determination of rotational axes on spin labeled TetF (left; five axes) and CpK (right; 10 axes)

The data gathered from the PRE measurements of the different Loqs-PD$_{\Delta NC}$ mutants with TetF and CpK confirmed our initial assumption that a shorter, less flexible unnatural amino acid is advantageous as bioorthogonal handle for spin labeling, since it gives a stronger signal due to less rotational freedom of the ligation product and a shorter distance of the radical to the protein backbone (Figure 2.25). The minimal distances between the Cα atom and the radical determined after energy minimization with MM2 in Chem3D constitute to 15.1 Å (14.9 Å without minimization) for one and 17.2 Å (17.8 Å without minimization) for the other isomer of Tet-TEMPO spin labeled CpK. This is 4-9 Å longer than the distance determined for two stereoisomers after labeling of TetF with Cp-PROXYL. Due to these findings, TetF was used for SDSL in further experiments. The low PRE effects of the spin label placed in one dsRBD on the second dsRBD are minimal and confirm the hypothesis that the two dsRBDs act independently without any dsRNA present.

Since the PRE data of the SDSL of the free Loqs-PD$_{\Delta NC}$ TetF mutants looked promising, the next step was the determination of PRE data of the Loqs-PD$_{\Delta NC}$ TetF mutants in complex with dsRNA to gather information about the domain orientation of the two dsRBDs of Loqs-PD upon binding of target siRNA. Unfortunately, the quality of the measured spectra was not good enough to confirm the hypothesis about a face-to-face arrangement of both dsRBDs and sliding behavior upon binding to dsRNA.

2.3.2.5.4 Determination of intra-domain distances in spin labeled Loqs-PD$_{\Delta NC}$ mutants by PELDOR measurements

Another question that remained to be elucidated for Loqs-PD$_{\Delta NC}$ was the distance between the two dsRBDs without and in presence of duplex RNA, which was addressed by PELDOR measurements of a double spin labeled mutant of Loqs-PD$_{\Delta NC}$. Positions F180 in dsRBD1 and Y292 in dsRBD2 were chosen as positions for a double TAG mutant, which was created by an additional quickchange mutagenesis on the single Loqs-PD$_{\Delta NC}$ mutant. Expression and purification was performed analogously to the single mutants with an increased concentration of TetF (Figure 2.26a), followed by labeling at 4 °C overnight with an excess of Cp-PROXYL. ESI-MS confirmed successful incorporation of TetF into Loqs-PD$_{\Delta NC}$-F180Y292TAG and quantitative modification with spin label (Figure 2.26b). PELDOR measurements of the double spin labeled Loqs-PD$_{\Delta NC}$-F180Y292TetF-Cp-PROXYL mutant without dsRNA show a broad distribution without a distinct distance, and only upon addition of dsRNA a defined distance of 7.4 nm with a peak width of 1.8 nm can be observed (Figure 2.26c-f). This experimental data confirms the hypothesis that the two dsRBDs act independently and adopt a distinct orientation towards each other upon binding of substrate RNA for processing.

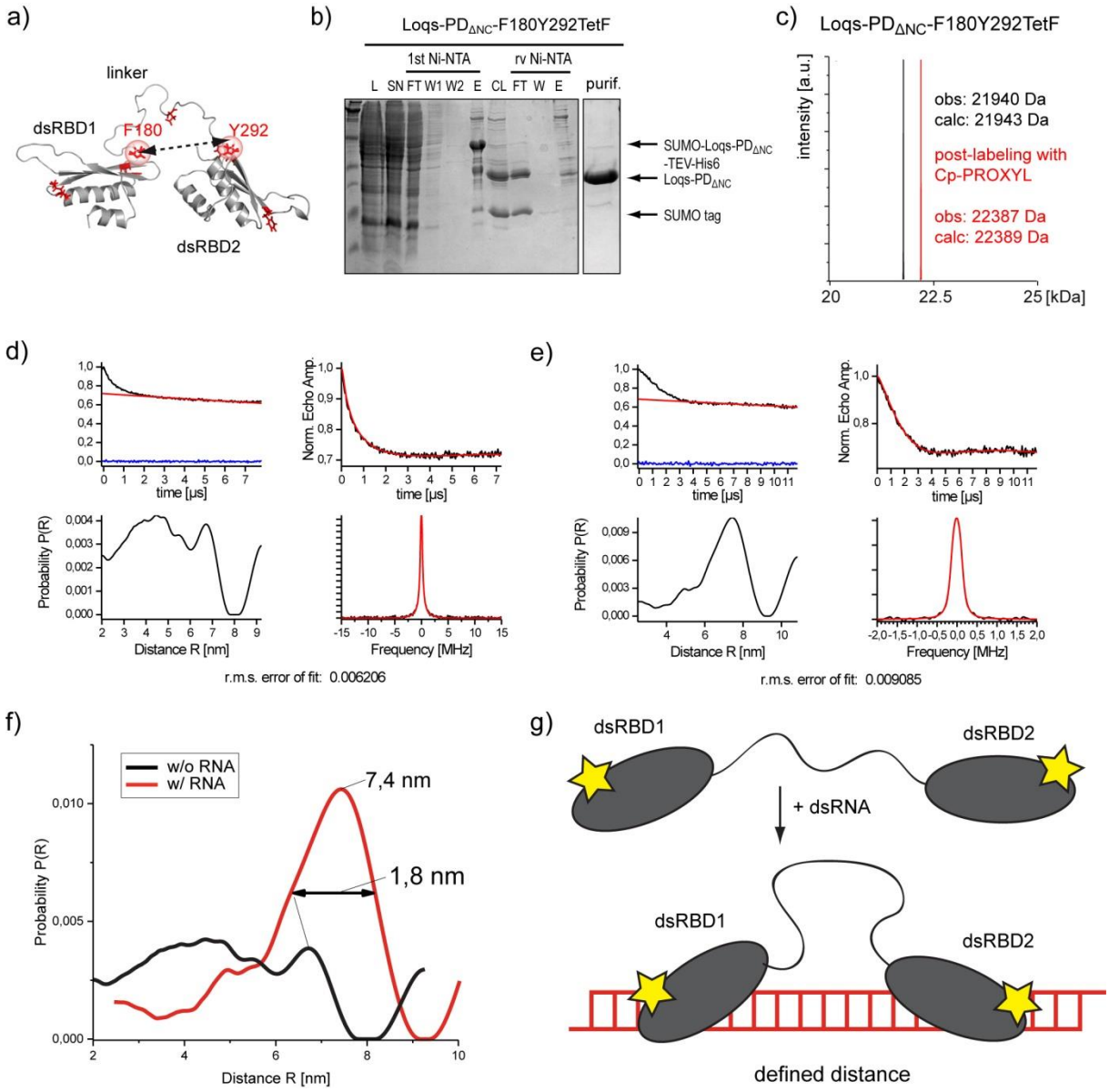


Figure 2.26: PELDOR on Loqs-PD$_{\Delta NC}$-F180Y292TetF-Cp-PROXYL

a) Positions of spin labels in Loqs-PD$_{\Delta NC}$ structure. b) Coomassie stained SDS-PAGE of expression and purification of Loqs-PD$_{\Delta NC}$-F180Y292TetF (FT: flow through, W: wash, E: elution, Cl: cleavage of SUMO and His-tags by TEV and SUMO hydrolase domain, reverse (rv) Ni-NTA purification to remove His-tag and proteases) c) ESI-MS spectra of Loqs-PD$_{\Delta NC}$-F180Y292TetF and after labeling with Cp-PROXYL d) PELDOR data of Loqs-PD$_{\Delta NC}$-F180Y292TetF-Cp-PROXYL without and e) with RNA. Top left: PELDOR time trace. Top right: PELDOR time trace normalized by background subtraction. Bottom left: distance distribution. Bottom right: Fourier transformed time trace. g) Overlay of distance distributions without (black) and with (red) RNA. g) Schematic illustration of the formation of a defined distance between dsRBD1 and dsRBD2 upon binding of dsRNA.

2.4 Summary and Outlook

We have developed a new site-directed spin labeling approach based on the incorporation of rigid 1,2,4,5-tetrazine containing uAAs, which can be decorated with a nitroxide spin label via chemoselective labeling with strained alkenes after protein expression. The inverse-electron demand Diels-Alder reaction occurs rapidly under physiological conditions and results in stable, rigid covalent linkages. We have synthesized a 1,2,4,5-tetrazine analog of biphenyl-alanine and the corresponding reaction partner, a cyclopropene moiety bearing a nitroxide spin label based on published synthetic routes. Incorporation of the new tetrazine uAA into proteins was performed using a mutated tRNA synthetase/tRNA pair from *M. jannaschii*. Model protein sfGFP was used to establish expression and characterization of the subsequent iEDDAC labeling reaction with the Cp-based spin label. Ubiquitin was chosen as a second model protein to characterize our SDSL approach on a rigid system by performing PELDOR measurements, resulting in good spin counting yields and a determined distance distribution between spin labels that is in accordance with Ubiquitin structure and only slightly bit broader than distanced gained via standard MTSSL labeling of cysteine mutants. MD simulations verified the quality of the gathered data.

The newly developed system to site-specifically spin-label proteins via genetic code expansion using tetrazine amino acids was then validated by applying it to the protein Loquacious-PD (Loqs-PD), which plays a role in recognition and processing of endogenous siRNA for downstream RNA interference in *D. melanogaster*. Loqs-PD contains two separate dsRBDs connected by a long and flexible linker. Even though the structure of Loqs-PD of each of the two dsRBDs without the linker was solved by our collaboration partner, questions regarding the distance between the dsRBDs and their orientation towards each other remained to be elucidated. Since Loqs-PD contains several natural cysteine residues, SDSL via MTSSL cannot be applied, requiring the implementation of an alternative strategy to introduce spin labels. Here, our approach could solve these limitations. Amber stop codons were placed at one or more different positions in the dsRBDs and in the linker to decorate them site-specifically with the spin labels. PRE-NMR experiments were then performed on the free Loqs-PD$_{\Delta NC}$ mutants. The introduction of our more rigid spin label based on TetF showed stronger PRE effects than incorporation and spin labeling of CpK, a more flexible lysine-based system. The PRE data showed that without the addition of dsRNA the two dsRBDs act independently in solution. This was confirmed by performance of PELDOR measurements of a Loqs-PD$_{\Delta NC}$ mutant bearing a spin labels in each domain. Only upon addition of dsRNA, a distinct distance is observed.

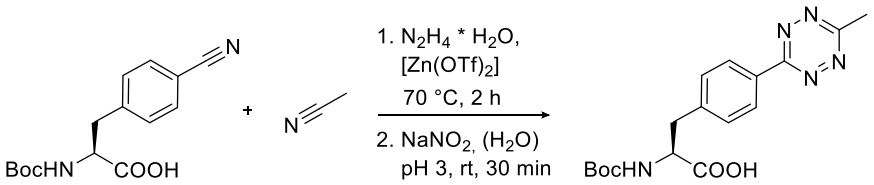
In conclusion, we have developed a powerful tool for site-directed introduction of spin labels into proteins, which have naturally occurring cysteines and therefore elude the standard labeling techniques for the analysis via EPR spectroscopy and NMR spectrometry.

2.5 Experimental

2.5.1 Synthesis routes for TetF

2.5.1.1 Synthesis of TetF from the nitrile precursors

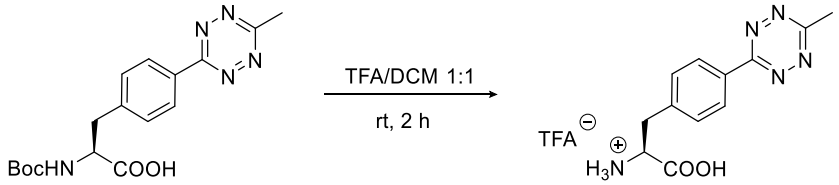
(*tert*-butoxycarbonyl)-*ß*-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alanine[322]



(*tert*-butoxycarbonyl)-*ß*-(4-cyano)phenyl)-L-alanine (2.0 g, 6.8 mmol, 1 eq), $Zn(OTf)_2$ (1.3 g, 3.5 mmol, 0.5 eq), and hydrazine hydrate (11.4 mL, 0.23 mol, 34 eq) were combined in a round flask and heated to 70 °C, before slow addition of acetonitrile (3.6 mL, 69 mmol, 10 eq) using a syringe pump (1.5 mL/h). The reaction mixture was stirred for two hours after addition of acetonitrile (ACN), before cooling to 0 °C. $NaNO_2$ (1.4 g, 0.14 mol, 10 eq) dissolved in minimal amount of water, added to the reaction mixture, which was then slowly acidified to pH 3 by the additon of 6 M HCl solution. The obtained mixture was stirred for 30 minutes on ice until the formation of nitrous fumes was completed, then the water phase was extracted with EtOAc (5x 100 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 before removing the solvent under reduced pressure. The crude residue was purified by flash chromatography (0-60 % EtOAc/DCM + 0.1 % AcOH), which resulted the product (0.85 g, 33 %) as a red solid.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 1.43 (s, 9H, Boc), 3.09 (s, 3H, CH_3), 3.14-3.39 (m, $^3J_{HH}$ = 6.6 Hz, 2H, CH_2), 4.64-4.73 (m, 1H, $CH_{\text{-}\alpha}$), 5.01 (d, $^3J_{HH}$ = 7.9 Hz, 1H, $NH_{\text{-}\alpha}$), 7.43 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar\text{-}2,2'}$), 8.53 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar\text{-}3,3'}$).
MS (ESI), *m/z* calc for $C_{17}H_{21}N_5O_4$ 359.16
found 260.0 $[M\text{-Boc+H}]^+$, 382.0 $[M+Na]^+$, 358.1 $[M\text{-H}]^-$

ß-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alanine trifluoroacetic acid (TetF)[322]



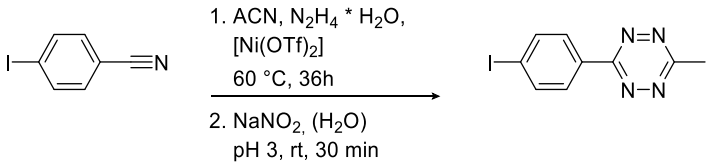
BocTetF (0.36 g, 1.06 mmol, 1 eq) was dissolved in 6 mL DCM, cooled to 0°C before addition of 1.45 mL TFA and a drop of H_2O. The reaction mixture was then stirred at room temperature for two hours. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (0.35 g, 92 %) as a pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 2.99 (s, 3H, CH_3), 3.23 (d, $^3J_{HH}$ = 6.6 Hz, 2H, CH_2), 4.23-4.36 (m, 1H, $CH_{-\alpha}$), 7.56 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar-2,2'}$), 8.36 (s, 3H, $NH_{-\alpha}$), 8.43 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar-3,3'}$).
^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 20.9 (1C, CH_3), 36.0 (1C, CH_2), 53.2 (1C, C_α), 127.9 (2C, $C_{Ar-3,3'}$), 130.8 (2C, $C_{Ar-2,2'}$), 131.1 (1C, C_{Ar-1}), 139.9 (1C, C_{Ar-4}), 163.5 (1C, C_{Tet-6}), 167.3 (1C, C_{Tet-3}), 171.2 (1C, COOH).
MS (ESI), *m/z* calc for $C_{12}H_{13}N_5O_2$ 259.11, found 259.8 $[M+H]^+$, 258.1 $[M-H]^-$

2.5.1.2 Synthesis of TetF via ß-CH-Arylation

2.5.1.2.1 Synthesis of starting materials

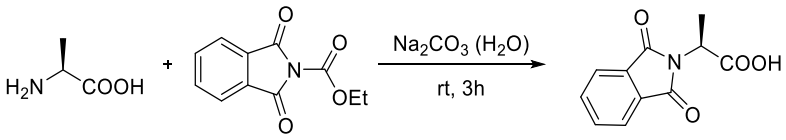
3-(4-iodophenyl)-6-methyl-1,2,4,5-tetrazine[187]



Hydrazine hydrate (13.1 mL, 0.27 mol, 52 eq) was added to a suspension of 4-iodobenzonitrile (1.2 g, 5.2 mmol, 1 eq), $Ni(OTf)_2$ (0.97 g, 2.6 mmol, 0.5 eq) in ACN (2.7 mL, 52 mmol, 10 eq) and heated to 60 °C. After 16 h more ACN (1.4 mL, 26 mmol, 5 eq) was added to the reaction mixture and then stirred for another 20 h. A solution of $NaNO_2$ (3.0 g, 52 mmol, 10 eq) in water was combined with the reaction mixture on ice and slowly acidified to pH 3 by the addition of 6 M HCl solution. After 30 minutes, when the formation of nitrous fumes was completed, the water phase was extracted with EtOAc (5x 100 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 before removing the solvent under reduced pressure. The crude residue was purified by flash chromatography (0-5 % EtOAc/pentane) to yield the product (0.89 g, 58 %) as a purple solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 2.99 (s, 3H, CH_3), 8.05 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar-2,2'}$), 8.23 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar-3,3'}$).

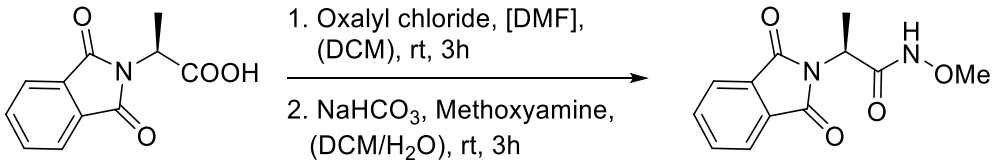
N^2,N^2-Phthaloyl-L-alanine[323]



L-alanine (2 g, 22.5 mmol, 1 eq) and Na_2CO_3 (2.4 g, 22.5 mmol, 1 eq) were dissolved in H_2O (22.5 mL), then *N*-ethoxycarbonylphthalimide (4.9 g, 22.5 mmol, 1 eq) was added in small portions and the resulting reaction mixture stirred at room temperature for 3 h. Upon acidification with 6 M HCl solution at 0 °C to pH 1-2 a white solid precipitated, which was gathered by filtration. It was washed with 1 M HCl solution and a pentane/EtOAc (5:1) mixture and then dried to obtain the product (4.5 g, 91 %) as a white solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.55 (d, $^3J_{HH}$ = 7.3 Hz, 3H, CH_3), 4.87 (q, $^3J_{HH}$ =7.3 Hz, 1H, CH_α), 7.84-7.94 (m, 4H, CH_{Ar}).

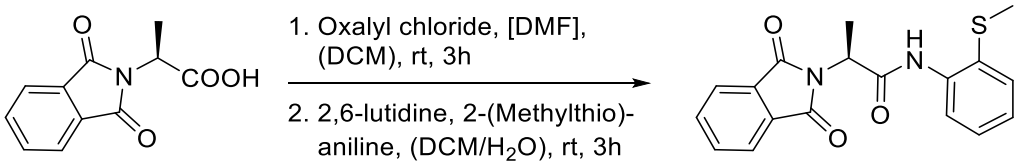
N^1-Methoxy-*N*^{2},*N*^{2}-phthaloyl-L-alaninamide[323]



N^2,*N*^{2}-Phthaloyl-L-alanine (1 g, 4.6 mmol, 1 eq) was dissolved in 10 mL DCM and DMF (5 µL), then oxalyl chloride (0.78 mL, 9.1 mmol, 2 eq) was slowly added and the resulting reaction mixture stirred at room temperature for 3 h. Afterwards the reaction was concentrated under reduced pressure and the crude product redissolved in 5.5 mL DCM. This was slowly added to a vigorously stirred mixture of methoxyamine hydrochloride (0.46 g, 5.5 mmol, 1.2 eq) dissolved in 5.5 mL DCM and $NaHCO_3$ (0.9 g, 11 mmol, 2.4 eq) in 5.5 mL water at 0 °C. The resulting mixture was stirred for 3 h, before extraction of the product with DCM (3x 10 mL), which was then washed with 10 % citric acid (3x 10 mL). The organic phase was dried over Na_2SO_4 and evaporated under reduced pressure. The crude product was recrystallized from DCM/pentane to yield the product (1 g, 91 %) as a white solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 1.69 (d, $^3J_{HH}$ = 7.4 Hz, 3H, CH_3), 3.76 (s, 3H, OCH_3), 4.90-5.05 (m, 1H, CH_α), 7.67-7.79 (m, 2H, $CH_{Ar3,3'}$), 7.82-7.90 (m, 2H, $CH_{Ar-2,2'}$), 8.94 (bs, 1H, NH).

N^1-(2-(methylthio)phenyl)-*N*^{2},*N*^{2}-phthaloyl-L-alaninamide[326]

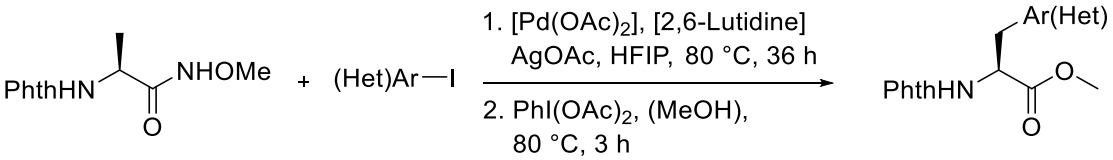


N^2,*N*^{2}-Phthaloyl-L-alanine (1 g, 4.6 mmol, 1 eq) was dissolved in 10 mL DCM and DMF (5 µL), then oxalyl chloride (0.78 mL, 9.1 mmol, 2 eq) was slowly added and the resulting reaction mixture stirred at room temperature for 3 h. Afterwards the reaction was concentrated under reduced pressure and the crude product redissolved in 5.5 mL DCM. This was slowly added to a solution of 2-(methylthio)aniline (0.69 mL, 5.5 mmol, 1.2 eq) and 2,6-lutidine (0.53 mL, 4.6 mmol, 1 eq) in 7.5 mL DCM at 0 °C. The resulting mixture was stirred for 3 h, then after dilution with DCM (10 mL) it was washed with 10 % citric acid (2x 10 mL) and saturated $NaHCO_3$ (1x 10 mL) solution. The organic phase was dried over Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by flash chromatography (0-5 % EtOAc/DCM) to yield the product (1.4 g, 90 %) as a slightly brown solid.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 1.87 (d, $^3J_{HH}$ = 7.3 Hz, 3H, CH_3), 2.26 (s, 3H, SCH_3), 5.15 (q, $^3J_{HH}$ =7.3 Hz, 1H, CH_α), 7.06 (dt, $^3J_{HH}$ = 7.7, 1.6 Hz, 1H, $CH_{Ar\text{-}4}$), 7.30 (ddd, $^3J_{HH}$ = 8.4, 7.3, 1.6 Hz, 1H, $CH_{Ar\text{-}5}$), 7.47 (dd, $^3J_{HH}$ = 7.7, 1.6 Hz, 1H, $CH_{Ar\text{-}3}$),7.76 (dd, $^3J_{HH}$ = 5.5, 3.1 Hz, 2H, $CH_{Ar(phthaloyl)\text{-}3,3'}$), 7.90 (dd, $^3J_{HH}$ = 5.5 Hz, $^4J_{HH}$ = 3.1 Hz, 2H, $CH_{Ar(phthaloyl)\text{-}2,2'}$), 8.33 (d, $^3J_{HH}$ = 8.4 Hz, 1H, $CH_{Ar\text{-}6}$), 8.93 (bs, 1H, NH).
MS (ESI), *m/z* calc for $C_{18}H_{16}N_2O_3S$ 340.09, found 363.2 $[M+Na]^+$, 395.2 $[M+MeOH+Na]^+$

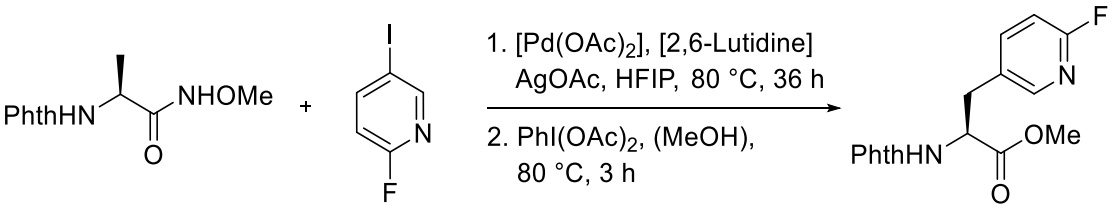
2.5.1.2.2 *β*-CH-Arylation reactions

General procedure using N^1-Methoxy-N^2,N^2-phthaloyl-L-alaninamide as starting material[323]



N^1-Methoxy-N^2,N^2-phthaloyl-L-alaninamide (62 mg, 0.25 mmol, 1 eq), the corresponding aryl iodide (0.38 mmol, 1.5 eq), $Pd(OAc)_2$ (8.4 mg, 0.038 mmol, 0.15 eq), AgOAc (84 mg, 0.5 mmol, 2 eq) 2,6-lutidine (8.7 µL, 0.075 mmol, 0.3 eq) and 2.5 mL HFIP were combined in a glass ampule (⌀ = 14.1 mm) including a stirrer. The glass ampule was then frozen in a dry ice/acetone ice bath, flushed with argon and evacuated, before being molten closed in liquid nitrogen. The reaction mixture was warmed up to room temperature before heating to 80 °C for 60 h. Then the ampule was broken and the reaction mixture filtered over Celite. The solvents were removed and the crude product dissolved in MeOH (2.5 mL). $PhI(OAc)_2$ (84 mg, 0.25 mmol, 1 eq) was added and the reaction heated to 80 °C for 3 h. Then the mixture was cooled, filtered over Celite and the solvent removed. The product was purified by flash chromatography (0-20 % EtOAc/pentane) to obtain the product.

Methyl N^2,N^2-phthaloyl-ß-(6-fluoropyridin-3-yl)-L-alaninate



The reaction was performed with 2-fluoro-5-iodopyridine (84 mg, 0.38 mmol, 1.5 eq) and yielded the product (33 mg, 40 %) as slightly brown solid.

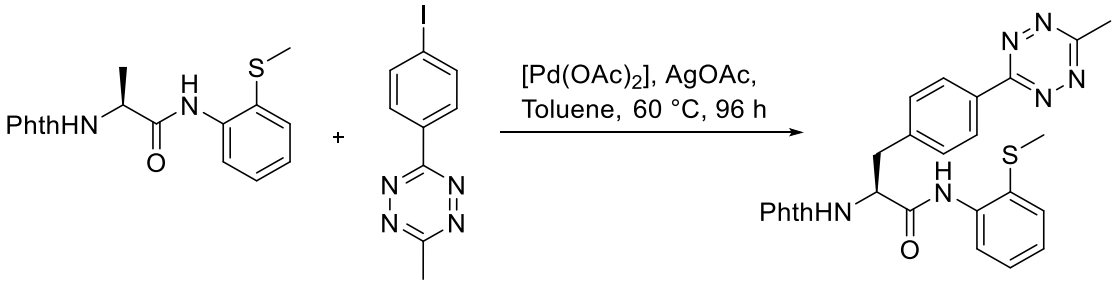
^{1}H-NMR (400 MHz, $CDCl_3$): δ = 3.53-3.57 (m, 2H, CH_2), 3.77 (s, 3H, CH_3), 5.08 (dd, $^3J_{HH}$ = 9.8, 6.5 Hz, 1H, CH_α), 6.80 (dd, $^3J_{HH}$ = 8.4 Hz, $^4J_{HH}$ = 2.7 Hz, 1H, $CH_{Ar\text{-}4}$), 7.65 (ddd, $^3J_{HF}$ =

7.3 Hz, $^3J_{HH}$ = 8.4, $^4J_{HH}$ = 2.7 Hz, 1H, $CH_{Ar\text{-}5}$), 7.69-7.73 (m, 2H, $CH_{Ar(phthaloyl)\text{-}3,3'}$), 7.77-7.80 (m, 2H, $CH_{Ar(phthaloyl)\text{-}2,2'}$), 7.96 (d, $^4J_{HH}$ = 2.2 Hz, 1H, $CH_{Ar\text{-}2}$).
MS (ESI), *m/z* calc for $C_{17}H_{13}FN_2O_4$ 328.09, found 383.2 $[M+MeOH+Na]^+$

Methyl N^2,N^2-phthaloyl-ß-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alaninate
The reaction was performed with 3-(4-iodophenyl)-6-methyl-1,2,4,5-tetrazine (0.11 g, 0.38 mmol, 1.5 eq) and did not yield any product.

General procedure using N^1-(2-(methylthio)phenyl)-N^2,N^2-phthaloyl-L-alaninamide as starting material[324]

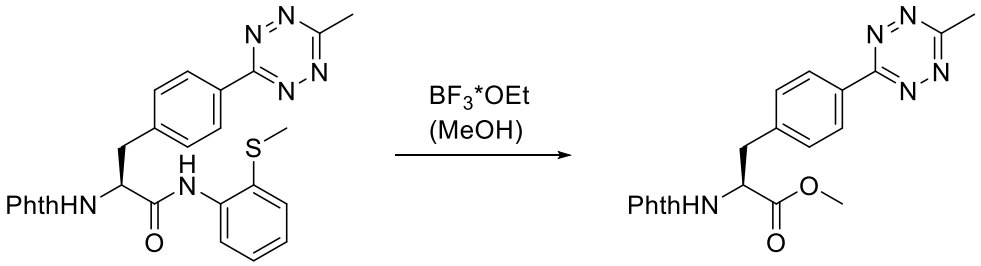
N^1-(2-(methylthio)phenyl)-N^2,N^2-phthaloyl-ß-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alaninamide



N^1-(2-(methylthio)phenyl)-N^2,N^2-phthaloyl-L-alaninamide (0.17 g, 0.5 mmol, 1 eq), the corresponding aryl iodide (0.45 g, 1.5 mmol, 3 eq), $Pd(OAc)_2$ (6 mg, 0.027 mmol, 0.05 eq), Ag(OAc) (0.21 g, 1.25 mmol, 2.5 eq) and 0.4 mL toluene were combined in a flask and the reaction heated to 60 °C and stirred for 96 h. The reaction was then diluted with 10 mL brine and extracted with DCM (3x 10 mL). The organic phase was dried over Na_2SO_4 and the solvent removed under reduced pressure. The product was purified by flash chromatography (0-50 % EtOAc/pentane) to obtain the product (75 mg, 30 %) as a pink oil.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 2.19 (s, 3H, SCH_3), 3.06 (s, 3H, CH_3), 3.81-3.86 (m, 2H, CH_2), 5.38 (dd, $^3J_{HH}$ =10.0, 6.6 Hz, 1H, CH_α), 7.08 (dt, $^3J_{HH}$ = 7.6, 1.4 Hz, 1H, $CH_{Ar\text{-}4}$), 7.32 (dd, $^3J_{HH}$ = 7.9, 1.5 Hz, 1H, $CH_{Ar\text{-}5}$), 7.38-7.45 (dd, $^3J_{HH}$ = 7.7, 1.6 Hz, 1H, $CH_{Ar\text{-}3}$),7.45 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar(phenyl)\text{-}2,2'}$), 7.70-7.74 (m, 2H, $CH_{Ar(phthaloyl)\text{-}3,3'}$), 7.80-7.84 (m, 2H, $CH_{Ar(phthaloyl)\text{-}2,2'}$), 8.35 (d, $^3J_{HH}$ = 8.4 Hz, 1H, $CH_{Ar\text{-}6}$), 8.43 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar(phenyl)\text{-}3,3'}$), 8.98 (bs, 1H, NH).
MS (ESI), *m/z* calc for $C_{27}H_{22}N_6O_3S$ 510.15, found 533.2 $[M+Na]^+$, 565.2 $[M+MeOH+Na]^+$

Methyl N^2,N^2-phthaloyl-ß-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alaninate

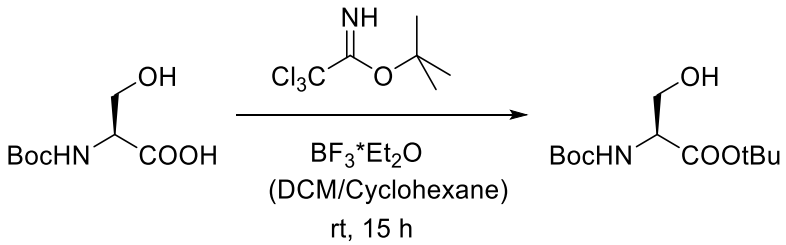


N^1-(2-(methylthio)phenyl)-N^2,N^2-phthaloyl-ß-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alaninamide (60 mg, 0.12 mmol, 1 eq) was suspended in 3.5 mL MeOH, then BF_3*Et_2O (0.17 mL, 1.4 mmol, 12 eq) was added and the reaction heated to 100 °C for 24 h. NEt_3 (0.29 mL, 2.1 mmol, 17.5 eq) was slowly added after cooling to quench the reaction. After evaporation of the solvent the crude product was purified by flash chromatography (0-50 % EtOAc/pentane) to obtain the product (19 mg, 39 %) as a pink oil.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.05 (s, 3H, CH_3), 3.65-3.72 (m, 2H, CH_2), 3.80 (s, 3H, OCH_3), 5.23 (dd, $^3J_{HH}$ =10.4, 6.1 Hz, 1H, CH_α), ,7.40 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar(phenyl)-2,2'}$), 7.67-7.72 (m, 2H, $CH_{Ar(phthaloyl)-3,3'}$), 7.77-7.81 (m, 2H, $CH_{Ar(phthaloyl)-2,2'}$), 8.42 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar(phenyl)-3,3'}$).

2.5.1.3 Synthesis of TetF via Negishi cross-coupling reaction

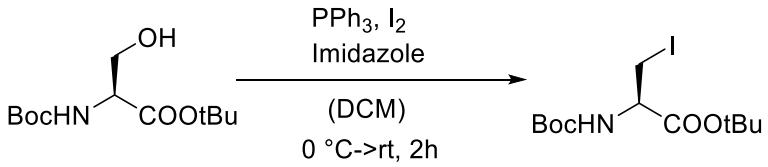
***tert*-butyl (*tert*-butoxycarbonyl)-L-serinate[329b]**



The reaction was performed under Schlenk conditions. *tert*-Butyl 2,2,2-trichloroacetimidate (8.9 mL, 50 mmol, 2 eq) was dissolved in 50 mL cyclohexane and then added to a solution of (*tert*-butoxycarbonyl)-L-serine (5.1 g, 25 mmol, 1 eq) in 100 mL DCM. After addition of BF_3*Et_2O (0.3 mL, 2.5 mmol, 0.1 eq) the resulting reaction mixture was stirred at room temperature overnight before quenching with solid $NaHCO_3$. The filtrate was evaporated and the crude product purified by flash chromatography (0-15 % EtOAc/pentane) to obtain the product (2.56 g, 58 %) as a white solid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.44 (s, 9H, Boc), 1.47 (s, 9H, OtBu) 3.88 (d, $^3J_{HH}$ = 3.9 Hz, 2H, CH_2), 4.17-4.28 (m, 1H, $CH_{-\alpha}$), 5.45 (s, 1H, $NH_{-\alpha}$).

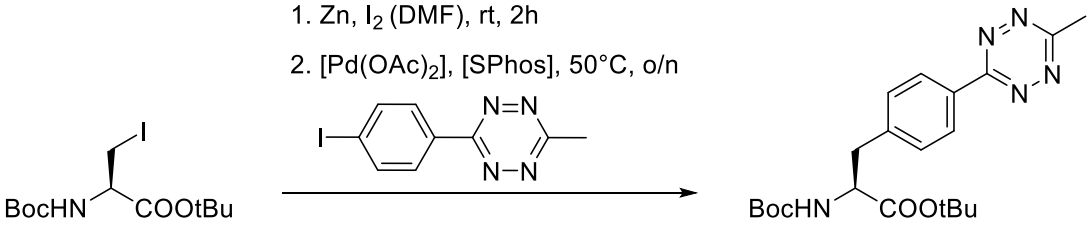
***tert*-butyl (*tert*-butoxycarbonyl)-*ß*-iodo-L-alaninate[329a]**



The reaction was performed under Schlenk conditions. Triphenylphosphine (3.2 g, 12.3 mmol, 1.25 eq) and imidazole (0.87 g, 12.7 mmol, 1.3 eq) were dissolved in 100 mL DCM and cooled to 0 °C. Then iodine (3.2 g, 12.7 mmol, 1.3 eq) was added in three portions over 30 minutes before a solution of *tert*-butyl (*tert*-butoxycarbonyl)-L-serinate (2.6 g, 9.8 mmol, 1 eq) in DCM (10 mL) was slowly added over 15 minutes. The resulting reaction mixture was stirred at 0 °C for 1 hour before warming to room temperature, where it was stirred for another 1.5 hours before quenching with 5 mL methanol. The organic phase was washed with a 20 % $Na_2S_2O_3$ solution (3x) and brine (1x), then dried over Na_2SO_4 and the solvent evaporated under reduced pressure. The crude product was purified by flash chromatography (0-2 % EtOAc/pentane) to result the product (2.9 g, 63 %) as a white solid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.45 (s, 9H, Boc), 1.50 (s, 9H, OtBu) 3.56 (d, $^3J_{HH}$ = 3.7 Hz, 2H, CH_2), 4.34 (dt, $^3J_{HH}$ = 7.4 Hz, 3.7 Hz, 1H, $CH_{-\alpha}$), 5.45 (d, $^3J_{HH}$ = 7.4 Hz, 1H, $NH_{-\alpha}$).
^{13}C-NMR (75 MHz, $CDCl_3$): δ = 9.1 (1C, CH_2), 28.2 (3C, CH_3-Boc), 28.4 (3C, CH_3-OtBu), 53.9 (1C, $C_{-\alpha}$), 80.4 (1C, NCOO**C**$(CH_3)_3$), 83.4 (1C, COO**C**$(CH_3)_3$), 155.1 (1C, N**C**OOtBu), 168.6 (1C, **C**OOtBu).

***tert*-butyl (*tert*-butoxycarbonyl)-*ß*-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alaninate (BocTetFOtBu)[332]**

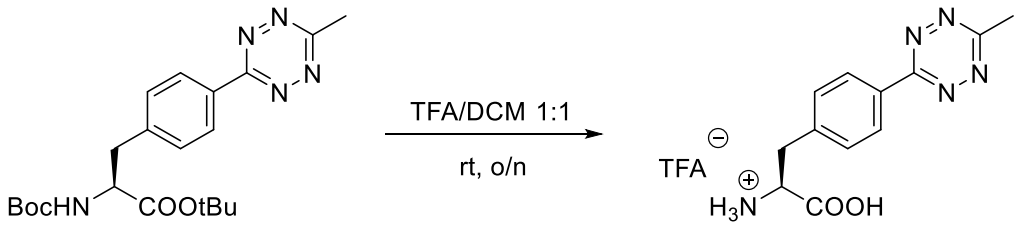


The reaction was performed under Schlenk conditions. 3 mL DMF and iodine (19 mg, 0.08 mmol, 0.15 eq) were added to a 3-neck flask containing zinc dust (98 mg, 1.5 mmol, 3 eq) and stirred vigorously (~750 rpm) for 15 minutes at room temperature. After the solution turned from orange to grey Boc-*β*-iodo-Ala-OtBu (0.19 g, 0.5 mmol, 1 eq) and iodine (19 mg, 0.08 mmol, 0.15 eq) were added and the reaction vigorously stirred (~750 rpm) for 2 hours at room temperature. Then, 3-(4-iodophenyl)-6-methyl-1,2,4,5-tetrazine (0.19 g, 0.65 mmol, 1.3 eq), $Pd(OAc)_2$ (3 mg, 0.013 mmol, 0.025 eq) and SPhos (10 mg, 0.025 mmol, 0.05 eq) were added and the resulting reaction mixture heated to 50 °C and vigorously (~750 rpm) stirred for 5 hours. The reaction was filtered over Celite and the solvent

evaporated under reduced pressure before purifiying the crude product by flash chromatography (0-20 % EtOAc/pentane) to obtain the product (0.17 g, 78 %) as a red oil.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.42 (s, 9H, Boc), 1.43 (s, 9H, OtBu), 3.09 (s, 3H, CH_3), 3.17 (t, $^3J_{HH}$ = 6.3 Hz, 2H, CH_2), 4.46-4.57 (m, 1H, $CH_{\text{-}\alpha}$), 5.08 (d, $^3J_{HH}$ = 8.1 Hz, 1H, $NH_{\text{-}\alpha}$), 7.40 (d, $^3J_{HH}$ = 8.3 Hz, 2H, $CH_{Ar\text{-}2,2'}$), 8.52 (d, $^3J_{HH}$ = 8.3 Hz, 2H, $CH_{Ar\text{-}3,3'}$).

ß-3-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alanine trifluoracetate (TetF)

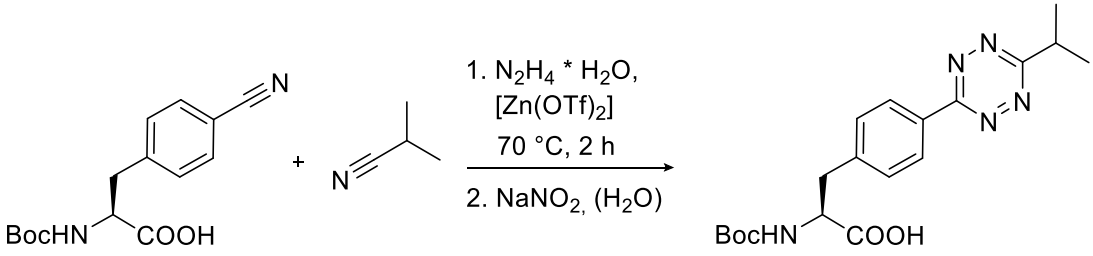


BocTetFOtBu (0.13 g, 0.32 mmol, 1 eq) was dissolved in 1.5 mL DCM, cooled to 0°C before addition of 1.5 mL TFA and a drop of H_2O. The reaction mixture was then stirred at room temperature overnight. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (0.11 g, 90 %) as pink solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 3.01 (s, 3H, CH_3), 3.22 (d, $^3J_{HH}$ = 6.0 Hz, 2H, CH_2), 4.26-4.36 (m, 1H, $CH_{\text{-}\alpha}$), 7.56 (d, $^3J_{HH}$ = 8.3 Hz, 2H, $CH_{Ar\text{-}2,2'}$), 8.34 (s, 3H, $NH_{\text{-}\alpha}$), 8.43 (d, $^3J_{HH}$ = 8.3 Hz, 2H, $CH_{Ar\text{-}3,3'}$).
MS (ESI), *m/z* calc for $C_{12}H_{13}N_5O_2$ 259.11, found 259.8 $[M+H]^+$, 258.1 $[M-H]^-$

2.5.2 Synthesis of IpTetF

(*tert*-butoxycarbonyl)-*ß*-(4-(6-isopropyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alanine (BocIpTetF)[322]

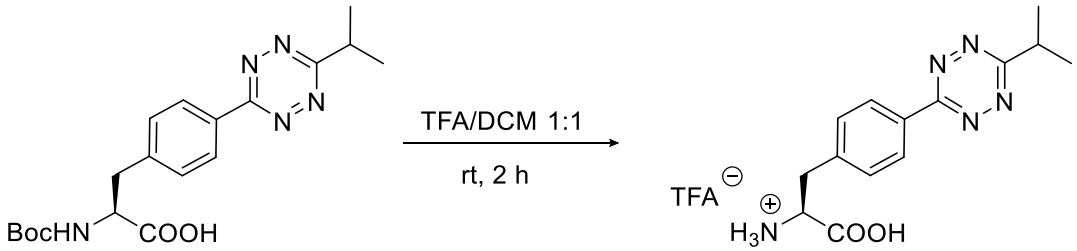


(*tert*-butoxycarbonyl)-ß-(4-cyano)phenyl)-L-alanine (5.0 g, 17.2 mmol, 1 eq), $Zn(OTf)_2$ (3.1 g, 8.6 mmol, 0.5 eq), and hydrazine hydrate (18.4 mL, 0.38 mol, 22 eq) were combined in a round flask and heated to 70 °C, before slow addition of isobutyronitrile (15.5 mL, 0.172 mol, 10 eq) using a syringe pump (3 mL/h). The reaction mixture was stirred for two hours after addition of isobutyronitrile, before cooling to 0 °C. $NaNO_2$ (11.9 g, 0.172 mol, 10 eq) dissolved in 5 mL H_2O was added to the reaction mixture, which was then slowly acidified to

pH 3 by the additon of 6 M HCl solution. It was stirred for 30 minutes on ice until the formation of nitrous fumes was completed, then the water phase was extracted with EtOAc (5x 100 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 before removing the solvent under reduced pressure. The crude residue was purified by flash chromatography (0-5 % MeOH/DCM + 0.1 % AcOH), which resulted the product (1.46 g, 17 %) as a red solid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.42 (s, 9H, Boc), 1.56 (d, $^3J_{HH}$ = 7.0 Hz, 6H, CH_3 Ip), 3.14-3.39 (m, $^3J_{HH}$ = 6.6 Hz, 2H, CH_2), 3.68 (h, $^3J_{HH}$ = 7.0 Hz, 1H, CH Ip), 4.64-4.73 (m, 1H, $CH_{-\alpha}$), 5.02 (d, $^3J_{HH}$ = 8.0 Hz, 1H, $NH_{-\alpha}$), 7.43 (d, $^3J_{HH}$ = 8.3 Hz, 2H, $CH_{Ar\text{-}2,2'}$), 8.55 (d, $^3J_{HH}$ = 8.3 Hz, 2H, $CH_{Ar\text{-}3,3'}$).
MS (ESI), *m/z* calc for $C_{19}H_{25}N_5O_4$ 387.19, found 288.2 $[M\text{-Boc+H}]^+$, 386.1 $[M\text{-H}]^-$

ß-(4-(6-isopropyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alanine trifluoroacetate (IpTetF)



BocIpTetF (0.21 g, 0.55 mmol, 1 eq) was dissolved in 1.5 mL DCM, cooled to 0°C before addition of 1.5 mL TFA. The reaction mixture was then stirred at room temperature overnight. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (0.15 g, 70 %) as a pink solid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.48 (d, $^3J_{HH}$ = 7.0 Hz, 6H, CH_3 Ip), 3.25 (d, $^3J_{HH}$ = 6.6 Hz, 2H, CH_2), 3.60 (h, $^3J_{HH}$ = 6.9 Hz, 1H, CH Ip), 4.25-4.36 (m, 1H, CH_α), 7.57 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar\text{-}2,2'}$), 8.43 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $CH_{Ar\text{-}3,3'}$), 8.40-8.51 (s, 3H, NH_α).
^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 21.0 (3C, CH_3-Ip), 33.6 (1C, CH-Ip), 35.9 (1C, CH_2), 53.1 (1C, C_α), 127.8 (2C, $C_{Ar\text{-}3,3'}$), 130.6 (2C, $C_{Ar\text{-}2,2'}$), 131.0 (1C, $C_{Ar\text{-}1}$), 139.8 (1C, $C_{Ar\text{-}4}$), 163.6 (1C, $C_{Tet\text{-}6}$), 170.3 (1C, $C_{Tet\text{-}3}$), 173.1 (1C, COOH).
MS (ESI), *m/z* calc for $C_{14}H_{17}N_5O_2$ 287.14, found 288.2 $[M+H]^+$, 286.1 $[M\text{-H}]^-$

2.5.3 Synthesis routes for TetA

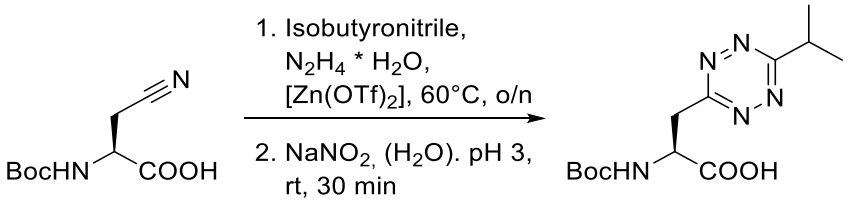
2.5.3.1 Synthesis of TetA from the nitriles and nitrile derivatives

General procedure:

(*tert*-butoxycarbonyl)-*ß*-cyano-L-alanine (1 eq), catalyst (0.5-3 eq), and hydrazine hydrate (20-50 eq) were combined in a round flask and heated to 60-70 °C, before addition of acetonitrile (10 eq), isobutyronitrile (10 eq) using a syringe pump (1.5 mL/h) or acetamidine hydrochloride (10 eq). The reaction mixture was stirred overnight, before cooling to 0 °C. $NaNO_2$ (10 eq) dissolved in minimal amount of water was added to the reaction mixture, which was then slowly acidified to pH 3 by the additon of 6 M HCl solution. It was stirred for 30 minutes on ice until the formation of nitrous fumes was completed, then the water phase was extracted with EtOAc (5x 10 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 before removing the solvent under reduced pressure.

Reaction conditions using (*tert*-butoxycarbonyl)-*ß*-cyano-L-alanine as starting material[322]

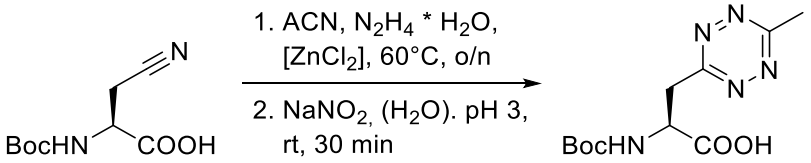
Condition A – (*tert*-butoxycarbonyl)-*ß*-(6-isopropyl-1,2,4,5-tetrazin-3-yl)-L-alanine



The reaction was performed with (*tert*-butoxycarbonyl)-*ß*-(4-cyano)-L-alanine (0.5 g, 4.7 mmol, 5 eq), isobutyronitrile (42 µL, 0.46 mmol, 1 eq), $Zn(OTf)_2$ (84 mg, 0.23 mmol, 0.5 eq) and did yield the product (15 mg, 10 %) as a red oil.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 1.40 (s, 9H, Boc), 1.53 (d, $^3J_{HH}$ = 7.0 Hz, 6H, CH_3-Ip), 3.64 (sp, $^3J_{HH}$ = 7.0 Hz, 1H, CH-Ip), 3.82 (dd, $^3J_{HH}$ = 15.7, 6.7 Hz, 1H, CH_2), 3.91 (dd, $^3J_{HH}$ = 15.7, 4.8 Hz, 1H, CH_2), 4.92-4.99 (m, 1H, CH_α), 5.43 (d, $^3J_{HH}$ = 8.0 Hz, 1H, NH_α).
MS (ESI), *m/z* calc for $C_{13}H_{21}N_5O_4$ 311.16, found 334.2 $[M+Na]^+$, 310.1 $[M-H]^-$

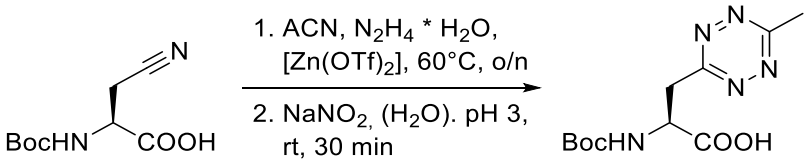
Condition B - (*tert*-butoxycarbonyl)-*ß*-(6-methyl-1,2,4,5-tetrazin-3-yl)-L-alanine



The reaction was performed with ACN (0.4 mL, 4.7 mmol, 10 eq) and $ZnCl_2$ (32 mg, 0.23 mmol, 0.5 eq) at 70 °C and did yield the product (4 mg, 3 %) as a red oil.

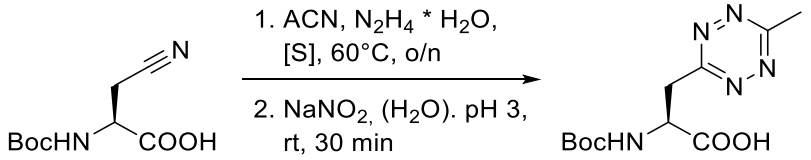
MS (ESI), *m/z* calc for $C_{11}H_{17}N_5O_4$ 283.13, found 284.2 $[M+H]^+$, 283.3 $[M-H]^-$

Condition C



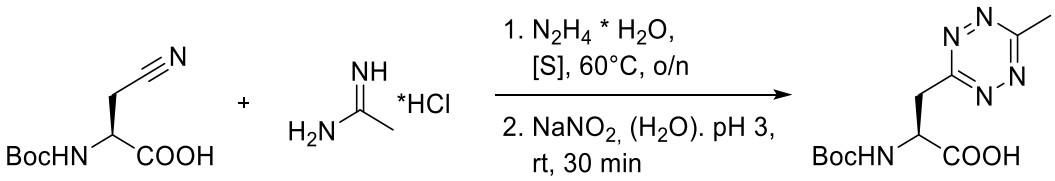
The reaction was performed with $Zn(OTf)_2$ (0.17 g, 0.47 mmol, 0.5 eq), hydrazine hydrate (22 eq) and ACN (0.83 mL, 9.3 mmol, 10 eq), which was added using a syringe pump (1.5 mL/h). The reaction did not yield any product.

Condition D[174a]



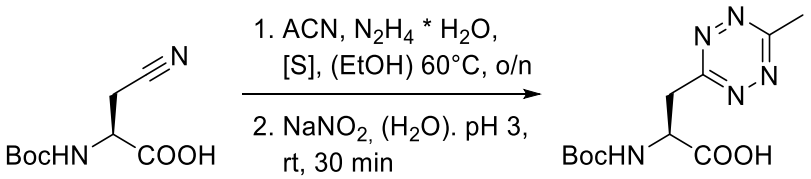
The reaction was performed with ACN (0.24 mL, 4.65 mmol, 10 eq), sulfur (69 mg, 1.87 mmol, 4 eq) and hydrazine hydrate (20 eq) and did not yield any product.

Condition D



The reaction was performed with acetamidine hydrochloride (0.44 g, 4.65 mmol, 10 eq) and sulfur (45 mg, 1.39 mmol, 3 eq) and hydrazine hydrate (50 eq) and did not yield any product.

Condition E[174b]

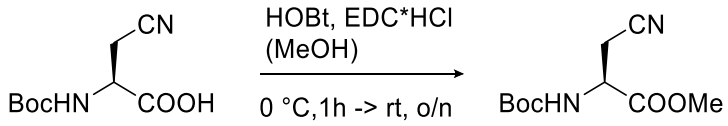


(*tert*-butoxycarbonyl)-*ß*-cyano-L-alanine (0.1 g, 0.47 mmol, 1 eq), sulfur (69 mg, 1.87 mmol, 4 eq), ACN (0.24 mL, 4.7 mmol, 10 eq) and hydrazine hydrate (0.45 mL, 9.3 mmol, 20 eq) were combined in a round flask with 3 mL EtOH and heated to reflux overnight, before cooling to 0 °C. $NaNO_2$ (0.64 g, 9.3 mmol, 5 eq) dissolved in minimal amount of water, added to the reaction mixture, which was then slowly acidified to pH 3 by the additon of 6 M

HCl solution. It was stirred for 30 minutes on ice until the formation of nitrous fumes was completed, then the water phase was extracted with EtOAc (5x 10 mL) until it wasn't pink anymore. The organic phase was dried over Na_2SO_4 before removing the solvent under reduced pressure. The reaction did not yield any product.

<u>Synthesis route starting from (*tert*-butoxycarbonyl)-*ß*-(ethoxyimido)-L-alanine</u>

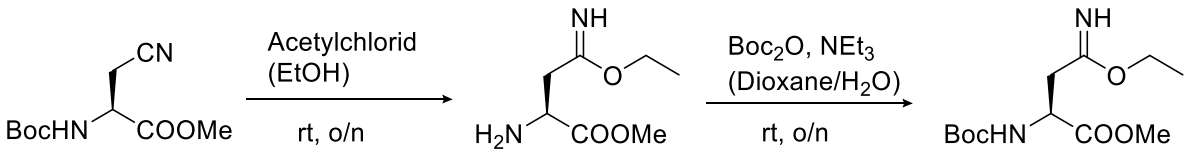
Methyl (*tert*-butoxycarbonyl)-*ß*-(4-cyano)-L-alaninate



HOBt (0.86 g, 5.6 mmol, 1.2 eq) was added to a solution of (*tert*-butoxycarbonyl)-*ß*-(4-cyano)-L-alanine (1 g, 4.7 mmol, 1 eq) in 10 mL MeOH and then cooled to 0 °C. After addition of EDC*HCl (1.1 mL, 5.6 mmol, 1.2 eq) the resulting reaction mixture was stirred for 1 h at 0 °C, followed by 5 h at room temperature. The solvent was evaporated and the crude product purified by flash chromatography (0-20 % EtOAc/pentane) to obtain the product (0.89 g, 84 %) as a white solid.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 1.46 (s, 9H, Boc), 2.97 (dd, $^3J_{HH}$ = 16.8, 4.5 Hz, 1H, $CH_ß$), 3.01 (dd, $^3J_{HH}$ = 16.8, 5.5 Hz, 1H, $CH_ß$) 3.85 (s, 3H, OCH_3), 4.49-4.56 (m, 1H, $CH_α$), 5.45 (d, $^3J_{HH}$ = 7.8 Hz, 1H, $NH_α$).

Methyl *ß*-(ethoxy(imino)methyl)-L-alaninate [331]



Methyl (*tert*-butoxycarbonyl)-*ß*-(4-cyano)-L-alaninate (0.43 g, 1.9 mmol, 1 eq) was dissolved in EtOH (1.2 mL, 23 mmol, 12 eq), cooled to 0 °C, then slowly acetyl chloride (1.1 mL, 15 mmol, 8 eq) was added and the reaction stirred at room temperature overnight. Afterwards the solvents were evaporated under reduced pressure to obtain the product as a white solid, which was used crude in the following step. Therefore it was dissolved in 5 mL dioxane/H_2O (9:1) and NEt_3 (0.7 mL, 4.8 mmol, 2.5 eq), cooled in an ice bath and slowly Boc_2O (0.48 mL, 2.1 mmol, 1.1 eq) was added. The resulting reaction mixture was stirred at room temperature overnight. Dioxane was removed under reduced pressure and the product extracted from the waterphase with DCM, which was dried over Na_2SO_4, filtered and the solvent evaporated. The crude product was purified by flash chromatography (0-20 % Et_2O/pentane) to obtain the product (0.29 g, 55 % over two steps) as a white solid.

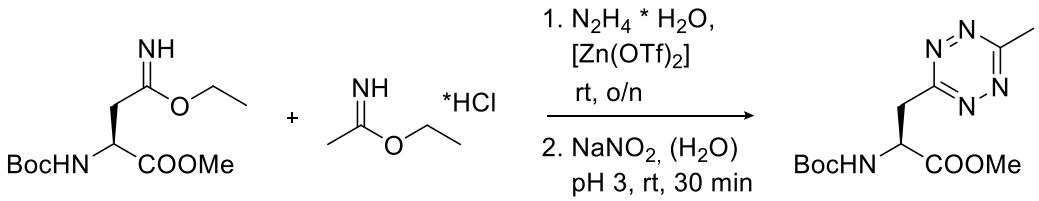
Methyl *ß*-(ethoxy(imino)methyl)-L-alaninate hydrochloride

^{1}H-NMR (500 MHz, D_2O): δ = 1.26 (t, $^3J_{HH}$ = 7.1 Hz, 3H, CH_3), 3.11 (dd, $^3J_{HH}$ = 18.2, 4.6 Hz, 1H, $CH_ß$), 3.20 (dd, $^3J_{HH}$ = 18.2, 6.1 Hz, 1H, $CH_ß$), 3.85 (s, 3H, OCH_3), 4.17-4.26 (m, 2H, OCH_2), 4.50 (dd, $^3J_{HH}$ = 6.1, 4.6 Hz, 1H, $CH_α$).

Methyl (*tert*-butoxycarbonyl)-*ß*-(ethoxy(imino)methyl)-L-alaninate

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.25 (t, $^3J_{HH}$ = 7.1 Hz, 3H, CH_3), 1.44 (s, 9H, Boc), 2.79 (dd, $^3J_{HH}$ = 17.0, 4.7 Hz, 1H, $CH_ß$), 2.98 (dd, $^3J_{HH}$ = 17.0, 4.7 Hz, 1H, $CH_ß$) 3.75 (s, 3H, OCH_3), 4.14 (q, $^3J_{HH}$ = 7.1 Hz, 2H, OCH_2), 4.51-4.60 (m, 1H, $CH_α$), 5.45 (d, $^3J_{HH}$ = 8.8 Hz, 1H, $NH_α$).

Condition F

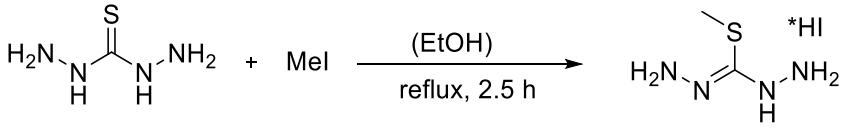


The reaction was performed with methyl (*tert*-butoxycarbonyl)-*ß*-(ethoxy(imino)methyl)-L-alaninate (0.1 g, 0.37 mmol, 1 eq), ethyl acetimidate hydrochloride (0.22 g, 1.8 mmol, 5 eq), $Zn(OTf)_2$ (66 mg, 0.18 mmol, 0.5 eq) and hydrazine hydrate (0.35 mL, 7.3 mmol, 20 eq) and did not yield any product.

2.5.3.3 Synthesis of TetA via Negishi cross-coupling reactions

2.5.3.3.1 Synthesis of starting materials

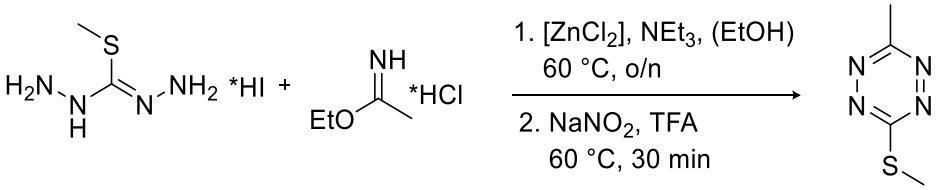
S-Methylisothiocarbohydrazide hydroiodide[333b]



Thiocarbohydrazide (5 g, 47 mmol, 1 eq) was suspended in a three-neck-flask in 160 mL absolute ethanol and heated to reflux. Methyl iodide (3.2 mL, 52 mmol, 1.1 eq) was added to the suspension, which was then stirred for 2.5 hours until everything was dissolved. The hot solution was filtered through a porous funnel without filter paper. Upon cooling the product crystallized from the solution as a white solid (7.4 g, 64 %), which was filtered off and dried under vacuum.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 2.38 (s, 3H, CH_3)
MS (ESI), *m/z* calc for $C_2H_8N_4S$ 120.05, found 121.1 $[M+H]^+$

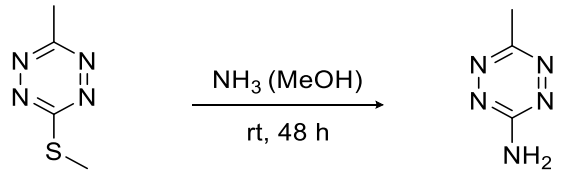
3-*S*-Methylthio-6-methyl-1,2,4,5-tetrazine[333b]



S-Methylisothiocarbohydrazide hydroiodide (5 g, 20 mmol, 1 eq) and zinc chloride (1.5 g, 10.1 mmol, 0.5 eq) were dissolved in 100 mL ethanol and the resulting solution heated to 60 °C. Then ethyl acetimidate hydrochloride (2.7 g, 22 mmol, 1.1 eq) and triethylamine (4.6 mL, 33 mmol, 1.65 eq) were added to the solution, which was stirred overnight. The next day, the reaction mixture was cooled in an ice bath, where $NaNO_2$ (2.8 g, 40 mmol, 2 eq) and TFA (2.56 mL, 33 mmol 1.65 eq) were slowly added. Then, the reaction was heated to 60 °C and stirred for 30 minutes. The solution was cooled to room temperature and washed with 20 % $Na_2S_2O_3$ solution (30 mL) before extracting the reaction mixture with DCM (3x 30 mL) until the water phase was not pink anymore. The organic phases were dried over Na_2SO_4 before the solvent was gently removed under reduced pressure. The crude product was purified by flash chromatography (0-2 % Et_2O/pentane) to yield the product (1.4 g, 48 %) as a volatile red liquid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 2.73 (s, 3H, CH_3), 2.98 (s, 3H, SCH_3).
^{1}H-NMR (300 MHz, DMSO-d_6): δ = 2.71 (s, 3H, CH_3), 2.88 (s, 3H, SCH_3).
^{13}C-NMR (75 MHz, DMSO-d_6): δ = 12.8 (1C, CH_3), 20.3 (1C, SCH_3), 164.9 (1C, $C_{Ar\text{-}3}$), 173.2 (1C, $C_{Ar\text{-}6}$).

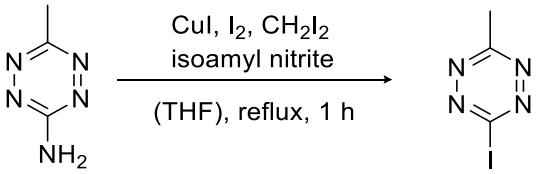
3-Amino-6-methyl-1,2,4,5-tetrazine[180a]



3-*S*-Methylthio-6-methyl-1,2,4,5-tetrazine (1.4 g, 10.1 mmol, 1 eq) was dissolved in 40 mL methanol and 7 M ammonia solution in methanol (8.7 mL, 60.5 mmol, 6 eq) was added to the reaction mixture, which was stirred at room temperature for 48 hours. After the reaction is complete the solvent was evaporated and the crude residue purified by flash chromatography (5-20 % EtOAc/pentane) to obtain the product (0.99 g, 88 %) as an orange solid.

^{1}H-NMR (300 MHz, DMSO-d_6): δ = 2.67 (s, 3H, CH_3), 7.74 (s, 2H, NH_2).
^{13}C-NMR (75 MHz, DMSO-d_6): δ = 19.6 (1C, CH_3), 160.7 (1C, $C_{Ar\text{-}3}$), 163.1 (1C, $C_{Ar\text{-}6}$).

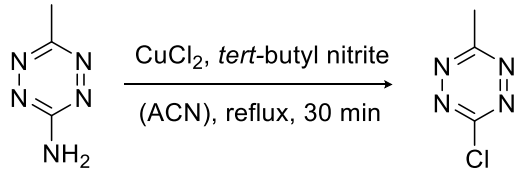
3-Iodo-6-methyl-1,2,4,5-tetrazine[333c]



3-Amino-6-methyl-1,2,4,5-tetrazine (1.45 g, 13.4 mmol, 1 eq), diiodomethane (10.8 mL, 0.134 mol, 10 eq), copper(I) iodide (2.68 g, 14.1 mmol, 1.05 eq), iodine (3.4 g, 13.4 mmol, 1 eq), isoamyl nitrite (5.4 mL, 40.1 mmol, 3 eq) and 22.5 mL THF were combined in a flask and heated to reflux for 1 hour. The reaction was filtered over Celite and the solvents removed under reduced pressure. The crude residue was dissolved in DCM (30 mL) and washed with 20 % $Na_2S_2O_3$ solution (3x 10 mL). The organic phase was dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude product was purified by two consecutive flash chromatographies (1. 0-3 % EtOAc/pentane to remove diiodomethane, 2. 0-3 % EtOAc/pentane to purify the product) to yield the product (1.8 g, 40 %) as a red solid.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 3.00 (s, 3H, CH_3).
^{1}H-NMR (300 MHz, DMSO-d_6): δ = 2.86 (s, 3H, CH_3).
^{13}C-NMR (75 MHz, DMSO-d_6): δ = 20.6 (1C, CH_3), 140.4 (1C, C_{Ar-3}), 167.2 (1C, C_{Ar-6}).
Elementary analysis: $C_3H_3N_4I$ calc: 16% C; 2 % H; 25 % N; 57 % I
found: 17.12% C; 1.48 % H; 24.83 % N; 55.1 % I

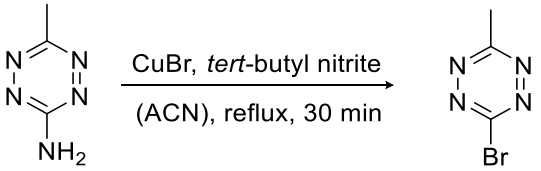
3-Chloro-6-methyl-1,2,4,5-tetrazine[333a]



3-Amino-6-methyl-1,2,4,5-tetrazine (0.2 g, 1.8 mmol, 1 eq), copper(II) chloride (0.29 g, 2.2 mmol, 1.2 eq), *tert*-butyl nitrite (0.32 mL, 2.7 mmol, 1.5 eq) and 5 mL ACN were combined in a flask and heated to reflux for 30 minutes. The reaction was quenched with 20 % aqueous HCl solution and extracted with Et_2O. The organic phase was dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude product was purified by flash chromatography (0-5 % Et_2O/pentane) to yield the product (0.15 g, 64 %) as a red, slightly volatile solid.

^{1}H-NMR (400 MHz, DMSO-d_6): δ = 2.97 (s, 3H, CH_3).
^{13}C-NMR (100 MHz, DMSO-d_6): δ = 20.3 (1C, CH_3), 167.1 (1C, C_{Ar-3}), 167.8 (1C, C_{Ar-6}).
Elementary analysis: $C_3H_3N_4Cl$ calc: 28% C; 2 % H; 43 % N; 27 % Cl
found: 29.37% C; 2.66 % H; 41.4 % N; 26.4 % Cl

3-Bromo-6-methyl-1,2,4,5-tetrazine[333a]

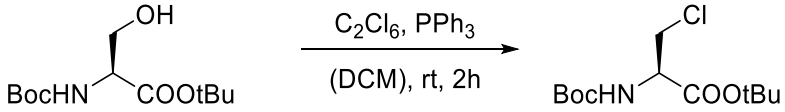


3-Amino-6-methyl-1,2,4,5-tetrazine (0.2 g, 1.8 mmol, 1 eq), copper(I) bromide (0.49 g, 2.2 mmol, 1.2 eq), *tert*-butylnitrite (0.32 mL, 2.7 mmol, 1.5 eq) and 5 mL ACN were combined in a flask and heated to reflux for 30 minutes. The reaction was quenched with 20 % aqueous HCl solution and extracted with Et_2O. The organic phase was dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude product was purified by flash chromatography (0-5 % Et_2O/pentane) to yield the product (0.18 g, 57 %) as a red, slightly volatile solid. The analytical data showed that the product is a mix of 3-bromo and 3-chloro substituted 6-methyl-1,2,4,5,-tetrazine.

^{1}H-NMR (400 MHz, DMSO-d_6): δ = 2.93 (s, 3H, CH_3)
^{13}C-NMR (100 MHz, DMSO-d_6): δ = 20.3 (1C, CH_3), 160.5 (1C, $C_{Ar\text{-}3}$), 167.8 (1C, $C_{Ar\text{-}6}$), 167.8 (1C, $C_{Ar\text{-}6}$).
Elementary analysis: $C_3H_3N_4Br$ calc: 21% C; 1 % H; 32 % N; 46 % Br
found: 31.98 % Br; 8.33 % Cl

Side product:
^{1}H-NMR (400 MHz, DMSO-d_6): δ = 2.97 (s, 3H, CH_3).

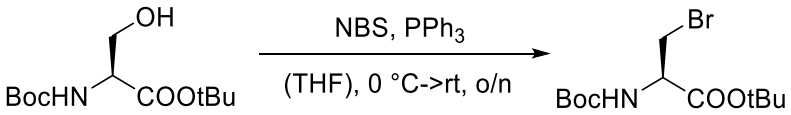
***tert*-butyl (*tert*-butoxycarbonyl)-*ß*-chloro-L-alaninate**[346]



A solution of triphenylphosphine (0.98 g, 3.7 mmol, 1.1 eq) and hexachloroethane (0.89 g, 3.7 mmol, 1.1 eq) in DCM (2.5 mL) was added to a solution of Boc-Ser-OtBu (0.9 g, 3.4 mmol, 1 eq) in DCM (13 mL). The reaction was stirred at room temperature for 2 h, before being quenched with a saturated solution of $NaHCO_3$ (2.5 mL). The organic phase was washed with brine and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (0-5 % EtOAc/pentane) to obtain the product (0.69 g, 73 %) as a white solid.

^{1}H-NMR (500 MHz, DMSO-d_6): δ = 1.39 (s, 9H, Boc), 1.41 (s, 9H, OtBu) 3.76 (dd, $^3J_{HH}$ = 11.2, 7.5 Hz, 1H, CH_2), 3.84 (dd, $^3J_{HH}$ = 11.2, 4.5 Hz, 1H, CH_2), 4.34 (ddd, $^3J_{HH}$ = 7.5 Hz, 4.5 Hz, 1H, $CH_{\text{-}\alpha}$), 7.27 (d, $^3J_{HH}$ = 8.2 Hz, 1H, $NH_{\text{-}\alpha}$).

***tert*-butyl (*tert*-butoxycarbonyl)-*ß*-bromo-L-alaninate[346]**

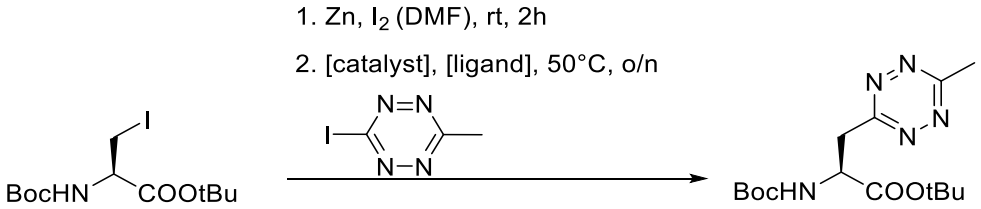


Triphenylphosphine (1.18 g, 4.5 mmol, 1.3 eq) was added to a solution of Boc-Ser-OtBu (0.9 g, 3.4 mmol, 1 eq) in THF (10 mL) and cooled to 0 °C. Then *N*-bromosuccinimide (3.2 g, 12.7 mmol, 1.3 eq) was slowly added in small portions over 30 minutes and the resulting reaction mixture was stirred at room temperature overnight. The white precipitate was filtered off and the filtrate concentrated under reduced pressure. The crude residue was purified by flash chromatography (0-5 % EtOAc/pentane) to obtain the product (0.69 g, 63 %) as a white solid.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 1.46 (s, 9H, Boc), 1.50 (s, 9H, OtBu) 3.70 (dd, $^3J_{HH}$ = 10.3, 3.2 Hz, 1H, CH_2), 3.80 (dd, $^3J_{HH}$ = 10.3, 3.2 Hz, 1H, CH_2), 4.34 (ddd, $^3J_{HH}$ = 7.7 Hz, 3.2 Hz, 3.2 Hz, 1H, $CH_{-\alpha}$), 5.39 (d, $^3J_{HH}$ = 7.7 Hz, 1H, $NH_{-\alpha}$).
MS (ESI), *m/z* calc for $C_{12}H_{22}BrNO_4$ 323.07, found 346.8, 347.8 $[M+Na]^+$

1.5.3.3.2 Negishi cross-coupling reactions

General procedure for Negishi cross-coupling reaction with various catalysts and ligands:



The reaction was performed under Schlenk conditions. DMF (1.5 mL/mmol) and iodine (0.15 eq) were added to a 3-neck flask containing zinc dust (1.5 eq) and stirred vigorously (~750 rpm) for 15 minutes at room temperature. After the solution turned from orange to grey Boc-*β*-iodo-Ala-OtBu (1 eq) and iodine (0.15 eq) were added and the reaction vigorously stirred (~750 rpm) for 2 hours at room temperature. Then, 3-iodo-6-methyl-1,2,4,5-tetrazine (1.3 eq), catalyst (0.05 eq) and ligand (0.01 eq) were added and the resulting reaction mixture heated to 50 °C and vigorously (~750 rpm) stirred overnight. The reaction was filtered over Celite and the solvent evaporated under reduced pressure before purifying the crude product by flash chromatography (0-10 % EtOAc/pentane) to obtain the product as a red oil.

***tert*-butyl (*tert*-butoxycarbonyl)-*ß*-3-(6-methyl-1,2,4,5-tetrazin-3-yl)-*L*-alaninate (BocTetAOtBu) - Condition A[332]**

The reaction was performed under Schlenk conditions. 3 mL DMF and iodine (19 mg, 0.08 mmol, 0.15 eq) were added to a 3-neck flask containing zinc dust (98 mg, 1.5 mmol, 1.5 eq) and stirred vigorously (~750 rpm) for 15 minutes at room temperature. After the solution turned from orange to grey Boc-*β*-Iodo-Ala-OtBu (0.19 g, 0.5 mmol, 1 eq) and iodine (19 mg, 0.08 mmol, 0.15 eq) were added and the reaction vigorously stirred (~750 rpm) for 2 hours at room temperature. Then, 3-iodo-6-methyl-1,2,4,5-tetrazine (0.15 g, 0.65 mmol, 1.3 eq), $Pd(OAc)_2$ (6 mg, 0.026 mmol, 0.05 eq) and SPhos (21 mg, 0.05 mmol, 0.1 eq) were added and the resulting reaction mixture heated to 50 °C and vigorously (~750 rpm) stirred overnight. The reaction was filtered over Celite and the solvent evaporated under reduced pressure before purifying the crude product by flash chromatography (0-10 % EtOAc/pentane) to obtain the product (64 mg, 37 %) as a red oil.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 1.38 (s, 9H, Boc), 1.42 (s, 9H, OtBu), 3.04 (s, 3H, CH_3), 3.71 (dd, $^3J_{HH}$ = 14.6, 5.1 Hz, 1H, CH_2), 3.82 (dd, $^3J_{HH}$ = 14.6, 6.9 Hz, 1H, CH_2), 4.70-4.77 (m, 1H, CH_α), 5.43 (d, $^3J_{HH}$ = 7.9 Hz, 1H, NH_α).
MS (ESI), *m/z* calc for $C_{15}H_{25}N_5O_4$ 339.19, found 362.2 $[M+Na]^+$

Condition B

The reaction was performed with $Pd(OAc)_2$ (6 mg, 0.013 mmol, 0.05 eq) and SPhos (23 mg, 0.1 mmol, 0.2 eq) at 75°C and yielded the product (61 mg, 32 %) as a red oil.

MS (ESI), *m/z* calc for $C_{15}H_{25}N_5O_4$ 339.19, found 362.2 $[M+Na]^+$

Condition C[329a]

The reaction was performed with $Pd_2(dba)_3$ (12 mg, 0.0125 mmol, 0.025 eq) and tri(o-tolyl)phospine (15 mg, 0.05 mmol, 0.1 eq) and yielded the product (17 mg, 10 %) as a red oil.

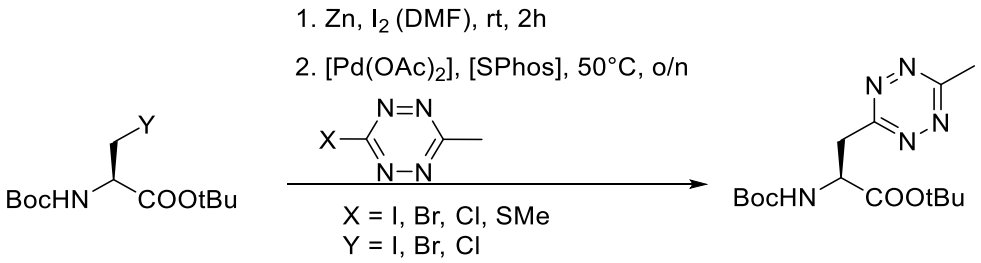
MS (ESI), *m/z* calc for $C_{15}H_{25}N_5O_4$ 339.19, found 362.2 $[M+Na]^+$

Condition D[336]

The reaction was performed with $Pd(OAc)_2$ (6 mg, 0.013 mmol, 0.05 eq) and tri(2-furyl)phosphine (23 mg, 0.1 mmol, 0.2 eq) and yielded the product (0.11 g, 60 %) as a red oil.

MS (ESI), *m/z* calc for $C_{15}H_{25}N_5O_4$ 339.19, found 362.2 $[M+Na]^+$

General procedure Negishi cross-coupling reaction with varying substituents on starting materials:



The reaction was performed under Schlenk conditions. DMF (1.5 mL/mmol) and iodine (0.15 eq) were added to a 3-neck flask containing zinc dust (1.5 eq) and stirred vigorously (~750 rpm) for 15 minutes at room temperature. After the solution turned from orange to grey Boc-*β*-X-Ala-OtBu (1 eq) and iodine (0.15 eq) were added and the reaction vigorously stirred (~750 rpm) for 2 hours at room temperature. Then, 3-X-6-methyl-1,2,4,5-tetrazine (1.3 eq), $Pd(OAc)_2$ (0.05 eq) and SPhos (0.01 eq) were added and the resulting reaction mixture heated to 50 °C and vigorously (~750 rpm) stirred overnight. The reaction was filtered over Celite and the solvent evaporated under reduced pressure before purifying the crude product by flash chromatography (0-10 % EtOAc/pentane) to obtain the product as a red oil.

Condition E[334]

The reaction was performed with Boc-*β*-iodo-Ala-OtBu (0.19 g, 0.5 mmol, 1 eq) and 3-S-methyl-6-methyl-1,2,4,5-tetrazine (93 mg, 0.65 mmol, 1.3 eq) at room temperature and 50°C and did not yield any product.

Condition F

The reaction was performed with Boc-*β*-bromo-Ala-OtBu (0.16 g, 0.5 mmol, 1 eq) and 3-iodo-6-methyl-1,2,4,5-tetrazine (0.15 g, 0.65 mmol, 1.3 eq) at 50 °C and did not yield any product.

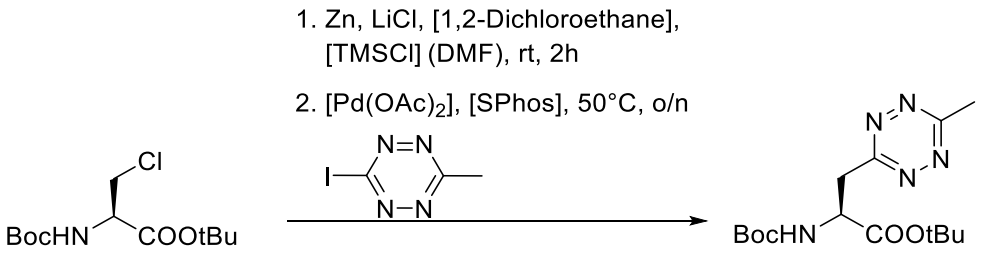
Condition G

The reaction was performed with Boc-*β*-bromo-Ala-OtBu (0.1 g, 0.3 mmol, 1 eq) and 3-iodo-6-methyl-1,2,4,5-tetrazine (97 mg, 0.4 mmol, 1.3 eq) at 80 °C (even zinc insertion) and did not yield any product.

Condition H

The reaction was performed with Boc-*β*-chloro-Ala-OtBu (0.16 g, 0.5 mmol, 1 eq) and 3-iodo-6-methyl-1,2,4,5-tetrazine (0.15 g, 0.65 mmol, 1.3 eq) at 50 °C and did not yield any product.

Condition I[328]



The reaction was performed under Schlenk conditions. Zinc dust (35 mg, 0.54 mmol, 1.5 eq) and 1 mL DMF were added to a heated out flask already containing LiCl (23 mg, 0.54 mmol, 1.5 eq). Then, 1,2-dichloroethane (2 µL, 18 µmol, 0.05 eq) was added and the reaction heated to ebullition. After cooling trimethylsilyl chloride (0.5 µL, 3.6 µmol, 0.01 eq) was added and heated to ebullition. Then, Boc-*β*-chloro-Ala-OtBu (0.1 g, 0.36 mmol, l eq) was added and the reaction heated to 50 °C and stirred overnight before 3-iodo-6-methyl-1,2,4,5-tetrazine (0.1 g, 0.46 mmol, 1.3 eq), $Pd(OAc)_2$ (4 mg, 0.017 mmol, 0.05 eq) and SPhos (15 mg, 0.035 mmol, 0.1 eq) were added and stirred at 50 °C overnight. The reaction did not yield any product.

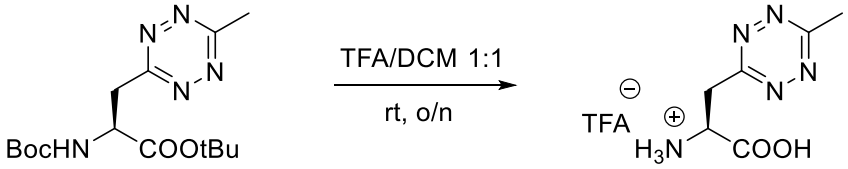
Condition J

The reaction was performed with Boc-*β*-iodo-Ala-OtBu (0.19 g, 0.5 mmol, 1 eq) and 3-chloro-6-methyl-1,2,4,5-tetrazine (83 mg, 0.65 mmol, 1.3 eq) at 50 °C and did yield the product in lower than 10 % yield.

Condition K

The reaction was performed with Boc-*β*-iodo-Ala-OtBu (0.19 g, 0.5 mmol, 1 eq) and 3-bromo-6-methyl-1,2,4,5-tetrazine (0.11 g, 0.65 mmol, 1.3 eq) at 50 °C and did yield the product (25 mg, 14 %) as a red oil.

***ß*-3-(6-methyl-1,2,4,5-tetrazin-3-yl)-L-alanine trifluoroacetate (TetA)**



BocTetAOtBu (0.15 g, 0.32 mmol, 1 eq) was dissolved in 2 mL DCM, cooled to 0°C before addition of 2 mL TFA and a drop of H_2O. The reaction mixture was then stirred at room temperature overnight. The solvent was evaporated under reduced pressure, the product dissolved in a minimal amount of MeOH and then precipitated from cold Et_2O and subsequently dried to obtain the product (0.13 g, 95 %) as a pink solid.

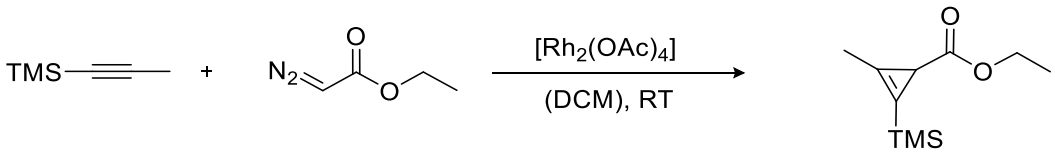
^{1}H-NMR (400 MHz, DMSO-d_6): δ = 3.01 (s, 3H, CH_3), 3.78 (dd, $^3J_{HH}$ = 15.7, 7.0 Hz, 1H, CH_2), 3.84 (dd, $^3J_{HH}$ = 15.7, 6.0 Hz, 1H, CH_2), 4.60-4.69 (m, 1H, $CH_{-\alpha}$), 8.42 (s, 3H, $NH_{-\alpha}$).

^{13}C-NMR (75 MHz, DMSO-d_6 + TFA): δ = 20.8 (1C, CH_3), 34.8 (1C, CH_2), 50.6 (1C, C_α), 165.4 (1C, C_{Tet-6}), 167.6 (1C, C_{Tet-3}), 169.7 (1C, COOH).
MS (ESI), *m/z* calc for $C_6H_9N_5O_2$ 183.08, found 183.99 $[M+H]^+$

2.5.4 Synthesis of CpK

2.5.4.1 Synthesis of CpK starting from TMS-propyne[192, 196, 198b]

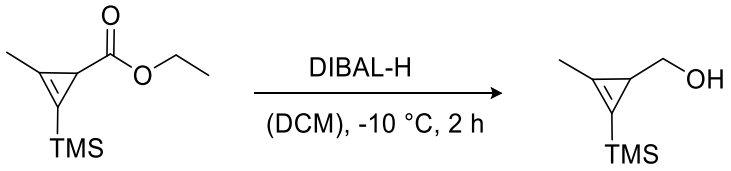
Ethyl 2-methyl-3-(trimethylsilyl)cycloprop-2-ene-1-carboxylate



TMS-propyne (9.24 mL, 62.4 mmol, 2 eq) was added to a suspension of rhodium(II) acetate (70 mg, 0.156 mmol, 0.005 eq) in 15 mL dry DCM. Then, ethyl diazoacetate (3.28 mL, 32.2 mmol, 1 eq) was slowly added at room temperature at a speed of 0.75 mL/h using a syringe pump. The resulting reaction mixture was stirred for 2 hours after addition of ethyl diazoacetate was complete. After solvent removal under reduced pressure the crude product was purified by flash column chromatography (eluding with 2-5% Et_2O in pentane) to result the product (4.44 g, 72%) as a colourless, oily residue.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 0.18 (s, 9H, $Si(CH_3)_3$), 1.23 (t, $^3J_{HH}$ = 7.1 Hz, 3H, $CH_{3\text{-}Ethoxy}$), 1.97 (s, 1H, CH), 2.19 (s, 3H, CH_3), 4.10 (m, 2H, CH_2).
MS (ESI), *m/z* calc for $C_{10}H_{18}O_2Si$ 198.11, found 221.2 $[M+Na]^+$

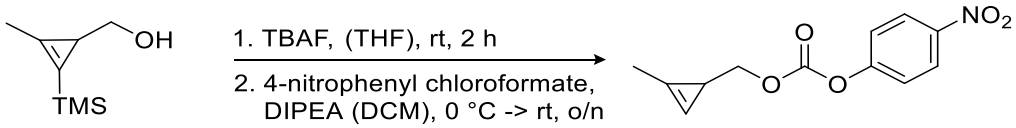
(2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methanol



Ethyl-2-methyl-3-(trimethylsilyl)cycloprop-2-ene-1-carboxylate (1.91 g, 9.6 mmol, 1 eq) was dissolved in 19 mL dry dichloromethane, then placed in a cooling bath made from ice and sodium chloride. 1 M DIBAL-H solution in DCM (19.2 mL, 19.2 mmol, 2 eq) was added dropwise via syringe over 45 min, followed by stirring the reaction mixture for 45 min to 1 h before quenching it with 1 M NaOH (3-5 mL) and water (1 mL). After stirring the reaction mixture for 1 to 2 hours, the emulsion broke down and a white solid formed, which was filtered off. The organic phase was dried over Na_2SO_4 and the solvent evaporated under reduced pressure. The crude product was purified the same day by flash column chromatography (gradient 4-20% Et_2O in pentane) to result the product (1.13 g, 75%) as a colourless, oily residue.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 0.20 (s, 9H, $Si(CH_3)_3$), 1.59 (t, $^3J_{HH}$ = 4.6 Hz, 1H, CH), 2.24 (s, 3H, CH_3), 3.51 (d, $^3J_{HH}$ = 4.6 Hz, 2H, CH_2).

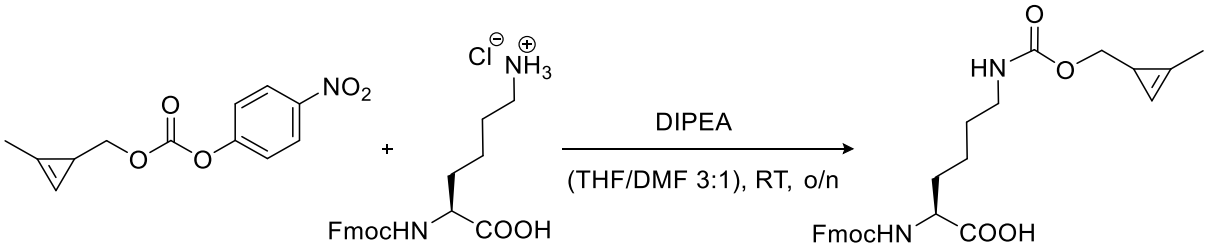
(2-methylcycloprop-2-en-1-yl)methyl (4-nitrophenyl) carbonate



1.4 M Tetrabutylammonium fluoride solution in THF (9.8 mL, 9.8 mmol, 1.16 eq) was added to a solution of 2-methyl-3-(trimethylsilyl)cycloprop-2-ene-1-yl)methanol (1.3 g, 8.5 mmol, 1 eq) in 19 mL THF and stirred at room temperature for 2 hours. The reaction mixture was directly purified by flash chromatography (0-100% Et_2O/pentane). The product was concentrated at 20 °C under reduced pressure and then directly dissolved in DCM (15 mL), followed by addition DIPEA (2.2 mL, 12.7 mmol, 1.5 eq) and 4-nitrophenyl chloroformate (2.1 g, 10.2 mmol, 1.2 eq) in small portions. The reaction mixture was stirred at room temperature for 2 h, then the solvent removed and the crude product was purified by flash column chromatography (0-5 % Et_2O/pentane) to result the product (0.95 g, 45 %) as a colourless, oily residue.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.78 (t, $^3J_{HH}$ = 5.3, 1.5 Hz, 1H, $C(sp^3)H$), 2.18 (d, $^4J_{HH}$ = 1.2 Hz, 3H, CH_3), 4.15 (dd, $^2J_{HH}$ = 10.8 Hz, $^3J_{HH}$ = 5.4 Hz, 1H, CH_2), 4.21 (ddd, $^2J_{HH}$ = 10.8 Hz, $^3J_{HH}$ = 5.2 Hz, $^4J_{HH}$ = 0.5 Hz, 1H, CH_2), 6.60-6.63 (m, 1H, $C(sp^2)H$), 7.39 (d, $^3J_{HH}$ = 9.3 Hz, 2H, $CH_{Ar\text{-}2,2'}$), 8.28 (d, $^3J_{HH}$ = 9.3 Hz, 2H, $CH_{Ar\text{-}3,3'}$).

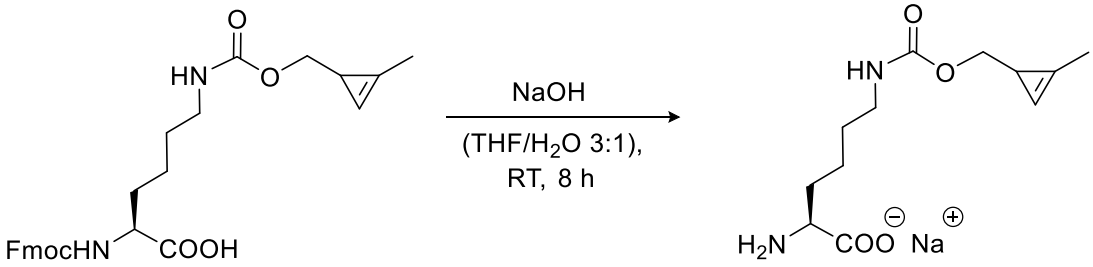
***N*^{2}-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-*N*^{6}-(((2-methylcycloprop-2-en-1-yl)methoxy)-carbonyl)-L-lysine (FmocCpK)**



Fmoc-Lys-OH*HCl (5.5 g, 13.5 mmol, 1.5 eq) was dissolved in 33 mL THF/DMF (3:1), then DIPEA (7.5 mL, 44.7 mmol, 5 eq) and a solution of (2-methyl-cycloprop-2-en-1-yl)methyl(4-nitrophenyl)carbonate (2.62 g, 9 mmol, 1 eq) in 5 mL THF were added and the reaction mixture was stirred at room temperature overnight. The organic phase was washed with 10 % citric acid (2-3x 50 mL) and brine (2x 50 mL) and then dried over Na_2SO_4 and the solvent evaporated under reduced pressure. The crude product was purified by flash column chromatography (50-100 % EtOAc/pentane + 0.1 % AcOH). Coevaporatation of acetic acid with toluene resulted the product (3.12 g, 73 %) as a slightly yellow solid.

^{1}H-NMR (400 MHz, DMSO-d_6): δ = 1.23-1.44 (m, 4H, $CH_{2\text{-}\gamma,\delta}$), 1.47-1.50 (m, 1H, C($sp^3$)H), 1.54-1.74 (m, 2H, $CH_{2\text{-}\beta}$), 2.09 (d, $^{4}J_{HH}$ = 1.1 Hz, 3H, CH_3), 2.91-2.99 (m, 2H, $CH_{2\text{-}\varepsilon}$), 3.75 (dd, $^{3}J_{HH}$ = 11.1, 5.2 Hz, 1H, OCH_2), 3.81 (dd, $^{3}J_{HH}$ = 11.1, 4.8 Hz, 1H, OCH_2), 3.85-3.95 (m, 1H, CH_α), 4.19-4.25 (m, 1H, CH_{Fmoc}), 4.25-4.32 (m, 2H, $CH_{2\text{-}Fmoc}$), 6.84 (s, 1H, C(sp^2)H), 7.04 (t, $^{3}J_{HH}$ = 5.7 Hz, 1H, NH_ε), 7.33 (dd, $^{3}J_{HH}$ = 7.4, 7.0 Hz, 2H, $CH_{Ar(Fmoc\text{-}2,2')}$), 7.41 (dd, $^{3}J_{HH}$ = 7.4, 7.3 Hz, 2H, $CH_{Ar(Fmoc\text{-}3,3')}$), 7.59 (d, $^{3}J_{HH}$ = 7.9 Hz, 1H, NH_α), 7.73 (d, $^{3}J_{HH}$ = 7.5 Hz, 2H, $CH_{Ar(Fmoc\text{-}1,1')}$), 7.89 (d, $^{3}J_{HH}$ = 7.5 Hz, 2H, $CH_{Ar(Fmoc\text{-}4,4')}$), 12.56 (bs, 1H, COOH).
^{13}C-NMR (75 MHz, DMSO-d_6): δ = 11.4 (1C, CH_3), 16.8 (1C, C(sp^3)H), 22.9 (1C, C_γ), 29.1 (1C, C_δ), 30.5 (1C, C_β), 39.9 (1C, C_ε), 46.7 (1C, CH_{Fmoc}), 53.8 (1C, C_α), 65.6 (1C, $CH_{2\text{-}Fmoc}$), 70.7 (1C, OCH_2), 102.3 (1C, C(sp^2)H), 120.1 (2C, $C_{Ar(Fmoc\text{-}4,4')}$), 125.3 ($C_{Ar(Fmoc\text{-}1,1')}$), 127.0 (2C, $C_{Ar(Fmoc\text{-}2,2')}$), 127.6 (2C, $C_{Ar(Fmoc\text{-}3,3')}$), 140.7 (2C, $C_{Ar(Fmoc\text{-}5,5')}$), 143.8 (2C, $C_{Ar(Fmoc\text{-}6,6')}$), 156.1 (1C, $N_\alpha COO$), 156.4 (1C, $N_\varepsilon COO$), 174.0 (1C, COOH).
MS (ESI), *m/z* calc for $C_{27}H_{30}N_2O_6$ 478.21, found 501.2 $[M+Na]^+$, 477.2 $[M-H]^-$

N^6-(((2-methylcycloprop-2-en-1-yl)methoxy)carbonyl)-L-lysine (CpK)

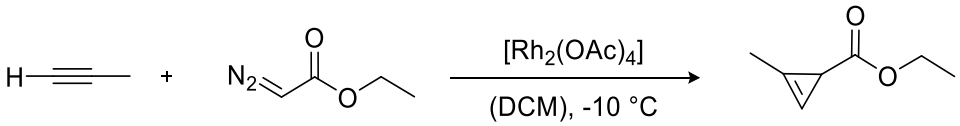


N^2-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-N^6-(((2-methylcycloprop-2-en-1-yl)methoxy)-carbonyl)-L-lysine (3 g, 6.26 mmol, 1 eq) was dissolved in 36 mL THF/H_2O (3:1), followed by addition of solid sodium hydroxide (0.78 g, 19.4 mmol, 3.1 eq). The reaction mixture was stirred at room temperature for 8 hours before adding water (50 mL) and washing the water phase with Et_2O (5x 50 mL). The water phase was concentrated and the resulting solid was dissolved in 100 mM NaOH to make a 100 mM stock solution, which is stored at -20 °C.

^{1}H-NMR (400 MHz, D_2O): δ = 1.32-1.49 (m, 2H, $CH_{2\text{-}\gamma}$), 1.49-1.60 (m, 2H, $CH_{2\text{-}\delta}$), 1.60 (t, $^{3}J_{HH}$ = 4.9 Hz, 1H, C(sp^3)H), 1.83-1.96 (m, 2H, $CH_{2\text{-}\beta}$), 2.13 (d, $^{4}J_{HH}$ = 1.2 Hz, 3H, CH_3), 3.15 (t, $^{3}J_{HH}$ = 6.8 Hz, 2H, $CH_{2\text{-}\varepsilon}$), 3.82 (dd, $^{3}J_{HH}$ = 11.0, 5.0 Hz, 1H, OCH_2), 4.01 (dd, $^{3}J_{HH}$ = 11.0, 4.8 Hz, 1H, OCH_2), 3.75 (dd, $^{3}J_{HH}$ = 5.9, 4.9 Hz, 1H, CH_α), 6.66 (s, 1H, C(sp^2)H).
^{13}C-NMR (75 MHz, D_2O): δ = 10.8 (1C, CH_3), 16.6 (1C, C(sp^3)H), 21.7 (1C, C_γ), 28.6 (1C, C_δ), 30.1 (1C, C_β), 40.0 (1C, C_ε), 54.8 (1C, C_α), 72.5 (1C, OCH_2), 101.2 (1C, C(sp^2)H), 120.5 (1C, $N_\varepsilon COO$), 174.8 (1C, COOH).
MS (ESI), *m/z* calc for $C_{12}H_{20}N_2O_4$ 256.14, found 257.2 $[M+H]^+$, 279.2 $[M+Na]^{+-}$

2.5.4.2 Synthesis of CpK starting from propyne gas[198b]

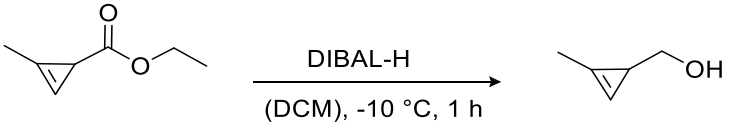
Ethyl 2-methylcycloprop-2-ene-1-carboxylate



A 100 mL 3-neck round bottom flask was fitted with a dry ice condenser and a pressure release valve and charged with rhodium acetate (80 mg, 0.18 mmol, 0.006 eq). The glass assembly was evacuated and heated out, followed by filling three times with argon. Afterwards dry DCM (2 mL) was added and the condenser filled with a mixture of acetone and dry ice. Propyne (approx. 15 mL) was condensed into the rhodium(II) acetate suspension and then the flask lowered into a salt-ice bath to obtain a steady reflux of propyne. Ethyl diazoacetate (3.3 mL, 31.2 mmol, 1 eq) was added to the stirred propyne solution drop-wise using a syringe pump (1.1 mL/min) and was stirred for 2 h more. After completion, the reaction was slowly warmed up to evaporate residual propyne. The cyclopropene product was then purified by flash chromatography (0-2 % Et_2O/pentane). The eluate was carefully concentrated (not to dryness!) at 20 °C to give the desired product as a colourless volatile liquid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.25 (t, $^3J_{HH}$ = 7.2 Hz, 3H, $CH_{3\text{-Ethoxy}}$), 2.11 (d, $^4J_{HH}$ = 1.5 Hz, 1H, C(sp^3)H), 2.16 (d, $^4J_{HH}$ = 1.3 Hz, 3H, CH_3), 4.13 (dq, $^3J_{HH}$ = 7.1 Hz, $^3J_{HH}$ = 2.6 Hz, 2H, CH_2), 6.32-6.35 (m, 1H, C(sp^2)H).

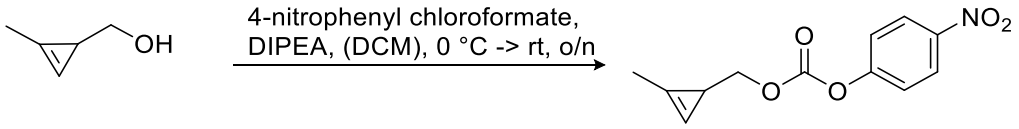
(2-methylcycloprop-2-en-1-yl)methanol



A 1 M solution of DIBAL-H in DCM (30 mL, 30 mmol, approx. 1.5 eq) was added drop-wise to a stirred solution of ethyl 2-methylcycloprop-2-ene-1-carboxylate (approx. 2.6 g, approx. 20 mmol, 1 eq) in DCM (30 mL) at -10 ºC for 1 h. Due to residual starting material present in the mixture more DIBAL-H (8 mL, 8 mmol, 0.4 eq) in DCM (8 mL) was added and the reaction stirred for another 30 minutes, before being quenched by cautious addition of H_2O (1 mL), 1 M aqueous NaOH solution (1 mL) and H_2O (2.3 mL). The mixture was stirred for a further 2 h at room temperature until the emulsion broke down, before the filtrate was dried over Na_2SO_4, filtered and purified by flash chromatography (50 % Et_2O/pentane). The eluate was carefully concentrated (not to dryness!) at 20 °C to give the desired product as a colourless volatile liquid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.67 (dt, $^3J_{HH}$ = 4.3, 1.6 Hz, 1H, C(sp^3)H), 2.16 (d, $^4J_{HH}$ = 1.2 Hz, 3H, CH_3), 3.50 (ddd, $^2J_{HH}$ = 10.6 Hz, $^3J_{HH}$ = 4.5 Hz, $^4J_{HH}$ = 0.7 Hz, 1H, CH_2), 3.57 (ddd, $^2J_{HH}$ = 10.6 Hz, $^3J_{HH}$ = 4.2 Hz, $^4J_{HH}$ = 0.6 Hz, 1H, CH_2), 6.62-6.64 (m, 1H, C(sp^2)H).

(2-methylcycloprop-2-en-1-yl)methyl (4-nitrophenyl) carbonate



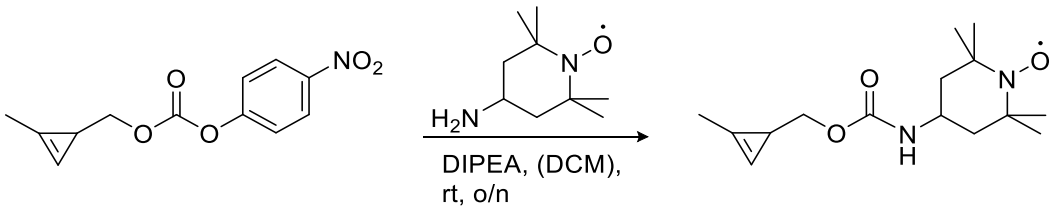
(2-methylcycloprop-2-en-1-yl)methanol (approx. 2.4 g, approx. 30 mmol, 1 eq) was dissolved in 40 mL DCM, then DIPEA (7 mL, 40 mmol, approx. 1.5 eq) and 4-nitrophenyl chloroformate (6.4 g, 32 mmol, approx. 1.2 eq) were added on ice and the reaction stirred at room temperature overnight. The reaction was concentrated under reduced pressure and the crude product purified by flash chromatography (0-5 % Et_2O/pentane) to obtain the product (2.5 g, 32 % over three steps) as a colourless oil.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 1.78 (dt, $^3J_{HH}$ = 5.3, 1.5 Hz, 1H, C(sp^3)H), 2.18 (d, $^4J_{HH}$ = 1.1 Hz, 3H, CH_3), 4.15 (dd, $^2J_{HH}$ = 10.9 Hz, $^3J_{HH}$ = 5.5 Hz, 1H, CH_2), 4.21 (ddd, $^2J_{HH}$ = 10.9 Hz, $^3J_{HH}$ = 5.2 Hz, $^4J_{HH}$ = 0.6 Hz, 1H, CH_2), 6.60-6.63 (m, 1H, C(sp^2)H), 7.39 (d, $^3J_{HH}$ = 9.3 Hz, 2H, $CH_{Ar\text{-}2,2'}$), 8.28 (d, $^3J_{HH}$ = 9.3 Hz, 2H, $CH_{Ar\text{-}3,3'}$).

The following steps are equal to the ones described in method 1.6.3.1

2.5.5 Synthesis of Cp-Spin labels

2.5.5.1 (*N*-((2-methylcycloprop-2-en-1-yl)methyl))-2,2,5,5-tetramethyl-1-pyrrolidinyl-oxyl–3-carboxamide (Cp-TEMPO)

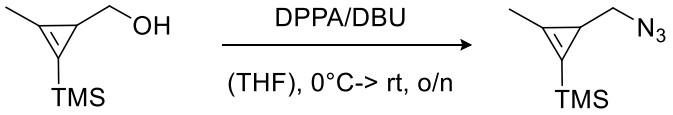


(2-methylcycloprop-2-en-1-yl)methyl-(4-nitrophenyl)-carbonate (0.1 g, 0.4 mmol, 1 eq) was dissolved in 2 mL DCM, then 4-amino-TEMPO (83 mg, 0.5 mmol, 1.2 eq) and DIPEA (84 µL, 0.5 mmol, 1.2 eq) were added and the reaction stirred at room temperature overnight. After evaporation of the solvent the crude residue was purified by flash chromatography (0-10 % EtOAc/pentane) to obtain the product (68 mg, 60 %) as a yellowish oil.

MS (ESI), *m/z* calc for $C_{15}H_{25}N_2O_3$ 281.19, found 281.2 $[M+H]^+$

2.5.5.2 (*N*-((2-methylcycloprop-2-en-1-yl)methyl))-2,2,5,5-tetramethyl-1-pyrrolidinyl-oxyl-3-carboxamide (Cp-Proxyl)

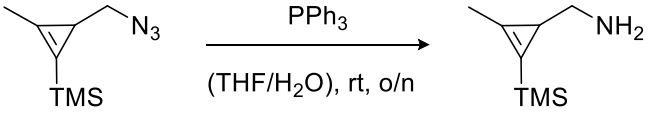
(3-(azidomethyl)-2-methyl-1-(trimethylsilyl))cycloprop-1-ene[158]



(2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methanol (0.37 g, 2.4 mmol, 1 eq) was dissolved in 6 mL THF, cooled to 0 °C, then DBU (0.46 mL, 3.1 mmol, 1.2 eq) and diphenylphosphoryl azide (DPPA) (0.66 mL, 3.1 mmol, 1.2 eq) were added, the reaction slowly warmed to room temperature and stirred overnight. THF was carefully evaporated in an air stream and the crude product directly purified by flash chromatography (100 % pentane). The eluate was carefully concentrated (not to dryness!) at 22 °C and 400-500 mbar to give the desired product as a colourless volatile liquid.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 0.18 (s, 9H, $SiCH_3$), 1.57 (dd, $^3J_{HH}$ = 5.7, 5.0 Hz, 1H, CH), 2.22 (s, 3H, CH_3), 3.02 (dd, $^2J_{HH}$ = 12.6 Hz, $^3J_{HH}$ = 4.9 Hz, 1H, CH_2), 3.15 (dd, $^2J_{HH}$ = 12.6 Hz, $^3J_{HH}$ = 4.9 Hz, 1H, CH_2).

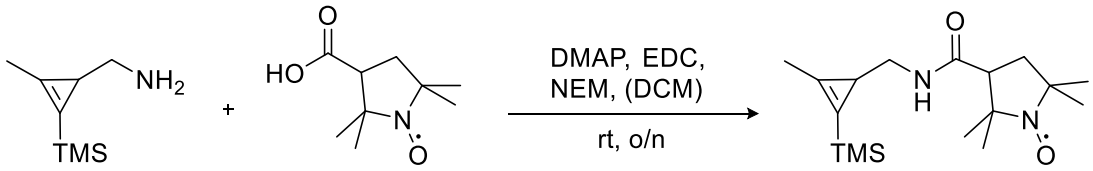
(2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methanamine[158]



(3-(azidomethyl)-2-methyl-1-(trimethylsilyl))cycloprop-1-ene was dissolved in 4 mL THF/H_2O (5.5:1), then triphenylphosphine (0.8 mL, 3.1 mmol, 1.3 eq) was added and the reaction stirred at room temperature overnight. 5 mL 1 M HCl solution was added and the water phase washed with Et_2O (3x 5 mL). The water phase was then basified by addition of 1 M NaOH solution to pH 9 and the product extracted with DCM (3x 10 mL). The organic phase was dried over Na_2SO_4, followed by careful concentration at 20 °C and 100 mbar (minimum!) to give the desired product (72 mg, 20 % over two steps) as a colourless solid.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 0.16 (s, 9H, $SiCH_3$), 1.53 (dd, $^3J_{HH}$ = 5.7, 4.1 Hz, 1H, CH), 2.26 (s, 3H, CH_3), 2.53 (dd, $^2J_{HH}$ = 12.7 Hz, $^3J_{HH}$ = 5.7 Hz, 1H, CH_2), 2.91 (dd, $^2J_{HH}$ = 12.7 Hz, $^3J_{HH}$ = 4.2 Hz, 1H, CH_2).

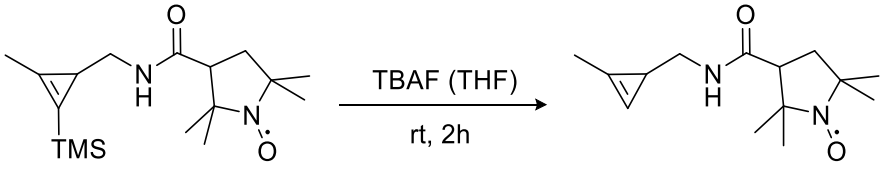
(*N*-((2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methyl))-2,2,5,5-tetramethyl-1-pyrrolidinyl-oxyl–3-carboxamide



3-Carboxy-PROXYL (79 mg, 0.42 mmol, 1 eq) was dissolved in 2 mL DCM and cooled in an ice bath, before DMAP (26 mg, 0.21 mmol, 0.5 eq), EDC*HCl (121 mg, 0.63 mmol, 1.5 eq), NEM (0.16 mL, 0.84 mmol, 2 eq) and (2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methanamine (72 mg, 0.46 mmol, 1.1 eq) were subsequently added and the reaction was stirred overnight at room temperature. After removal of the solvent under reduced pressure the crude product was purified by flash chromatography (0-10 % EtOAc/pentane) to yield the product (76 mg, 56 %) as a yellow solid.

MS (ESI), *m/z* calc for $C_{17}H_{31}N_2O_2Si$ 323.22, found 323.9 $[M+H]^+$

(*N*-((2-methylcycloprop-2-en-1-yl)methyl))-2,2,5,5-tetramethyl-1-pyrrolidinyl-oxyl–3-carboxamide (Cp-Proxyl)

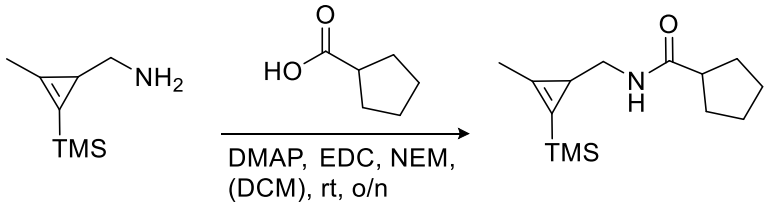


(*N*-((2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methyl))-2,2,5,5-tetramethyl-1-pyrrolidinyl-oxyl-3-carboxamide (76 mg, 0.23 mmol, 1 eq) was dissolved in 0.5 mL THF, followed by addition of 1 M TBAF solution in THF (0.27 mL, 0.27 mmol, 1.16 eq). The resulting reaction mixture was stirred at room temperature for 2 h. The solvent was then removed under reduced pressure and the crude residue purified by flash chromatography (50 % Et_2O/pentane) to obtain the product (57 mg, 98 %) as yellow solid.

MS (ESI), *m/z* calc for $C_{14}H_{23}N_2O_2$ 251.18, found 251.8 $[M+H]^+$, 273.8 $[M+Na]^+$

2.5.5.3 *N*-((2-methylcycloprop-2-en-1-yl)methyl)cyclopentanecarboxamide (Cp-SLA)

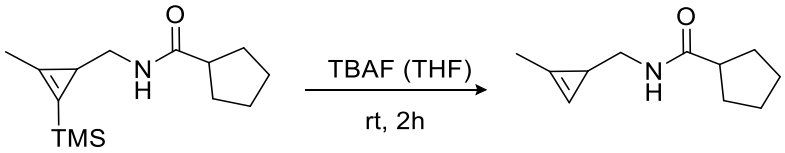
***N*-((2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methyl)cyclopentanecarboxamide**



Cyclopentanecarboxylic acid (0.13 mL, 1.2 mmol, 1 eq) was dissolved in 6 mL DCM and cooled in an ice bath, before DMAP (72 mg, 0.6 mmol, 0.5 eq), EDC*HCl (0.34 g, 1.8 mmol, 1.5 eq), NEM (0.28 mL, 2.4 mmol, 2 eq) and (2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methanamine (0.2 g, 1.3 mmol, 1.1 eq) were subsequently added and the reaction was stirred overnight at room temperature. After removal of the solvent under reduced pressure the crude product was purified by flash chromatography (0-10 % EtOAc/pentane) to yield the product (0.24 g, 74 %) as a yellow solid.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 0.16 (s, 9H, $SiCH_3$), 1.44 (t, $^3J_{HH}$ = 4.5 Hz, 1H, $CH_{cyclopropene}$), 1.51-1.61 (m, 2H, $CH_{2cyclopentane}$), 1.68-1.88 (m, 6H, 3x $CH_{2cyclopentane}$), 2.18 (s, 3H, CH_3), 2.42-2.52 (m, 1H, $CH_{cyclopentane}$), 3.15 (dd, $^3J_{HH}$ = 4.9, 4.8 Hz, 2H, CH_2), 5.26 (bs, 1H, NH).
MS (ESI), *m/z* calc for $C_{14}H_{25}NOSi$ 251.17, found 252.2 $[M+H]^+$

***N*-((2-methylcycloprop-2-en-1-yl)methyl)cyclopentanecarboxamide (Cp-SLA)**



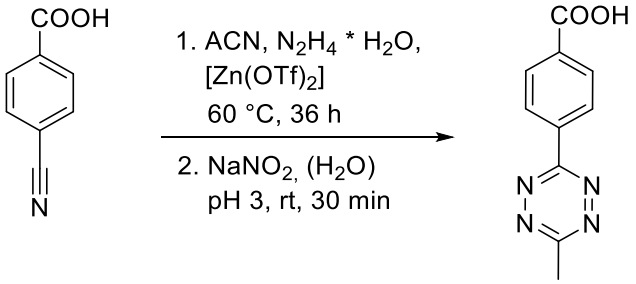
N-((2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methyl)cyclopentanecarboxamide (0.24 g, 0.95 mmol, 1 eq) was dissolved in 3 mL THF, followed by addition of 1 M TBAF solution in THF (1.4 mL, 1.4 mmol, 1.5 eq). The resulting reaction mixture was stirred at room temperature for 2 h. The solvent was then removed under reduced pressure and the crude residue purified by flash chromatography (0-20 % EtOAc/pentane) to obtain the product (0.15 g, 88 %) as yellow solid.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 1.52-1.56 (m, 1H, $C(sp^3)H_{cyclopropene}$), 1.53-1.61 (m, 2H, $CH_{2cyclopentane}$), 1.68-1.88 (m, 6H, 3x $CH_{2cyclopentane}$), 2.11 (d, $^4J_{HH}$ = 1.1 Hz 3H, CH_3), 2.42-2.52 (m, 1H, $CH_{cyclopentane}$), 3.15 (dddd, $^2J_{HH}$ = 13.7 Hz, $^3J_{HH}$ = 5.5, 4.4 Hz, $^4J_{HH}$ = 0.8 Hz, 1H, CH_2), 3.24 (dddd, $^2J_{HH}$ = 13.7 Hz, $^3J_{HH}$ = 5.5, 4.1 Hz, $^4J_{HH}$ = 0.8 Hz, 1H, CH_2), 5.33 (bs, 1H, NH), 6.56-6.59 (m, 1H, $C(sp^2)H_{cyclopropene}$).
MS (ESI), *m/z* calc for $C_{11}H_{17}NO$ 179.13, found 180.2 $[M+H]^+$

2.5.6 Synthesis of Tetrazine spin labels

2.5.6.1 *N*-(2,2,6,6-Tetramethylpiperidin-1-oxyl)-4-(6-methyl-1,2,4,5-tetrazin-3-yl)-benzamide (Tet-TEMPO)

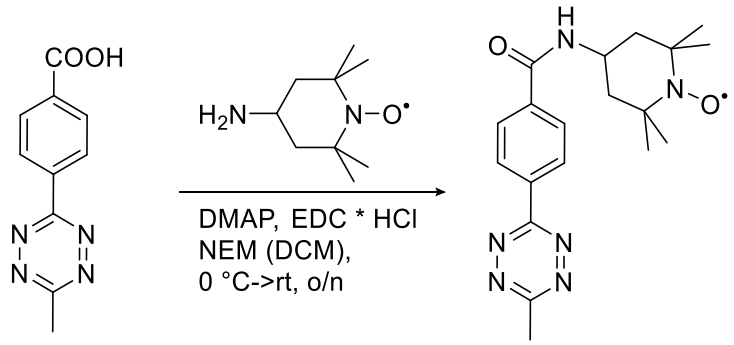
3-(4-Carboxyphenyl)-6-methyl-1,2,4,5-tetrazine[347]



4-Cyanobenzoic acid (1.8 g, 12 mmol, 1 eq) and $Zn(OTf)_2$ (2.2 g, 6 mmol, 0.5 eq) were combined in a flask, then ACN (6.3 mL, 0.12 mol, 10 eq) and hydrazine hydrate (30 mL, 0.6 mol, 50 eq) were added and the reaction heated to 60 °C for 24 h. After addition of $NaNO_2$ (8.3 g, 0.12 mol, 10 eq) dissolved in H_2O (20 mL) the reaction mixture was acidified to pH 3 on ice with a 6 M HCl solution. After stirring for 30 minutes, the pink precipitate was collected by centrifugation. The crude residue was purified by flash chromate-graphy (0-10 % MeOH/DCM + 0.1 % FA), which yielded the product (1.6 g, 61 %) as a pink solid.

^{1}H-NMR (300 MHz, $CDCl_3$): δ = 3.02 (s, 3H, CH_3), 8.20 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $H_{Ar\text{-}2,2'}$), 8.58 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $H_{Ar\text{-}3,3'}$).

***N*-(2,2,6,6-Tetramethylpiperidin-1-oxyl)-4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzamide (Tet-TEMPO)**

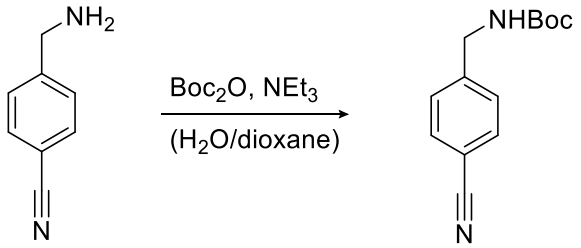


DMAP (28 mg, 0.23 mmol, 0.5 eq) and 3-(4-Carboxyphenyl)-6-methyl-1,2,4,5-tetrazine (0.1 g, 0.46 mmol, 1 eq) were dissolved in 1 mL DCM and cooled in an ice bath before addition of EDC*HCl (0.13 g, 0.7 mmol, 1.5 eq), 4-amino-TEMPO (87 mg, 0.5 mmol, 1.1 eq) and NEM (0.1 mL, 0.92 mmol, 2 eq). The resulting reaction mixture was stirred at room temperature overnight, diluted with 10 mL DCM and washed with 10 % citric acid (2x 10 mL) and saturated $NaHCO_3$ solution (1x 10 mL). The organic phase was collected, dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude residue was purified by flash chromatography (0-1 % MeOH/DCM) to obtain the product (115 mg, 67 %) as a pink solid.

MS (ESI), *m/z* calc for $C_{19}H_{25}N_6O_2$ 369.20, found 370.04 $[M+H]^+$

2.5.6.2 (*N*-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)-2,2,5,5-tetramethyl-1-pyrrolidinyl-oxyl–3-carboxamide (Tet-PROXYL)

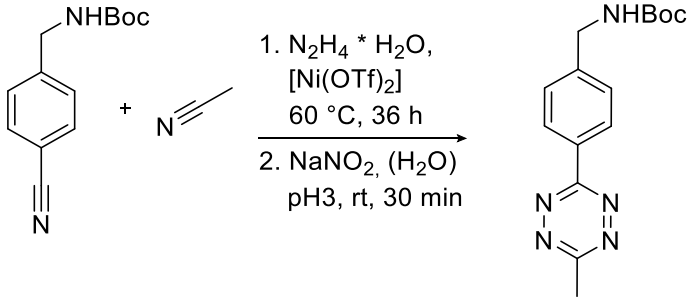
4-(*N*-(tert-butoxycarbonyl)aminomethyl)benzonitrile[244]



4-(Aminomethyl)benzonitrile hydrochloride (3 g, 18 mmol, 1 eq) was evaporated twice with toluene to remove traces of acid, before it was suspended in 40 mL DCM. Then NEt_3 (6.2 mL, 45 mmol, 2.5 eq) and Boc_2O (4.5 mL, 20 mmol, 1.1 eq) were added and the white suspension stirred at room temperature overnight. The reaction was diluted with DCM and washed with 10 % citric acid (2x 50 mL) and saturated $NaHCO_3$ solution (1x 50 mL). The organic phase was collected, dried over Na_2SO_4 and the solvent removed under reduced pressure to give the product (4.8 mg, 92 %) as a white solid.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 1.46 (s, 9H, Boc), 4.37 (d, $^3J_{HH}$ = 4.9 Hz, 2H, CH_2), 4.95 (s, 1H, NH), 7.39 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $H_{Ar\text{-}3,3'}$), 7.62 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $H_{Ar\text{-}2,2'}$).

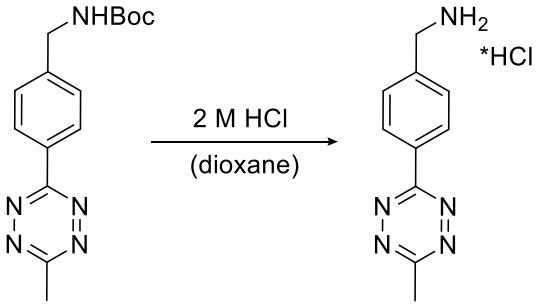
***N*-(*tert*-butoxycarbonyl)-3-(4-aminomethyl)phenyl-6-methyl-1,2,4,5-tetrazine[187]**



4-(*N*-(*tert*-butoxycarbonyl)aminomethyl)benzonitrile (0.5 g, 2.2 mmol, 1 eq) and $Ni(OTf)_2$ (0.38 g, 1.1 mmol, 0.5 eq) were combined in a flask, then ACN (1.8 mL, 33 mmol, 16 eq) and hydrazine hydrate (5.2 mL, 0.11 mol, 50 eq) were added and the reaction heated to 60 °C for 48 h. After addition of $NaNO_2$ (1.5 g, 22 mol, 10 eq) dissolved in H_2O (10 mL) the reaction mixture was carefully acidified to pH 3 on ice with a 6 M HCl solution and stirred for 30 minutes. The reaction was extracted with EtOAc (3x 30 mL), the organic phase dried over Na_2SO_4 and the solvent removed under reduced pressure. The crude residue was purified by flash chromatography (0-30 % EtOAc/pentane) to obtain the product (0.27 g, 42 %) as a purple solid.

^{1}H-NMR (500 MHz, $CDCl_3$): δ = 1.48 (s, 9H, Boc), 3.10 (s, 3H, CH_3), 4.44 (d, $^3J_{HH}$ = 5.9 Hz, 2H, CH_2), 4.95 (s, 1H, NH), 7.51 (d, $^3J_{HH}$ = 8.2 Hz, 2H, $H_{Ar\text{-}3,3'}$), 8.56 (d, $^3J_{HH}$ = 8.2 Hz, 2H, $H_{Ar\text{-}2,2'}$).

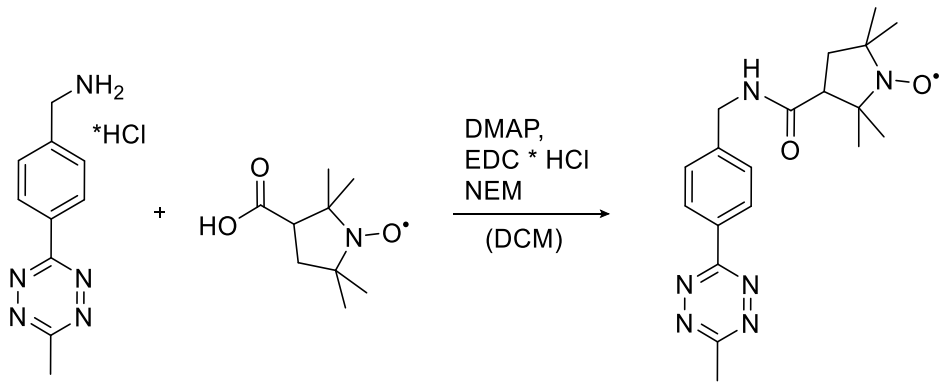
3-(4-aminomethyl)phenyl-6-methyl-1,2,4,5-tetrazine hydrochloride



N-(*tert*-butoxycarbonyl)-3-(4-aminomethyl)phenyl-6-methyl-1,2,4,5-tetrazine (0.13 g, 0.43 mmol, 1 eq) was dissolved in 0.9 mL dioxane, then 4 M HCl solution in dioxane (0.9 mL, 0.86 mmol, 2 eq) was added and the reaction stirred at room temperature for 1 h. Afterwards the solvent was evaporated to give the product (95 g, 93 %) as a pink solid.

^{1}H-NMR (500 MHz, MeOD): δ = 3.06 (s, 3H, CH_3), 4.26 (s, 2H, CH_2), 7.71 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $H_{Ar\text{-}3,3'}$), 8.64 (d, $^3J_{HH}$ = 8.4 Hz, 2H, $H_{Ar\text{-}2,2'}$).

(*N*-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl)-2,2,5,5-tetramethyl-1-pyrrolidinyloxyl–3-carboxamide (Tet-PROXYL)



3-Carboxy-PROXYL (46 mg, 0.25 mmol, 1 eq) was dissolved in 1 mL DCM and cooled in an ice bath, before DMAP (15 mg, 0.13 mmol, 0.5 eq), EDC*HCl (71 mg, 0.37 mmol, 1.5 eq), NEM (54 µL, 0.5 mmol, 2 eq) and 3-(4-aminomethyl)phenyl-6-methyl-1,2,4,5-tetrazine hydrochloride (64 mg, 0.27 mmol, 1.1 eq) were subsequently added and the reaction was stirred overnight at room temperature. After dilution with DCM the organic phase was washed with 10 % citric acid (1x 5 mL), saturated $NaHCO_3$ solution (1x10 mL) and dried over Na_2SO_4. The crude product was purified by flash chromatography (0-1 % MeOH/DCM) to yield the product (84 mg, 93 %) as a pink solid.

MS (ESI), *m/z* calc for $C_{19}H_{25}N_6O_2$ 369.20, found 370.2 $[M+H]^+$, 368.2 $[M-H]^-$

CHAPTER 3

Expanding the genetic code with a lysine derivative bearing an enzymatically removable phenylacetyl group

The work in the following chapter has been published as:

M. Reille-Seroussi, S.V. Mayer, W. Dörner, K. Lang and H.D. Mootz. Expanding the genetic code with a lysine derivative bearing an enzymatically removable phenylacetyl group. *ChemComm*, **2019**, *55*, 4793-4796.

Chapter 3: Expanding the genetic code with a lysine derivative bearing an enzymatically removable phenylacetyl group

3.1 Aim and introduction

Developing adequate tools to study biomolecules in their natural environment to understand complex cellular processes is of great interest. One elegant approach that has emerged over the last years is the development of decaging reactions in combination with genetic code expansion to gain artificial control over enzyme function and structure.[64] GCE can be used to site-specifically incorporate a caged amino acid into the active center or a structurally important feature of an enzyme thereby abolishing its activity. An external, orthogonal stimulus then removes the masking group and reveals the original molecule, leading to full recovery of enzymatic function.

Activation of masked uAAs, especially lysine, tyrosine, serine and cysteine derivatives, has been achieved by either a chemical[251a, 251c,65d, 252] or a light trigger.[65a-c, 348] Irradiation with UV light is the most common approach towards control over biological processes, however it has several limitations. UV light (far and middle UV) is phototoxic for living cells and has difficulty to penetrate tissue or animal samples.[250] Secondly, the *ortho*-nitrobenzyl photo-cage, which is used in most cases, is not completely chemically stable *in vivo*,[65a, 72c, 348a, 349] which can lead to premature deprotection and loss of control. Other complications of the reported deprotection strategies include metabolically[72c, 349] instable compounds or the use of harsh deprotection conditions.[72d] The development of deprotection methods with external triggers like light or small molecules was aimed at *in vivo* applications, which requires the decaging to be very rapid in order to follow fast processes in biological settings. Studying proteins *in vitro* does not necessitate such fast reaction kinetics, but rather relies on stable uAAs that can be easily deprotected using mild deprotection. An elegant stimulus that would fulfill these demands is the use of enzymes to facilitate the cleavage reaction.

So far no genetically encoded uAAs have been reported where enzyme-mediated deprotection has been applied. There are only a few uAAs known bearing post-translational modifications (PTM), which can be considered as masking groups (for example acetyllysine, phosphoserine and phosphotyrosine), as there are eraser proteins that cleave the corresponding moieties. For an artificially triggered encaging approach, the existence of such endogenous enzymes recognizing these PTMs is however problematic.[69a, 350]

The aim of this project, which was a collaboration with the group of Henning Mootz (Universität Münster), was therefore to develop an approach for the enzyme-triggered unmasking of a genetically encoded, caged uAAs. We chose the chemically stable phenylacetyl group (Pac) as protecting group for lysine (PacK, Figure 3.1), which we found can be removed by penicillin G acylase (PGA; E.C.3.5.1.11),[351] and should be bulky enough to affect enzyme activity by locally altering protein structure upon incorporation. So far, the use of Pac deprotection by PGA has been demonstrated on artificially synthesized peptides only.[352]

3.2 Results and Discussion

3.2.1 Site-specific incorporation of Pac-bearing lysine uAAs into proteins

The first step of the project was the synthesis of amino acid PacK carried out by Marie Reille-Seroussi (Universität Münster), based on previously reported procedures in three steps,[72a, 352b] followed by investigation of the cleavage reaction on small molecule level. Therefore, Fmoc-Lys(Pac)-OH was incubated with commercially available PGA and the reaction analyzed by HPLC to confirm complete deprotection of Pac after 10 min (Figure 3.1).

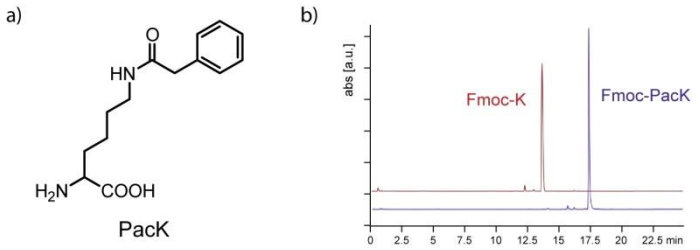


Figure 3.1: a) Structure of PacK. b) HPLC analysis of Pac group removal from Fmoc protected PacK.[66] Adapted by permission of The Royal Society of Chemistry

Next, Marie Reille-Seroussi analyzed the metabolic stability of PacK against endogenous deacylases in *E. coli*, since one of these, CobB, was observed to hydrolyze amide bonds of lysine derivatives with alkyl side chains[63, 350a] Therefore, the growth of lysine auxotrophic *E. coli* strain JW2806-1 in minimal media supplemented with PacK was compared to their growth in the presence of lysine without PacK. No growth was observed for cells grown with PacK, indicating that no endogenous enzyme was able to remove the protecting group.

With these preliminary, positive results in hand, we set out to site-specifically incorporate PacK into proteins, which was performed by myself. Being a lysine derivative, it was hypothesized that the chances of finding a PylRS mutant are quite good, due to structural resemblance with other uAAs that are successfully introduced by PylRS mutants. Therefore, a number of known PylRS mutants (Figure 3.2a,b) were screened for the incorporation of PacK into sfGFP with a premature stop codon and resulted a positive hit. The previously described, very promiscuous mutant BrCnKRS[63] (Y271M, L274A, C313A, *Mb* numbering) selectively and efficiently incorporates PacK into sfGFP-N150TAG-His6 in very good yields of 54 mg/L of culture (Figure 3.2a). The specific incorporation and stability of PacK was confirmed by ESI-MS analysis of the purified protein (Figure 3.2c).

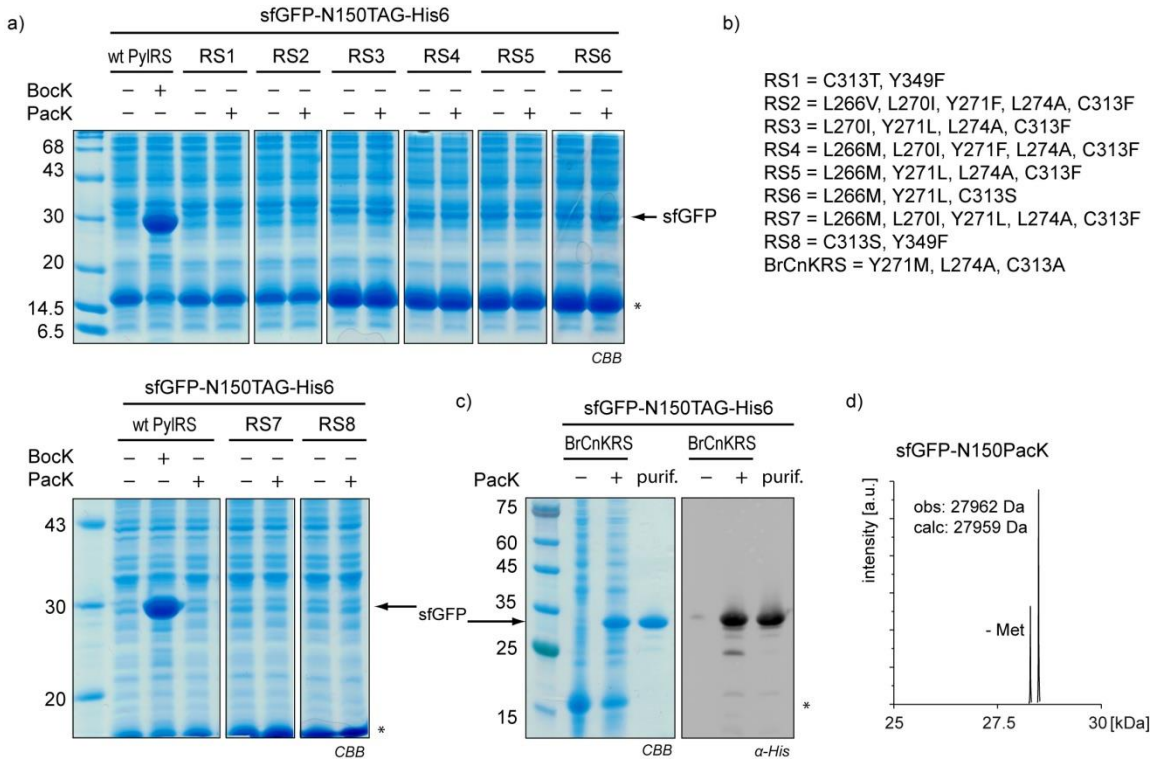


Figure 3.2: Incorporation of PacK into sfGFP-N150TAG-His6.
a) Synthetase screen for the incorporation of PacK into proteins with known PylRS mutants. b) Mutations of PylRS mutants. c) RS9, also known as BrCnKRS, successfully and efficiently incorporates PacK into sfGFP-N150TAG-His6. (Left: Coomassie stained SDS-PAGE, right: α-His6 western blot) d) ESI-MS analysis of purified sfGFP-N150PacK-His6.

Having found a synthetase for the successful incorporation of PacK into proteins, next the unmasking of the Pac moiety on the protein was tested. Therefore, sfGFP-N150PacK-His6 (10 µM) was incubated with PGA (up to 1 U) and samples were taken at different time points, followed by LC-ESI-MS analysis. When removal of the Pac moiety could not be observed (even after 24 to 48 h), our collaboration partners in the Mootz group took a closer look at the active site of PGA. There, Pac is located at the end of a long narrow channel pointing to the active site (Figure 3.3a). At position 150 of sfGFP, PacK sits on the surface of the non-flexible β-barrel, probably not reaching far enough into the active site of PGA for cleavage of Pac, which is supported by a few reports using Pac/PGA deprotection on small proteins like insulin.[352] Therefore, our collaboration partners selected human SUMO-2 (Small Ubiquitin-like MOdifier) as a new model protein, since it contains an unstructured *N*- and *C*-terminal region flanking the ß-grasp fold.[353] K11, one of the 3 lysine residues (8 in total) located in the *N*-terminal region, was chosen as position for amber suppression due to improved accessibility of PacK for deprotection with PGA (Figure 3.3c). SUMO-2-K11PacK was expressed in *E. coli* as an intein-CBP (chitin-binding domain) fusion protein (Figure 3.3b) in good yields (2.5 mg/L of culture) and analyzed by ESI-MS (Figure 3.3c). The purified protein was then subjected to PGA deprotection overnight at rt, which led to the desired, complete deprotection of Pac with 10 U of enzyme (Figure 3.3c).

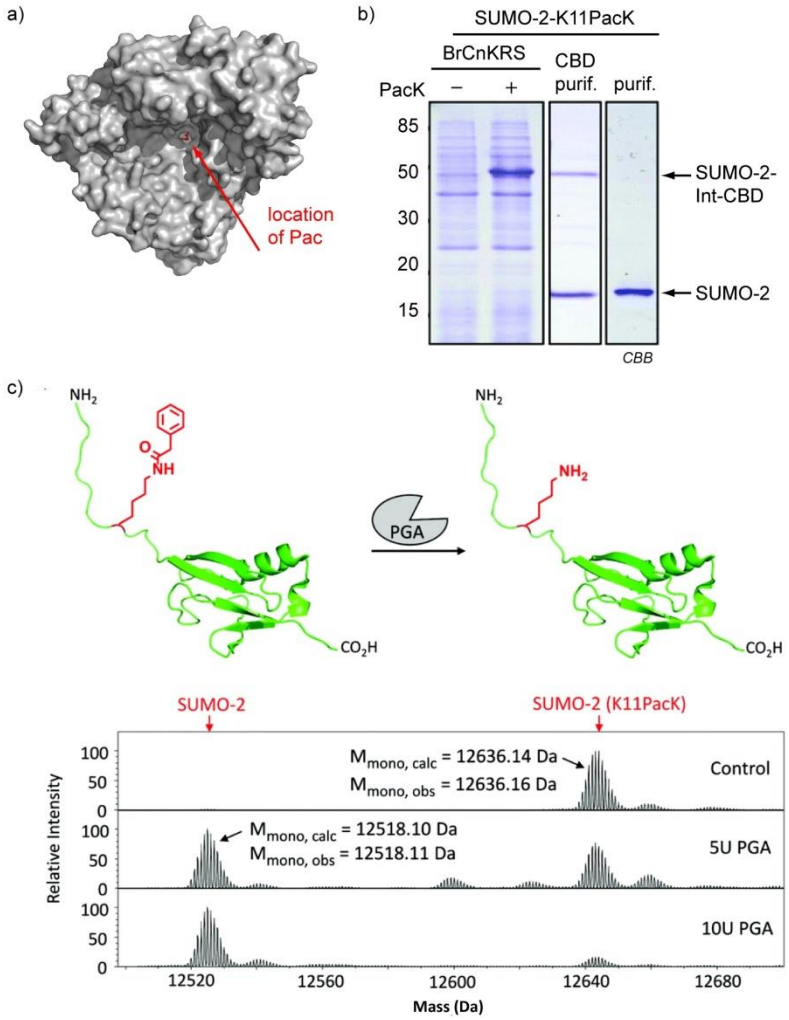


Figure 3.3: Characterization of Pac deprotection of SUMO-2-K11PacK.[66] Adapted by permission of The Royal Society of Chemistry

a) Surface representation of PGA crystal structure (1PNL) with location of Pac indicated in red. b) Successful incorporation of PacK into human SUMO-2-K11TAG using BrCnKRS. c) Top: Schematic illustration of Pac deprotection of SUMO-2-K11PacK by PGA. Bottom: ESI-MS analysis of the cleavage reaction at 25 °C overnight.

3.2.2 Incorporation of PacK into GrsA to control protein interaction

After testing our Pac/PGA unmasking approach on model proteins, such as sfGFP and SUMO-2, the Mootz group used this approach to study a more complex system. Multi-domain protein gramicidin S synthetase A (GrsA) is a nonribosomal peptide synthetase (NRPS) in the protein-templated biosynthesis of gramicidin S.[354] It is of great interest for the Mootz group to study the intermolecular interactions between the different domains of GrsA, as well as with downstream partners, to elucidate the up-to-date not completely understood mechanism of NRPS-mediated peptide assembly.

In the biosynthesis of gramicidin GrsA facilitates the transformation of L-phenylalanine (L-Phe) to D-Phe by activation, binding and epimerization. For this it comprises an adenylation (A) domain (PheA), a peptidyl carrier protein (PCP) domain and an epimerization (E) domain (Figure 3.4a). The PheA domain performs two different reactions, adenylation of L-Phe and transfer to the neighboring PCP domain by thiolation. To rearrange the active site in between these two different reactions A has to undergo domain alternation, where the *C*-terminal

subdomain A (A^C) rotates ~140° relative to the *N*-terminal subdomain A (A^N).[355] Each of these two conformations comprises a conserved lysine residue that is identically located relative to the A^N subdomain, K517 (adenylation conformation) and K434 (thiolation conformation) in GrsA (Figure 3.4b). An alanine scan of these two positions revealed K517 to be important for catalytic activity, while the mutant K434A still showed activity in both reactions.[356]

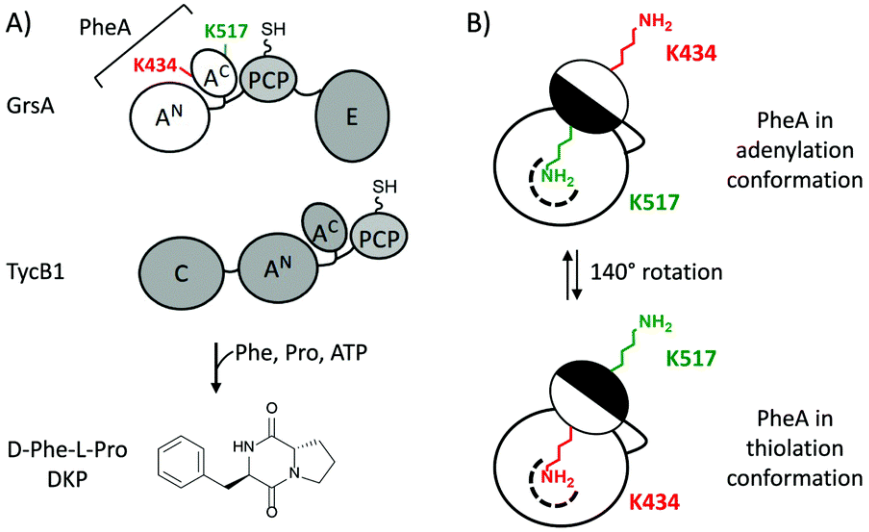


Figure 3.4: GrsA activity.[66] Reproduced by permission of The Royal Society of Chemistry
A) Schematic representation of NRPSs constructs and their reaction product. B) Illustration of the two lysine residues in the conformational states of the A-domain of GrsA (PheA). The black dotted line represents the active site.[66]

The Mootz group wanted to use our newly developed enzyme-mediated decaging approach to study the influence of both on the adenylation and thiolation activity of the PheA domain of GrsA by introducing our Pac masked lysine uAA into these conserved lysine residues K434 and K517. Addition of PGA should then lead to a recuperation of activity. For this, our collaboration partners successfully amber suppressed positions K434 or K517 in multi-domain protein GrsA with PacK (Figure 3.5a), which was confirmed by ESI-MS (Figure 3.5b). To test if the uAA PacK is able to mask enzyme activity, our collaboration partners performed an activity assay with GrsA K434PacK or K517PacK mutants before and after incubation with PGA. The assay monitors formation of dipeptide D-phenylalanyl-L-prolyl-diketopiperazine (DKP) from L-Phe and L-proline (L-Pro) by the interaction of GrsA with TycB1, the second module of tyrocidine biosynthesis.[357] The formation of DKP was monitored by HPLC and the fully inactive, enzymatically masked mutants GrsA-K434PacK and GrsA-K517PacK were characterized compared to the wt (Figure 3.5c). Subjection to 5 U PGA almost fully restored GrsA-K517PacK activity, while only 10 % of wt production of DKP was observed for the GrsA-K434PacK mutant. This suggests that deprotection of PacK in this position was incomplete, which might be due to lower accessibility in the surrounding of the PCP and E domain. Surprisingly, treatment with a higher amount of PGA (10 U) lead to a decrease of DKP generation, indicating lower deprotection levels of both GrsA mutants. A reason therefore might be precipitation of PGA at higher concentrations, due to being commercially available as suspension of low purity´.

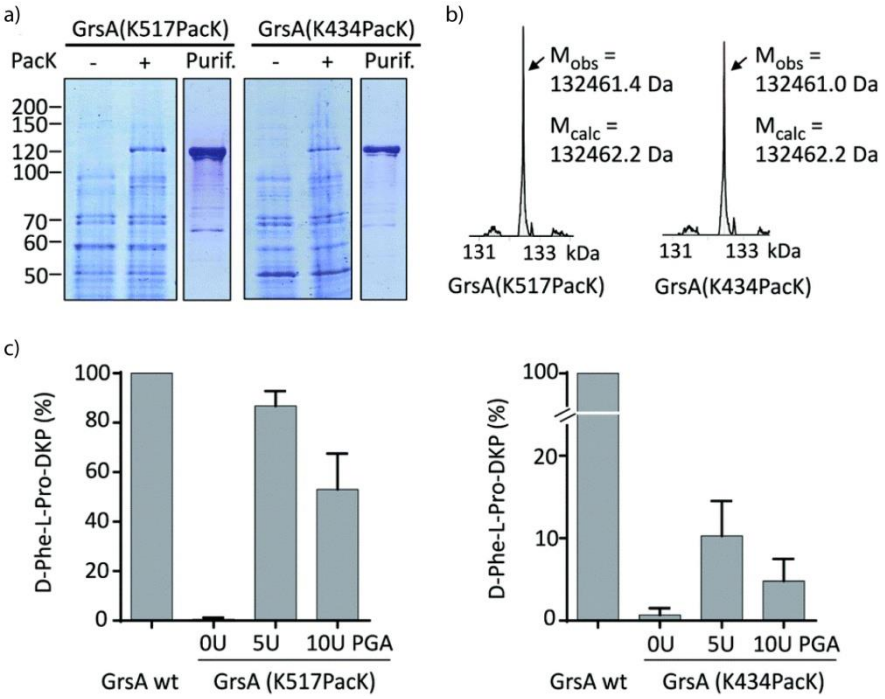


Figure 3.5: Enzymatic control of GrsA activity.[66] Reproduced by permission of The Royal Society of Chemistry
a) Coomassie stained SDS-PAGE of Incorporation of PacK into GrsA constructs. b) ESI-MS analysis of the purified proteins. c) Activity of GrsA proteins in a DKP product formation assay with TycB1 as illustrated in (Figure 3.4), with or without prior incubation with PGA overnight at 25 °C. The results of at least four experiments are presented, error bars represent standard deviations.[66]

3.2.3 SrtN as alternative enzyme for the deprotection of PacK

The results using PGA clearly demonstrate our enzyme-triggered deprotection approach of a genetically introduced, masked uAA to gain artificial control over protein function to be quite effective, albeit there are some limitations for the use of PGA. Therefore, our collaboration partners were looking into an alternative class of enzymes suitable for our deprotection approach, the sirtuin family, known for their deacylase activity on lysine *N*-acyl modifications.[358] However, to our knowledge the removal of bulky, aromatic groups like Pac via sirtuins has not been reported so far. After recombinant expression of SrtN from *Bacillus subtilis* in *E. coli*, our collaboration partners tested the deacylase activity of SrtN against PacK. First, model protein SUMO-2-K11PacK was incubated with SrtN and rapid removal of Pac observed after 5 min at rt only (Figure 3.6a), which corresponds to a half-life time of 5.9 min under sub-stochiometric conditions (Figure 3.6b). SrtN deprotection of PacK on SUMO-2 is therefore considerably faster than PGA facilitated Pac removal, which in contrast takes several hours. As a next step, our collaboration partners tested SrtN activity on the PacK mutants of GrsA. GrsA-K434PacK and GrsA-K517PacK were only partially cleaved with a slight preference for K434PacK. Elevating temperatures to 37 °C significantly enhanced SrtN activity on these substrates. Even though, the deprotection did not proceed as fast as on SUMO-2, these results show that for the first time a sirtuin can deacylate the bulky Pac group from lysine. In structurally accessible regions like the C-terminus of SUMO the cleavage reaction is quantitative and possesses rapid reaction rates.

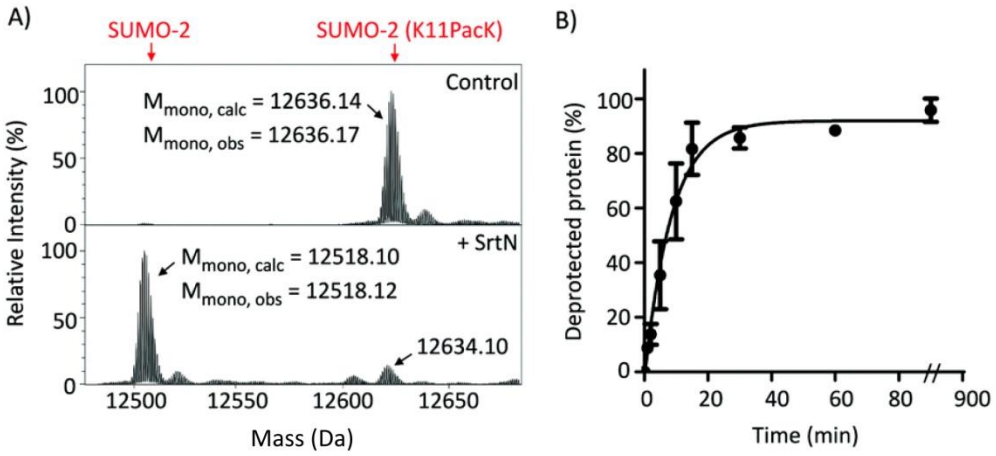


Figure 3.6: MS analysis of the deprotection of SUMO(K11PacK) by SrtN.[66] Adapted by permission of The Royal Society of Chemistry
A) SUMO-2-K11PacK (5 mM) before and after incubation with 4 equivalents of SrtN at 25 °C for 5 min. B) Time course analysis of deprotection using a sub-equimolar amount of enzyme ([Sumo-2-K11PacK] = 5 mM; [SrtN] =2 mM). The results of at least three experiments are presented, error bars represent standard deviations.

3.3 Summary and Outlook

In conclusion, together with our collaboration partners, we have reported the first enzyme-mediated decaging reaction in combination with genetic code expansion to orthogonally modulate conformation and biological activity of enzymes. We presented the site-specific incorporation of an unnatural lysine derivative bearing the enzyme-labile phenylacetyl group (Pac) into proteins in *E. coli* using a previously described, promiscuous PylRS mutant from *M. barkeri* and efficient deprotection of Pac using two different enzymes. Penicillin G acylase was already known to act on PacK, albeit has only been applied on small proteins synthesized via solid phase before. PGA efficiently removes Pac in readily accessible structural features, but shows decreased to no activity on more complex or rigid structures, due to a long and narrow channel leading to its active site. Another limitation is the quality of the commercially available PGA, our collaboration partners are therefore working on establishing their own expression system for PGA at the moment, which is unfortunately not trivial and requires periplasmic expression for correct folding.[359]

The second enzyme that was reported for the enzyme-mediated decaging of genetically encoded PacK is SrtN, a deacylase from the sirtuin family, which was previously not known to accept Pac bearing lysine as substrate. SrtN rapidly and efficiently removed Pac from SUMO-2-K11PacK within 5 min compared to hours necessary for PGA deprotection. SrtN also favors the more readily accessible sites for PacK deprotection; however cleavage activity can be enhanced by performing the reaction at physiological temperatures compared to 25 °C. Altogether, our newly developed enzyme-mediated deprotection approach has great potential to gain control of protein function and structure by an external enzyme trigger, due to its mild conditions and efficient and rapid removal of PacK by SrtN.

II. Materials and General Methods

II. Materials and General Methods

II.1 Chemical Methods

II.1.1 Chemicals

Commercially available chemicals were purchased at the highest available quality from the Suppliers Acros Organics, Alfa Aesar, Applichem, Carbolution, Carl Roth, Iris Biotech, Jena Bioscience, Sigma Aldrich, Thermo Fisher Scientific and Tokyo Chemical Industry. These chemicals were used without further purification. Solvents for reactions were purchased as anhydrous if not used in combinations with water. All reactions were performed under Argon atmosphere from an Argon balloon attached to the reaction vessel. Schlenk conditions were only used when indicated in the reaction procedure.

II.1.2 Solutions

Table II.1: Solutions for work up

Solution	Amount/Concentration
Citric acid	20 % (w/v)
HCl	1 M
HCl	3 M
HCl	6 M
LiOH	0.57 M
NaCl	saturated
$NaHCO_3$	saturated
NaOH	1 M
Na_2SO_3	20 % (w/v)

II.1.3 TLC

Thin layer chromatography (TLC) was used to analyze the conversion of starting material of a reaction on TLC plates Silica gel 60 F_{254} (Merck, Germany) in combination with UV at 254 nm and the following staining solutions:

Table II.2: Used staining solutions for TLC

Staining solution	Composition	Use	
Bromocresol Green	0.04 g bromocresol green 100 mL EtOH 0.1 M NaOH dropwise until blue	Carboxylic acids	yellow (acid) or blue (base) spots on light blue background
Cerium molybdate	2.2 g $Ce(SO_4)_2$	Universal	blue spots on light

	5.6 g Phosphomolybdic acid 200 mL H_2O 7.1 mL H_2SO_4		background
Potassium permanganate	1.5 g $KMnO_4$ 10 g K_2CO_3 1.25 mL 10% NaOH 200 mL H_2O	Olefins, oxidizable groups	yellow spots on purple background
Ninhydrin	0.72 g ninhydrin 200 mL EtOH	Amines	pink to blue spots on light background

II.1.3. Silica flash column purification

Solvents used for purification by flash chromatography were either purchased with analytical purity from Fisher Scientific (EtOAc, MeOH) or were purchased from the Materialverwaltung (TUM) as technical grade (DCM, pentane) and further purified by distillation before use. Compounds were separated over silica gel with a grain distribution of 40-60 µm and a pore size of 60 Å (Silica gel, Ultra Pure, 40-60µm, 60Å, Acros Organics, Germany). Solvents are given as mixtures of volumes.

II.1.4. Liquid chromatography mass spectrometry (LC-MS)

LC-MS spectra were recorded on an Agilent 1260 Infinity Series LC system with an Agilent 6210 ESI Single Quadrupole mass spectrometer. Solvents A (MilliQ H_2O+ 0.1 % FA) and B (ACN + 0.1 % FA) were used without filtration. Small molecules were analyzed via a Jupiter C18 5 µm (2 x 150 mm) capillary column (Phenomenex, Torrence, USA) at rt with a flow rate of 0.5 mL/min with a gradient from 5-95 % B in 6 min. Proteins samples were separated on a Jupiter C4 5 µm (2 x 150 mm) capillary column (Phenomenex, Torrence, USA) at rt with a flow rate of 0.3 mL/min with a gradient of 5-55 % B in 5 min. Measurements were carried out in positive and negative ion mode and the resulting spectra analyzed with OpenLab ChemStation Edition Software C.01.07.SR3 [465].

II.1.5. High performance liquid chromatography (HPLC)

HPLC purification were carried out with a Shimadzu LC-20AT Prominence system, consisting of two pumps (LC-20AT), a degassing unit (DGU-$20A_{3R}$), a diode array detector (SPD-M-20A), a fraction collector (FRC-10A) and a communications bus module (CBM-20A). A Phenomenex Luna C18, 5 µm (4.6 x 250 mm) column was used to separate compounds on an analytical scale; a Phenomenex Luna C18, 5 µm (10 x 250 mm) column and a Phenomenex Luna C18, 10 µm (21.2 x 250 mm) for preparative scale runs. Solvents A (MilliQ H_2O+ 0.1 % FA) and B (ACN + 0.1 % FA) were used without filtration.

II.1.6. NMR spektroscopy

^{1}H and ^{13}C NMR spectra were measured on Bruker Systems AVHD300 (300; 75 Hz) or AVHD400 (400; 100 Hz) and AVHD500 (500 Hz, ^{1}H only). The chemical shift δ [ppm] was calibrated on the residual proton peak of the used solvent. One to two drops of TFA were added to dissolve the final amino acids in DMSO. Signal multiplicity is characterized as s (singlet), d (doublet), t (triplet), q (quartett), h (heptet), m (multiplet) and their combinations.

II.1.7 Solid Phase Peptide Synthesis (SPPS)

SPPS was performed in plastic syringes with a frit. Shaking was performed by using a rotary unit. The equivalents (eq.) used were based on the maximal loading capacity of the CTC resin given by the supplier.

II.1.8 Decaging of cyclopropenone compounds at 365 nm

photo-DMBO compounds were decaged to the alkyne (DMBO) by loss of CO at 365 nm using a UV Lamp VL-215.L, or 365 nm LEDs with 1 or 3 W.

II.2 Biochemical Work

II.2.1 Material

II.2.1.1 Used Organisms

Table II.3: List of used prokaryotic organisms

Organism	Genotype	Use	Origin
Escherichia coli DH 10 B	F⁻ *ara*D139 Δ(*ara,leu*)7697 Δ*lac*X74 *gal*U *gal*K *rps*L *deo*R Φ80d*lac*ZΔM15 *end*A1 *nup*G *rec*A1 *mcr*A Δ(*mrr hsd*RMS *mcr*BC)	Cloning, Transformation, Expression	New England Biolabs (NEB)Inc. (Ipswich, USA)
Escherichia coli BL21 (DE3)	F⁻ *fhuA2 [lon] ompT gal (λ DE3) [dcm]* Δ*hsdS* *λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21* Δ*nin5*	Transformation, Expression	NEB (Ipswich, USA)
Escherichia coli BL21 AI	F⁻*omp*T *hsd*S_B (r_B^- m_B^-) *gal dcm ara*B::T7RNAP-*tet*A	Transformation, Expression	Lemke Lab, EMBL (Heidelberg)
Escherichia coli JW2203-1	F⁻ *Δ(araD-araB)567* *ΔlacZ4787*(::rrnB-3) *λ*⁻ *ΔompC768::kan rph-1 Δ(rhaD-rhaB)568 hsdR514*	Transformation, Expression	Chair for Microbiology, LMU (Munich)

II.2.1.2 Used Plasmids

Plasmids encoding constitutively expressed mutant *M. barkeri* or *mazei* Pyrolysl-tRNA synthetases (PylRS) on a pBK backbone were combined with pPylt plasmids containing the genes for a constitutively expressed mutant *M. bakeri* pyrolysyl-$tRNA_{CUA}$ as well as a *L*-arabinose indicuble araBAD promotor system for the expression of the proteins of interest harboring amber codons in different positions. Alternative expression systems comprised the genes for the tRNA synthetase and the tRNA on one plasmid (pEVOL, pDule) and the arabinose inducible gene for the amber-suppressible protein of interest on another wit pBAD backbone. pDule plasmids contained the mutant *M. jannaschii* Tyrosyl-tRNA synthetase and the corresponding tRNA, while pEVOL plasmids were available for both *M. jannaschii* TyrRS/tRNA and *M. barkeri* or *mazei* PylRS/ *M. mazei* Pylt pairs. pEVOL plasmids encoded two copies of synthetase, one constitutively expressed and one under a *L*-arabinose inducible araBAD promotor, and one copy of constitutively expressed tRNA.

Table II.4: List of used PylRS mutant plasmids

Plasmid	Organism	bb	Nucleic acid	AR	Original vector	Primer
D4	Mb	pBK	PylRS wt	Kan	-	-
PylRS wt (Amp)	Mb	pBK	PylRS wt	Amp	-	-

D1	Mb	pBK	PylRS Y271M, L274G, C313A	Amp	-	-
D2	Mb	pBK	PylRS (D76G, S123G, T269N) Y271V, L274M	Amp	-	-
D3	Mb	pBK	PylRS (M40L), (D76G), (S123G), M241F, A267S, Y271C, L274M	Amp	-	-
D19	Mb	pBK	PylRS (D76G), (V95A), (S123G), A267S, L274G, C313V	Amp	-	-
D43	Mb	pBK	PylRS (D76G), (S123G), A267S, C313V, M315F (D344G)	Amp	-	-
D60	Mm	pBK	PylRS wt	Kan	-	-
D78	Mm	pBK	PylRS Y271A, L274M, C313A (S123G not programmed)	Amp	-	-
D79	Mb	pBK	PylRS C313V	Kan	-	-
D161	Mb	pBK	PylRS Y271M, L274G, C313A, Y349W	Kan	-	-
D171	Mb	pBK	PylRS Y271M, L274A, C313A	Kan	-	-
D172	Mb	pBK	PylRS C313I	Kan	-	-
D173	Mb	pBK	PylRS Y271L, L274F, N311M, C313G	Kan	-	-
D181	Mb	pBK	PylRS (R85H), N311M, C313Q, V366G, W382N	Kan	-	-
D182	Mb	pBK	PylRS (S131N), N311Q, C313S, (E340Q), V366G, W382N	Kan	-	-
D187	Mb	pBK	PylRS Y271M, L274A, C313A, Y349F	Kan	-	-
D188	Mb	pBK	PylRS Y271M, L274A, C313A, Y349F	Amp	-	-
D189	Mb	pBK	PylRS Y271L, L274F, N311M, C313G	Kan	-	-
D197	Mb	pBK	PylRS V366G, W382N	Kan	-	-
TM3	Mb	pBK	PylRS Y349F	Kan	-	-
MC1	Mb	pBK	PylRS Y271A, L274M	Kan	-	-
MF3	Mb	pBK	PylRS Y271G, C313V	Amp	-	-
MF4	Mb	pBK	PylRS Y271A, C313V	Amp	-	-
MF9	Mb	pBK	PylRS L266M, L270I, Y271F, L274A, C313F	Amp	-	-
MF10	Mb	pBK	PylRS C313S, Y349F	Kan	-	-
MF11	Mb	pBK	PylRS C313T, Y349F	Kan	-	-
MF12	Mb	pBK	PylRS L266V, L270I, Y271F, L274A, C313F	Amp	-	-
MF13	Mb	pBK	PylRS L270I, Y271L, L274A, C313F	Amp	-	-

MF14	Mb	pBK	PylRS L266M, Y271L, L274A, C313F	Amp	-	-
MF15	Mb	pBK	PylRS L266M, Y271L, C313S	Amp	-	-
MF16	Mb	pBK	PylRS L266M, L270I, Y271L, L274A, C313F	Amp	-	-
MF17	Mb	pBK	PylRS L266M, L270V, Y271L, L274A, C313F	Amp	-	-
MF18	Mb	pBK	PylRS L274A, N311Q, C313S	Kan	-	-
MF19	Mb	pBK	PylRS A267T, L274A, N311Q, C313S	Kan	-	-
pEVOL MF3	Mb	pEVOL	PylRS Y271G, C313V, Pylt	Cam	-	-
MF3 (Mm)	Mm	pBK	PylRS (Mm) Y306G, C348V	Kan	SM53	QC51, 52
pEVOL MF3 (Mm)	Mm	pEVOL	PylRS (Mm) Y306G, C348V, Pylt	Cam	MF3 (Mm)	RC: 50, 51 HiFi: 52-55
SM1	Mb	pBK	PylRS N311A, C313A	Kan	D4	QC1, 2
SM2	Mb	pBK	PylRS A267T	Kan	D4	QC3, 4
SM5	Mb	pBK	PylRS N311A, C313A, W382A	Kan	SM1	QC 5, 6
SM6	Mb	pBK	PylRS N311A, C313A, A267T	Kan	SM1	QC3, 4
SM7	Mb	pBK	PylRS A267T, L274A	Kan	SM2	QC7, 8
SM8	Mb	pBK	PylRS A267T, W382A	Kan	SM2	QC5, 6
SM9	Mb	pBK	PylRS N311A, C313A, A267T, W382A	Kan	SM6	QC5, 6
SM10	Mb	pBK	PylRS N311A, C313A, A267T, L274A	Kan	SM7	QC7, 8
SM17	Mb	pBK	PylRS L270I, Y349F	Kan	TM3	QC9, 10
SM18	Mb	pBK	PylRS L270I, Y271F, Y349F	Kan	SM17	QC11, 12
SM19	Mb	pBK	PylRS C313A	Kan	D4	QC17, 18
SM20	Mb	pBK	PylRS L274A	Kan	D4	QC21, 22
SM21	Mb	pBK	PylRS L270I, Y271F, C313F, Y349F	Kan	SM18	QC15, 16
SM22	Mb	pBK	PylRS L274A, C313A	Kan	SM19	QC21, 22
SM23	Mb	pBK	PylRS Y271A, Y349F	Kan	TM3	QC19, 20
SM24	Mb	pBK	PylRS L274A, Y349F	Kan	TM3	QC21, 22
SM25	Mb	pBK	PylRS C313A, Y349F	Kan	TM3	QC17, 18
SM26	Mb	pBK	PylRS L274A, C313A, Y349F	Kan	SM24	QC17, 18
SM27	Mb	pBK	PylRS L270I, Y271F, L274G, C313F, Y349F	Kan	SM21	QC35, 36
SM28	Mb	pBK	PylRS N311A, C313S	Kan	D4	QC29, 30

SM29	Mb	pBK	PylRS L274A, C313V, Y349F	Kan	SM24	QC23, 24
SM30	Mb	pBK	PylRS L274A, C313S, Y349F	Kan	SM24	QC25, 26
SM31	Mb	pBK	PylRS Y271A, Y349F, I378L	Kan	SM23	QC27, 28
SM33	Mb	pBK	PylRS Y271M, L274A, C313A, Y349W	Kan	D171	QC31, 32
SM34	Mb	pBK	PylRS M241A, Y271A, L274V, C313V, M315Y, V370R	Kan	D4	HiFi: 13 -16, GS17
SM35	Mb	pBK	PylRS Y271A, L274V, C313V, M315Y, V370R	Kan	D4	HiFi: 13 -16, GS18
SM36	Mb	pBK	PylRS Y271A, L274V, C313V, M315Y, Y349F, V370R	Kan	D4	HiFi: 13 -16, GS19
SM37	Mb	pBK	PylRS M241A, Y271A, L274V, C313V, M315Y, Y349F, V370R	Kan	SM34	QC39, 40
SM41	Mb	pBK	PylRS Y271G, L274A, C313V	Kan	MF3	QC43, 44
SM42	Mb	pBK	PylRS Y271G, C313F	Kan	MF3	QC45, 46
SM43	Mb	pBK	PylRS Y271G, L274A, C313F	Kan	SM42	QC43, 44
SM53	Mm	pBK	PylRS Y306G	Amp	D60	QC49, 50

Table II.5: List of used *M. jannaschii* TyrRS mutant plasmids

Plasmid	bb	Nucleic acid	AR	Original vector	Primer
SM15	pDule1	MjTyrRS_Y32G, L65Q, F108S, Q109D, D158S, L162N; tRNA	Tcn	-	-
SM16	pDule2	MjTyrRS_Y32G, L65Q, F108S, Q109D, D158S, L162N; tRNA	Sm	-	-
pET301	pET301		Amp	-	-
pET301 tet2.0	pET301	MjTyrRS_Y32G, L65Q, F108S, Q109D, D158S, L162N	Amp	pET301, SM16	HiFi: 38, 39, 44,45
pET301 GS5	pET301	MjTyrRS_Y32G, L65Q, H70A, F108S, Q155E, D158S, L162N	Amp	pET301 tet2.0	HiFi: 1-4, GS5
pET301 GS6	pET301	MjTyrRS_Y32G, L65Q, H70A, F108S, D158S, L162N	Amp	pET301 tet2.0	HiFi: 1-4, GS6
pET301 GS7	pET301	MjTyrRS_Y32G, L65Q, H70A, F108S, Q155E, D158S, L162N	Amp	pET301 tet2.0	HiFi: 1-4, GS7
pET301 GS8	pET301	MjTyrRS_Y32G, L65Q, H70A, Q155E, D158S, L162N	Amp	pET301 tet2.0	HiFi: 1-4, GS8
pET301 GS8QC	pET301	MjTyrRS_Y32G, L65Q, H70A, D158S, L162N	Amp	pET301 GS8	QC 41, 42
pET301 GS9	pET301	MjTyrRS_Y32G L65Y H70A Q155E D158G I159W L162S	Amp	pET301 tet2.0	HiFi: 1-4, GS9
SM44	pEVOL	MjTyrRS_Y32G, L65Q, F108S, Q109D, D158S, L162N; tRNA	Cam	pET301 tet2.0	RC1: 40, 41 RC2: 42, 43

SM45	pEVOL	MjTyrRS_Y32G, L65Q, H70A, F108S, Q155E, D158S, L162N; tRNA	Cam	pET301 GS5	RC1: 40, 41 RC2: 42, 43
SM46	pEVOL	MjTyrRS_Y32G, L65Q, H70A, F108S, D158S, L162N; tRNA	Cam	pET301 GS6	RC1: 40, 41 RC2: 42, 43
SM47	pEVOL	MjTyrRS_Y32G, L65Q, H70A, F108S, Q155E, D158S, L162N; tRNA	Cam	pET301 GS7	RC1: 40, 41 RC2: 42, 43
SM48	pEVOL	MjTyrRS_Y32G, L65Q, H70A, Q155E, D158S, L162N; tRNA	Cam	pET301 GS8	RC1: 40, 41 RC2: 42, 43
SM49	pEVOL	MjTyrRS_Y32G, L65Q, H70A, D158S, L162N; tRNA	Cam	pET301 GS8QC	RC1: 40, 41 RC2: 42, 43
SM50	pEVOL	MjTyrRS_Y32G L65Y H70A Q155E D158G I159W L162S; tRNA	Cam	pET301 GS9	RC1: 40, 41 RC2: 42, 43

Table II.6: List of used reporter plasmids for protein expression

Plasmid/Genomic DNA	bb	Nucleic acid	AR	Original vector	Primer
pPylt_sfGFP150TAG	pPylt	sfGFP150TAG-His, Pylt	Tcn	-	-
pPylt_sfGFP wt	pPylt	sfGFP wt-His, Pylt	Tcn	pPyltsfGFP150TAG	QC 53, 54
pPylt_sfGFP3TAG	pPylt	sfGFP3TAG-His, Pylt	Tcn	pPyltsfGFPwt	QC 57, 58
pPylt_sfGFP40TAG	pPylt	sfGFP40TAG-His, Pylt	Tcn	pPyltsfGFPwt	QC 55, 56
pPylt_sfGFP109,150 TAG	pPylt	sfGFP109,150TAG-His, Pylt	Tcn	-	-
pPylt_Ub6TAG	pPylt	Ub6TAG-His, Pylt	Tcn	-	-
pPylt_Myo4TAG	pPylt	Myo4TAG-His, Pylt	Tcn	-	-
E.coli Genomic DNA	-	OmpC	-	-	46, 47
pPylt_OmpC wt	pPylt	OmpC wt-His, Pylt	Tcn	pPyltsfGFP150TAG	HiFi: 46-49
pPylt_OmpC232TAG	pPylt	OmpC232TAG-His, Pylt	Tcn	pPyltOmpC wt	QC 47, 48
pBAD_sfGFP wt	pBAD	sfGFP wt-His	Amp	-	-
pBAD_sfGFP150TAG	pBAD	sfGFP150TAG-His	Amp	-	-

Table II.7: Sequences of used plasmids

Plasmid	Sequence
pBK_*Mb* PylRS wt (Kan)	CTCGGGTTGTCAGCCTGTCCCGCTTATAAGATCATACGCCGTTATACGTT GTTTACGCTTTGAGGAATCCCATATGATGGATAAAAAACCGCTGGATGT GCTGATTAGCGCGACCGGCCTGTGGATGAGCCGTACCGGCACCCTGCAT AAAATCAAACATCATGAAGTGAGCCGCAGCAAAATCTATATTGAAATG GCGTGCGGCGATCATCTGGTGGTGAACAACAGCCGTAGCTGCCGTACCG CGCGTGCGTTTCGTCATCATAAATACCGCAAAACCTGCAAACGTTGCCG

TGTGAGCGATGAAGATATCAACAACTTTCTGACCCGTAGCACCGAAAGC
AAAAACAGCGTGAAAGTGCGTGTGGTGAGCGCGCCGAAAGTGAAAAAA
GCGATGCCGAAAAGCGTGAGCCGTGCGCCGAAACCGCTGGAAAATAGC
GTGAGCGCGAAAGCGAGCACCAACACCAGCCGTAGCGTTCCGAGCCCG
GCGAAAAGCACCCCGAACAGCAGCGTTCCGGCGTCTGCGCCGGCACCG
AGCCTGACCCGCAGCCAGCTGGATCGTGTGGAAGCGCTGCTGTCTCCGG
AAGATAAAATTAGCCTGAACATGGCGAAACCGTTTCGTGAACTGGAACC
GGAACTGGTGACCCGTCGTAAAAACGATTTTCAGCGCCTGTATACCAAC
GATCGTGAAGATTATCTGGGCAAACTGGAACGTGATATCACCAAATTTT
TTGTGGATCGCGGCTTTCTGGAAATTAAAAGCCCGATTCTGATTCCGGC
GGAATATGTGGAACGTATGGGCATTAACAACGACACCGAACTGAGCAA
ACAAATTTTCCGCGTGGATAAAAACCTGTGCCTGCGTCCGATGCTGGCC
CCGACCCTGTATAACTATCTGCGTAAACTGGATCGTATTCTGCCGGGTCC
GATCAAAATTTTTGAAGTGGGCCCGTGCTATCGCAAAGAAAGCGATGGC
AAAGAACACCTGGAAGAATTCACCATGGTTAACTTTTGCCAAATGGGCA
GCGGCTGCACCCGTGAAAACCTGGAAGCGCTGATCAAAGAATTCCTGGA
TTATCTGGAAATCGACTTCGAAATTGTGGGCGATAGCTGCATGGTGTAT
GGCGATACCCTGGATATTATGCATGGCGATCTGGAACTGAGCAGCGCGG
TGGTGGGTCCGGTTAGCCTGGATCGTGAATGGGGCATTGATAAACCGTG
GATTGGCGCGGGTTTTGGCCTGGAACGTCTGCTGAAAGTGATGCATGGC
TTCAAAAACATTAAACGTGCGAGCCGTAGCGAAAGCTACTATAACGGCA
TTAGCACGAACCTGTAACTGCAGTTTCAAACGCTAAATTGCCTGATGCG
CTACGCTTATCAGGCCTACATGATCTCTGCAATATATTGAGTTTGCGTGC
TTTTGTAGGCCGGATAAGGCGTTCACGCCGCATCCGGCAAGAAACAGCA
AACAATCCAAAACGCCGCGTTCAGCGGCGTTTTTTCTGCTTTTCTTCGCG
AATTAATTCCGCTTCGCACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC
CCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC
CCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCG
TGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTT
CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCT
CAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCC
CCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGC
TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC
GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC
AGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC
TACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTG
GTCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGT
GTTATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAA
ATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAA
TGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGAT
GCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATG
TTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCT
TCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCA
CCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCC
TGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGG
TTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTT
CGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA
GTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAA
AGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATG
GTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGT
TGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTG
CCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGG
CTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTT
TCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGT
AACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGA
ATAAATCGAACTTTTGCTGAGTTGAAGGATC

pBK_*Mb* MF3 (Amp)	CTCGGGAGTTGTCAGCCTGTCCCGCTTATAAGATCATACGCCGTTATACG TTGTTTACGCTTTGAGGAATCCCATATGATGGATAAAAAACCGCTGGAT GTGCTGATTAGCGCGACCGGCCTGTGGATGAGCCGTACCGGCACCCTGC ATAAAATCAAACATCATGAAGTGAGCCGCAGCAAAATCTATATTGAAAT GGCGTGCGGCGATCATCTGGTGGTGAACAACAGCCGTAGCTGCCGTACC GCGCGTGCGTTTCGTCATCATAAATACCGCAAAACCTGCAAACGTTGCC GTGTGAGCGGTGAAGATATCAACAACTTTCTGACCCGTAGCACCGAAAG CAAAAACAGCGTGAAAGTGCGTGTGGTGAGCGCGCCGAAAGTGAAAAA AGCGATGCCGAAAAGCGTGAGCCGTGCGCCGAAACCGCTGGAAAATAG CGTGGGCGCGAAAGCGAGCACCAACACCAGCCGTAGCGTTCCGAGCCC GGCGAAAAGCACCCCGAACAGCAGCGTTCCGGCGTCTGCGCCGGCACC GAGCCTGACCCGCAGCCAGCTGGATCGTGTGGAAGCGCTGCTGTCTCCG GAAGATAAAATTAGCCTGAACATGGCGAAACCGTTTCGTGAACTGGAAC CGGAACTGGTGACCCGTCGTAAAAACGATTTTCAGCGCCTGTATACCAA CGATCGTGAAGATTATCTGGGCAAACTGGAACGTGATATCACCAAATTT TTTGTGGATCGCGGCTTTCTGGAAATTAAAAGCCCGATTCTGATTCCGGC GGAATATGTGGAACGTATGGGCATTAACAACGACACCGAACTGAGCAA ACAAATTTTCCGCGTGGATAAAAACCTGTGCCTGCGTCCGATGCTGGCT CCGACCCTGGGGAACTATCTTCGTAAACTGGATCGTATTCTGCCGGGTC CGATCAAAATTTTTGAAGTGGGCCCGTGCTATCGCAAAGAAAGCGATGG CAAAGAACACCTGGAAGAATTCACCATGGTTAACTTTGTGCAAATGGGC AGCGGCTGCACCCGTGAAAACCTGGAAGCGCTGATCAAAGAATTCCTGG ATTATCTGGAAATCGACTTCGAAATTGTGGGCGATAGCTGCATGGTGTA TGGCGATACCCTGGATATTATGCATGGCGATCTGGAACTGAGCAGCGCG GTGGTGGGTCCGGTTAGCCTGGATCGTGAATGGGGCATTGATAAACCGT GGATTGGCGCGGGTTTTGGCCTGGAACGTCTGCTGAAAGTGATGCATGG CTTCAAAAACATTAAACGTGCGAGCCGTAGCGAAAGCTACTATAACGGC ATTAGCACGAACCTGTAACGTTTCAAACGCTAAATTGCCTGATGCGCTA CGCTTATCAGGCCTACATGATCTCTGCAACACACCGAGCCCGCGTGCTTT TGCAGGCCGGATAAGGCGTTCGCGCCGCATCCGGCAAGAAACAGCAAA CAATCCAAAACGCCGCGTTCGGCGGCGTTTTTTCTGCTTTTCTTCGCGAA TTAATTCCGCTTCGCAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG AACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC CGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTC TCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTC AGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCC CCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG GTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCT CTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCT ACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGG TCATGATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATA AATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTC CGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCT CACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGT GCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTG AGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTT CTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGCCAAGAGCAAC TCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCA GTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGC AGTGCTGCCATAACCATGAGTGATAACACTGCAGCCAACTTACTTCTGA CAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGG GGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCC ATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACA ACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGC AACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCT

	GCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCG GTGAGCGTGGCTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATG GATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGC ATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTA AAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAA TCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAG AGGATC
pPylt_sfGFP150TAG	CCCCGGGAATCTAACCCGGCTGAACGGATTTAGAGTCCATTCGATCTAC ATGATCAGGTTCCCGCGGCCGCGAATTCAGCGTTACAAGTATTACACAA AGTTTTTTATGTTGAGAATATTTTTTTGATGGGGCGCCACTTATTTTTGAT CGTTCGCTCAAAGAAGCGGCGCCAGGGTTGTTTTTCTTTTCACCAGTGAG ACGGGCAACAGAACGCCATGAGCGGCCTCATTTCTTATTCTGAGTTACA ACAGTCCGCACCGCTGCCGGTAGCTCCTTCCGGTGGGCGCGGGGCATGA CTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGA CAGGTGCCGGCAGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCA TACCCACGCCGAAACAAGCGCCCTGCACCATTATGTTCCGGATCTGCAT CGCAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATTAACG AAGCGCTAACCGTTTTTATCATGCTCTGGGAGGCAGAATAAATGATCAT ATCGTCAATTATTACCTCCACGGGGAGAGCCTGAGCAAACTGGCCTCAG GCATTTGAGAAGCACACGGTCACACTGCTTCCGGTAGTCAATAAACCGG TAAACCAGCAATAGACATAAGCGGCTATTTAACGACCCTGCCCTGAACC GACGACCGGGTCGAATTTGCTTTCGAATTTCTGCCATTCATCCGCTTATT ATCACTTATTCAGGCGTAGCAACCAGGCGTTTAAGGGCACCAATAACTG CCTTAAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTTGACTGGGT CATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGG GCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGG GAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGG CAGCAGATCAATTCGCGCGCGAAGGCGAAGCGGCATGCATAATGTGCCT GTCAAATGGACGAAGCAGGGATTCTGCAAACCCTATGCTACTCCGTCAA GCCGTCAATTGTCTGATTCGTTACCAATTATGACAACTTGACGGCTACAT CATTCACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTGGCCCCG GTGCATTTTTTAAATACCCGCGAGAAATAGAGTTGATCGTCAAAACCAA CATTGCGACCGACGGTGGCGATAGGCATCCGGGTGGTGCTCAAAAGCA GCTTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATC CCTAACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAAGCAA ACATGCTGTGCGACGCTGGCGATATCAAAATTGCTGTCTGCCAGGTGAT CGCTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGGTGGATG GAGCGACTCGTTAATCGCTTCCATGCGCCGCAGTAACAATTGCTCAAGC AGATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGCCCGGCGTT AATGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCTTCATCC GGGCGAAAGAACCCCGTATTGGCAAATATTGACGGCCAGTTAAGCCATT CATGCCAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTGATACCATTC GCGAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGCGGGAAC AGCAAAATATCACCCGGTCGGCAAACAAATTCTCGTCCCTGATTTTTCA CCACCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATTCC CAGCGGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGC GTTAAACCCGCCACCAGATGGGCATTAAACGAGTATCCCGGCAGCAGG GGATCATTTTGCGCTTCAGCCATACTTTTCATACTCCCGCCATTCAGAGA AGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTT ACTGGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTTATTAAAAGCATT CTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGT GTCTATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTCA CACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACC TGACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTGGG CTAACAGGAGGAATTAACCATGGTTAGCAAAGGTGAAGAACTGTTTACC GGCGTTGTGCCGATTCTGGTGGAACTGGATGGTGATGTGAATGGCCATA AATTTAGCGTTCGTGGCGAAGGCGAAGGTGATGCGACCAACGGTAAACT GACCCTGAAATTTATTTGCACCACCGGTAAACTGCCGGTTCCGTGGCCG ACCCTGGTGACCACCCTGACCTATGGCGTTCAGTGCTTTAGCCGCTATCC GGATCATATGAAACGCCATGATTTCTTTAAAAGCGCGATGCCGGAAGGC

TATGTGCAGGAACGTACCATTAGCTTCAAAGATGATGGCACCTATAAAA
CCCGTGCGGAAGTTAAATTTGAAGGCGATACCCTGGTGAACCGCATTGA
ACTGAAAGGTATTGATTTTAAAGAAGATGGCAACATTCTGGGTCATAAA
CTGGAATATAATTTCAACAGCCATTAGGTGTATATTACCGCCGATAAAC
AGAAAAATGGCATCAAAGCGAACTTTAAAATCCGTCACAACGTGGAAG
ATGGTAGCGTGCAGCTGGCGGATCATTATCAGCAGAATACCCCGATTGG
TGATGGCCCGGTGCTGCTGCCGGATAATCATTATCTGAGCACCCAGAGC
GTTCTGAGCAAAGATCCGAATGAAAAACGTGATCATATGGTGCTGCTGG
AATTTGTTACCGCCGCGGGCATTACCCACGGTATGGATGAACTGTATAA
AGGCAGCCACCATCATCATCACCATTAAAGCTCGAGCGAAGCTTGGGCC
CGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCAT
CATCATCATCATCATTGAGTTTAAACGGTCTCCAGCTTGGCTGTTTTGGC
GGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGA
AGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCC
CACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGG
TAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAAT
AAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGT
TTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCTGTCC
CTCCTGTTCAGCTACTGACGGGGTGGTGCGTAACGGCAAAAGCACCGCC
GGACATCAGCGCTAGCGGAGTGTATACTGGCTTACTATGTTGGCACTGA
TGAGGGTGTCAGTGAAGTGCTTCATGTGGCAGGAGAAAAAAGGCTGCA
CCGGTGCGTCAGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGC
TCACTGACTCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGC
TTACGAACGGGGCGGAGATTTCCTGGAAGATGCCAGGAAGATACTTAAC
AGGGAAGTGAGAGGGCCGCGGCAAAGCCGTTTTTCCATAGGCTCCGCCC
CCCTGACAAGCATCACGAAATCTGACGCTCAAATCAGTGGTGGCGAAAC
CCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGCGGCTCCCTCG
TGCGCTCTCCTGTTCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTAT
GGCCGCGTTTGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGT
TCGCTCCAAGCTGGACTGTATGCACGAACCCCCCGTTCAGTCCGACCGC
TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGAAAGACATG
CAAAAGCACCACTGGCAGCAGCCACTGGTAATTGATTTAGAGGAGTTAG
TCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGG
TGACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCT
CAGAGAACCTTCGAAAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAG
AGCAAGAGTTTACGCGCAGACCAAAACGATCTCAAGAAGATCATCTTAT
TAATCAGATAAAATATTTCTAGATTTCAGTGCAATTTATCTCTTCAAATG
TAGCACCTGAAGTCAGCCCCATACGATATAAGTTGTAATTCTCATGTTTG
ACAGCTTATCATCGATAAGCTTTAATGCGGTAGTTTATCACAGTTAAATT
GCTAACGCAGTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATCGT
CATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATG
CCGGTACTGCCGGGCCTCTTGCGGGATATCGTCCATTCCGACAGCATCG
CCAGTCACTATGGCGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTA
TGCGCACCCGTTCTCGGAGCACTGTCCGACCGCTTTGGCCGCCGCCCAG
TCCTGCTCGCTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCG
ACCACACCCGTCCTGTGGATCCTCTACGCCGGACGCATCGTGGCCGGCA
TCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCAC
CGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTC
GGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCC
ATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAA
CCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGT
CGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGG
CGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATG
CAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCGGCGAGG
ACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTGCGGTATT
CGGAATCTTGCACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCA
AACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGCCGACGC
GCTGGGCTACGTCTTGCTGGCGTTCGCGACGCGAGGCTGGATGGCCTTC
CCCATTATGATTCTTCTCGCTTCCGGCGGCATCGGGATGCCCGCGTTGCA
GGCCATGCTGTCCAGGCAGGTAGATGACGACCATCAGGGACAGCTTCAA
GGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCATTGGACCGCTGAT
CGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCA

	TGGATTGTAGGCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTTGCGTCG CGGTGCATGGAGCCGGGCCACCTCGACCTGAATGGAAGCCGGCGGCAC CTCGCTAACGGATTCACCACTCCAAGAATTGGAGCCAATCAATTCTTGC GGAGAACTGTGAATGCGCAAACCAACCCTTGGCAGAACATATCCATCGC GTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCTCCTTGCAT GCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGG GCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCTGGCGAAAG GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCA GTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTTAAAAAAAATCCT TAGCTTTCGCTAAGATCTGCAGTGGCGGAAA
pEVOL_*Mm* MF3	TCCTGAAAATCTCGATAACTCAAAAAATACGCCCGGTAGTGATCTTATT TCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATTT TCGCCAAAAGTTGGCCCAGGGCTTCCCGGTATCAACAGGGACACCAGGA TTTATTTATTCTGCGAAGTGATCTTCCGTCACAGGTATTTATTCGGCGCA AAGTGCGTCGGGTGATGCTGCCAACTTACTGATTTAGTGTATGATGGTG TTTTTGAGGTGCTCCAGTGGCTTCTGTTTCTATCAGCTGTCCCTCCTGTTC AGCTACTGACGGGGTGGTGCGTAACGGCAAAAGCACCGCCGGACATCA GCGCTAGCGGAGTGTATACTGGCTTACTATGTTGGCACTGATGAGGGTG TCAGTGAAGTGCTTCATGTGGCAGGAGAAAAAAGGCTGCACCGGTGCGT CAGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCTCACTGACTC GCTACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGG GGCGGAGATTTCCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTG AGAGGGCCGCGGCAAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAA GCATCACGAAATCTGACGCTCAAATCAGTGGTGGCGAAACCCGACAGG ACTATAAAGATACCAGGCGTTTCCCCCTGGCGGCTCCCTCGTGCGCTCTC CTGTTCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTATGGCCGCGTT TGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAA GCTGGACTGTATGCACGAACCCCCCGTTCAGTCCGACCGCTGCGCCTTA TCCGGTAACTATCGTCTTGAGTCCAACCCGGAAAGACATGCAAAAGCAC CACTGGCAGCAGCCACTGGTAATTGATTTAGAGGAGTTAGTCTTGAAGT CATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGGTGACTGCGC TCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACC TTCGAAAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGCAAGAGA TTACGCGCAGACCAAAACGATCTCAAGAAGATCATCTTATTAATCAGAT AAAATATTTCTAGATTTCAGTGCAATTTATCTCTTCAAATGTAGCACCTG AAGTCAGCCCCATACGATATAAGTTGTAATTCTCATGTTTGACAGCTTAT CATCGATAAGCTTGGTACCCAATTATGACAACTTGACGGCTACATCATT CACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTGGCCCCGGTG CATTTTTTAAATACCCGCGAGAAATAGAGTTGATCGTCAAAACCAACAT TGCGACCGACGGTGGCGATAGGCATCCGGGTGGTGCTCAAAAGCAGCTT CGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATCCCT AACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAAGCAAACA TGCTGTGCGACGCTGGCGATATCAAAATTGCTGTCTGCCAGGTGATCGC TGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGGTGGATGGAG CGACTCGTTAATCGCTTCCATGCGCCGCAGTAACAATTGCTCAAGCAGA TTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGCCCGGCGTTAAT GATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCTTCATCCGGG CGAAAGAACCCCGTATTGGCAAATATTGACGGCCAGTTAAGCCATTCAT GCCAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTGATACCATTCGC GAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGCGGGAACAG CAAAATATCACTCGGTCGGCAAACAAATTCTCGTCCCTGATTTTTCACCA CCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATTCCCAG CGGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGTT AAACCCGCCACCAGATGGGCATTAAACGAGTATCCCGGCAGCAGGGGA TCATTTTGCGCTTCAGCCATACTTTTCATACTCCCGCCATTCAGAGAAGA AACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTTACT GGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTTATTAAAAGCATTCTG TAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTC TATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTCACAC TTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGA CGCTTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTTGGGCT

AACAGGAGGAATTAGATCTATGGACAAAAAACCGCTGAATACCCTGATT
AGCGCAACCGGTCTGTGGATGAGCCGTACCGGCACCATTCATAAAATTA
AACATCATGAAGTGAGCCGCAGCAAAATTTATATTGAAATGGCATGTGG
CGATCATCTGGTTGTTAATAATAGCCGTAGCAGCCGTACCGCACGTGCA
CTGCGTCATCATAAATATCGTAAAACCTGCAAACGTTGCCGTGTTAGTG
ATGAAGATCTGAATAAATTTCTGACCAAAGCCAATGAAGATCAGACCAG
CGTTAAAGTTAAAGTTGTTAGCGCACCGACCCGTACCAAAAAAGCAATG
CCGAAAAGCGTTGCACGTGCACCGAAACCGCTGGAAAATACCGAAGCA
GCACAGGCACAGCCGAGCGGTAGCAAATTTTCACCGGCAATTCCGGTTA
GCACCCAAGAAAGCGTTAGCGTTCCGGCAAGCGTTAGCACCAGCATTAG
CAGCATTTCAACCGGTGCAACCGCAAGCGCACTGGTTAAAGGTAATACC
AATCCGATTACCAGCATGAGCGCACCGGTTCAGGCAAGCGCACCGGCAC
TGACCAAAAGCCAGACCGATCGTCTGGAAGTTCTGCTGAATCCGAAAGA
TGAAATTAGCCTGAATAGCGGTAAACCGTTTCGTGAACTGGAAAGCGAA
CTGCTGAGCCGTCGTAAAAAAGATCTGCAACAAATTTATGCCGAAGAAC
GTGAAAATTATCTGGGTAAACTGGAACGCGAAATTACCCGTTTTTTTGTT
GATCGTGGCTTTCTGGAAATTAAAAGCCCGATTCTGATTCCGCTGGAAT
ATATTGAACGCATGGGCATTGATAATGATACCGAACTGAGCAAACAAAT
TTTTCGCGTGGATAAAAATTTTTGTCTGCGTCCGATGCTGGCACCGAATC
TGGGCAATTATCTGCGCAAACTGGATCGTGCACTGCCGGATCCGATTAA
AATTTTTGAAATTGGTCCGTGCTATCGCAAAGAAAGTGATGGTAAAGAA
CATCTGGAAGAATTTACCATGCTGAATTTTGTGCAGATGGGTAGCGGTT
GTACCCGTGAAAATCTGGAAAGCATTATTACCGATTTTCTGAATCATCTG
GGCATTGATTTTAAAATTGTGGGTGATAGCTGCATGGTGTATGGTGATA
CCCTGGATGTTATGCATGGTGATCTGGAACTGAGCAGCGCAGTTGTTGG
TCCGATTCCGCTGGATCGTGAATGGGGTATTGATAAACCGTGGATTGGT
GCAGGTTTTGGTCTGGAACGCCTGCTGAAAGTTAAACATGATTTTAAAA
ATATTAAACGTGCCGCACGCAGCGAAAGCTATTACAATGGTATTAGCAC
CAATCTGTAAGTCGACCATCATCATCATCATTGAGTTTAAACGGTCT
CCAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAG
ATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCG
GCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAA
ACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGG
AACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGC
CTTGTTTGTGAGCTCCCGGTCATCAATCATCCCCATAATCCTTGTTAGAT
TATCAATTTTAAAAAACTAACAGTTGTCAGCCTGTCCCGCTTTAATATCA
TACGCCGTTATACGTTGTTTACGCTTTGAGGAATCCCATATGGATAAAA
AACCGCTGAATACCCTGATTAGCGCAACCGGTCTGTGGATGAGCCGTAC
CGGCACCATTCATAAAATTAAACATCATGAAGTGAGCCGCAGCAAAATT
TATATTGAAATGGCATGTGGCGATCATCTGGTTGTTAATAATAGCCGTA
GCAGCCGTACCGCACGTGCACTGCGTCATCATAAATATCGTAAAACCTG
CAAACGTTGCCGTGTTAGTGATGAAGATCTGAATAAATTTCTGACCAAA
GCCAATGAAGATCAGACCAGCGTTAAAGTTAAAGTTGTTAGCGCACCGA
CCCGTACCAAAAAAGCAATGCCGAAAAGCGTTGCACGTGCACCGAAAC
CGCTGGAAAATACCGAAGCAGCACAGGCACAGCCGAGCGGTAGCAAAT
TTTCACCGGCAATTCCGGTTAGCACCCAAGAAAGCGTTAGCGTTCCGGC
AAGCGTTAGCACCAGCATTAGCAGCATTTCAACCGGTGCAACCGCAAGC
GCACTGGTTAAAGGTAATACCAATCCGATTACCAGCATGAGCGCACCGG
TTCAGGCAAGCGCACCGGCACTGACCAAAAGCCAGACCGATCGTCTGG
AAGTTCTGCTGAATCCGAAAGATGAAATTAGCCTGAATAGCGGTAAACC
GTTTCGTGAACTGGAAAGCGAACTGCTGAGCCGTCGTAAAAAAGATCTG
CAACAAATTTATGCCGAAGAACGTGAAAATTATCTGGGTAAACTGGAAC
GCGAAATTACCCGTTTTTTTGTTGATCGTGGCTTTCTGGAAATTAAAAGC
CCGATTCTGATTCCGCTGGAATATATTGAACGCATGGGCATTGATAATG
ATACCGAACTGAGCAAACAAATTTTTCGCGTGGATAAAAATTTTTGTCT
GCGTCCGATGCTGGCACCGAATCTGGGCAATTATCTGCGCAAACTGGAT
CGTGCACTGCCGGATCCGATTAAAATTTTTGAAATTGGTCCGTGCTATCG
CAAAGAAAGTGATGGTAAAGAACATCTGGAAGAATTTACCATGCTGAA
TTTTGTGCAGATGGGTAGCGGTTGTACCCGTGAAAATCTGGAAAGCATT
ATTACCGATTTTCTGAATCATCTGGGCATTGATTTTAAAATTGTGGGTGA
TAGCTGCATGGTGTATGGTGATACCCTGGATGTTATGCATGGTGATCTG
GAACTGAGCAGCGCAGTTGTTGGTCCGATTCCGCTGGATCGTGAATGGG

	GTATTGATAAACCGTGGATTGGTGCAGGTTTTGGTCTGGAACGCCTGCT GAAAGTTAAACATGATTTTAAAAATATTAAACGTGCCGCACGCAGCGAA AGCTATTACAATGGTATTAGCACCAATCTGTAACTGCAGTTTCAAACGC TAAATTGCCTGATGCGCTACGCTTATCAGGCCTACATGATCTCTGCAATA TATTGAGTTTGCGTGCTTTTGTAGGCCGGATAAGGCGTTCACGCCGCATC CGGCAAGAAACAGCAAACAATCCAAAACGCCGCGTTCAGCGGCGTTTTT TCTGCTTTTCTTCGCGAATTAATTCCGCTTCGCAACATGTGAGCACCGGT TTATTGACTACCGGAAGCAGTGTGACCGTGTGCTTCTCAAATGCCTGAG GCCAGTTTGCTCAGGCTCTCCCCGTGGAGGTAATAATTGACGATATGAT CAGTGCACGGCTAACTAAGCGGCCTGCTGACTTTCTCGCCGATCAAAAG GCATTTTGCTATTAAGGGATTGACGAGGGCGTATCTGCGCAGTAAGATG CGCCCCGCATTGGAAACCTGATCATGTAGATCGAATGGACTCTAAATCC GTTCAGCCGGGTTAGATTCCCGGGGTTTCCGCCAAATTCGAAAAGCCTG CTCAACGAGCAGGCTTTTTTGCATGCTCGAGCAGCTCAGGGTCGAATTT GCTTTCGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTA GCACCAGGCGTTTAAGGGCACCAATAACTGCCTTAAAAAAATTACGCCC CGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTCTGCCG ACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGC ATCAGCACCTTGTCGCCTTGCGTATAATATTTGCCCATGGTGAAAACGG GGGCGAAGAAGTTGTCCATATTGGCCACGTTTAAATCAAAACTGGTGAA ACTCACCCAGGGATTGGCTGAGACGAAAAACATATTCTCAATAAACCCT TTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTGCGAAT ATATGTGTAGAAACTGCCGGAAATCGTCGTGGTATTCACTCCAGAGCGA TGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAACAAGGGTGAACA CTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAATTCCGG ATGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAA CTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCT GAACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAA ATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCCAGTGA TTTTTTTCTCCATTTTAGCTTCCTTAGC
pEVOL_*Mj* (SM44)	TATCATCGATAAGCTTGGTACCCAATTATGACAACTTGACGGCTACATC ATTCACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTGGCCCCGG TGCATTTTTTAAATACCCGCGAGAAATAGAGTTGATCGTCAAAACCAAC ATTGCGACCGACGGTGGCGATAGGCATCCGGGTGGTGCTCAAAAGCAG CTTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATCC CTAACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAAGCAAA CATGCTGTGCGACGCTGGCGATATCAAAATTGCTGTCTGCCAGGTGATC GCTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGGTGGATGG AGCGACTCGTTAATCGCTTCCATGCGCCGCAGTAACAATTGCTCAAGCA GATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGCCCGGCGTTA ATGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCTTCATCCG GGCGAAAGAACCCCGTATTGGCAAATATTGACGGCCAGTTAAGCCATTC ATGCCAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTGATACCATTCG CGAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGCGGGAACA GCAAAATATCACTCGGTCGGCAAACAAATTCTCGTCCCTGATTTTTCACC ACCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTCATTCCCA GCGGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCGT TAAACCCGCCACCAGATGGGCATTAAACGAGTATCCCGGCAGCAGGGG ATCATTTTGCGCTTCAGCCATACTTTTCATACTCCCGCCATTCAGAGAAG AAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTTAC TGGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTTATTAAAAGCATTCT GTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGT CTATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTCACA CTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTG ACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTTGGGC TAACAGGAGGAATTAGATCTATGGACGAATTTGAAATGATAAAGAGAA ACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAGAGGTTTTAAAAA AAGATGAAAAATCTGCTGGGATAGGTTTTGAACCAAGTGGTAAAATACA TTTAGGGCATTATCTCCAAATAAAAAAGATGATTGATTTACAAAATGCT GGATTTGATATAATTATACAGTTGGCTGATTTACACGCCTATTTAAACCA GAAAGGAGAGTTGGATGAGATTAGAAAAATAGGAGATTATAACAAAAA

AGTTTTTGAAGCAATGGGGTTAAAGGCAAAATATGTTTATGGAAGTGAA
TCTGATCTTGATAAGGATTATACACTGAATGTCTATAGATTGGCTTTAAA
AACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTTATAGCAAGAGA
GGATGAAAATCCAAAGGTTGCTGAAGTTATCTATCCAATAATGCAGGTT
AATTCTATTCATTATAATGGCGTTGATGTTGCAGTTGGAGGGATGGAGC
AGAGAAAAATACACATGTTAGCAAGGGAGCTTTTACCAAAAAAGGTTG
TTTGTATTCACAACCCTGTCTTAACGGGTTTGGATGGAGAAGGAAAGAT
GAGTTCTTCAAAAGGGAATTTTATAGCTGTTGATGACTCTCCAGAAGAG
ATTAGGGCTAAGATAAAGAAAGCATACTGCCCAGCTGGAGTTGTTGAAG
GAAATCCAATAATGGAGATAGCTAAATACTTCCTTGAATATCCTTTAAC
CATAAAAAGGCCAGAAAAATTTGGTGGAGATTTGACAGTTAATAGCTAT
GAGGAGTTAGAGAGTTTATTTAAAAATAAGGAATTGCATCCAATGGATT
TAAAAAATGCTGTAGCTGAAGAACTTATAAAGATTTTAGAGCCAATTAG
AAAGAGATTATAAGTCGACCATCATCATCATCATCATTGAGTTTAAACG
GTCTCCAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGAT
ACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCT
GGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAG
TGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGT
AGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACT
GGGCCTTGTTTGTGAGCTCCCGGTCATCAATCATCCCCATAATCCTTGTT
AGATTATCAATTTTAAAAAACTAACAGTTGTCAGCCTGTCCCGCTTTAAT
ATCATACGCCGTTATACGTTGTTTACGCTTTGAGGAATCCCATATGGACG
AATTTGAAATGATAAAGAGAAACACATCTGAAATTATCAGCGAGGAAG
AGTTAAGAGAGGTTTTAAAAAAAGATGAAAAATCTGCTGGGATAGGTTT
TGAACCAAGTGGTAAAATACATTTAGGGCATTATCTCCAAATAAAAAAG
ATGATTGATTTACAAAATGCTGGATTTGATATAATTATACAGTTGGCTGA
TTTACACGCCTATTTAAACCAGAAAGGAGAGTTGGATGAGATTAGAAAA
ATAGGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTTAAAGGCA
AAATATGTTTATGGAAGTGAATCTGATCTTGATAAGGATTATACACTGA
ATGTCTATAGATTGGCTTTAAAAACTACCTTAAAAAGAGCAAGAAGGAG
TATGGAACTTATAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGTT
ATCTATCCAATAATGCAGGTTAATTCTATTCATTATAATGGCGTTGATGT
TGCAGTTGGAGGGATGGAGCAGAGAAAAATACACATGTTAGCAAGGGA
GCTTTTACCAAAAAAGGTTGTTTGTATTCACAACCCTGTCTTAACGGGTT
TGGATGGAGAAGGAAAGATGAGTTCTTCAAAAGGGAATTTTATAGCTGT
TGATGACTCTCCAGAAGAGATTAGGGCTAAGATAAAGAAAGCATACTG
CCCAGCTGGAGTTGTTGAAGGAAATCCAATAATGGAGATAGCTAAATAC
TTCCTTGAATATCCTTTAACCATAAAAAGGCCAGAAAAATTTGGTGGAG
ATTTGACAGTTAATAGCTATGAGGAGTTAGAGAGTTTATTTAAAAATAA
GGAATTGCATCCAATGGATTTAAAAAATGCTGTAGCTGAAGAACTTATA
AAGATTTTAGAGCCAATTAGAAAGAGATTATAACTGCAGTTTCAAACGC
TAAATTGCCTGATGCGCTACGCTTATCAGGCCTACATGATCTCTGCAATA
TATTGAGTTTGCGTGCTTTTGTAGGCCGGATAAGGCGTTCACGCCGCATC
CGGCAAGAAACAGCAAACAATCCAAAACGCCGCGTTCAGCGGCGTTTTT
TCTGCTTTTCTTCGCGAATTAATTCCGCTTCGCAACATGTGAGCACCGGT
TTATTGACTACCGGAAGCAGTGTGACCGTGTGCTTCTCAAATGCCTGAG
GCCAGTTTGCTCAGGCTCTCCCCGTGGAGGTAATAATTGACGATATGAT
CAGTGCACGGCTAACTAAGCGGCCTGCTGACTTTCTCGCCGATCAAAAG
GCATTTTGCTATTAAGGGATTGACGAGGGCGTATCTGCGCAGTAAGATG
CGCCCCGCATTCCGGCGGTAGTTCAGCAGGGCAGAACGGCGGACTCTAA
ATCCGCATGGCAGGGGTTCAAATCCCCTCCGCCGGACCAAATTCGAAAA
GCCTGCTCAACGAGCAGGCTTTTTTGCATGCTCGAGCAGCTCAGGGTCG
AATTTGCTTTCGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGG
CGTAGCAACCAGGCGTTTAAGGGCACCAATAACTGCCTTAAAAAAATTA
CGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAGCATTC
TGCCGACATGGAAGCCATCACAAACGGCATGATGAACCTGAATCGCCA
GCGGCATCAGCACCTTGTCGCCTTGCGTATAATATTTGCCCATGGTGAA
AACGGGGGCGAAGAAGTTGTCCATATTGGCCACGTTTAAATCAAAACTG
GTGAAACTCACCCAGGGATTGGCTGAGACGAAAAACATATTCTCAATAA
ACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACATCTTG
CGAATATATGTGTAGAAACTGCCGGAAATCGTCGTGGTATTCACTCCAG
AGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAACAAGGGT

	GAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAAT TCCGGATGAGCATTCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGA TAAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATC CAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCC TCAAAATGTTCTTTACGATGCCATTGGGATATATCAACGGTGGTATATCC AGTGATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAATCTCGATAA CTCAAAAAATACGCCCGGTAGTGATCTTATTTCATTATGGTGAAAGTTG GAACCTCTTACGTGCCGATCAACGTCTCATTTTCGCCAAAAGTTGGCCCA GGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTATTCTGCGAAG TGATCTTCCGTCACAGGTATTTATTCGGCGCAAAGTGCGTCGGGTGATG CTGCCAACTTACTGATTTAGTGTATGATGGTGTTTTTGAGGTGCTCCAGT GGCTTCTGTTTCTATCAGCTGTCCCTCCTGTTCAGCTACTGACGGGGTGG TGCGTAACGGCAAAAGCACCGCCGGACATCAGCGCTAGCGGAGTGTAT ACTGGCTTACTATGTTGGCACTGATGAGGGTGTCAGTGAAGTGCTTCAT GTGGCAGGAGAAAAAAGGCTGCACCGGTGCGTCAGCAGAATATGTGAT ACAGGATATATTCCGCTTCCTCGCTCACTGACTCGCTACGCTCGGTCGTT CGACTGCGGCGAGCGGAAATGGCTTACGAACGGGGCGGAGATTTCCTG GAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGCGGCAA AGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCATCACGAAATCTG ACGCTCAAATCAGTGGTGGCGAAACCCGACAGGACTATAAAGATACCA GGCGTTTCCCCCTGGCGGCTCCCTCGTGCGCTCTCCTGTTCCTGCCTTTC GGTTTACCGGTGTCATTCCGCTGTTATGGCCGCGTTTGTCTCATTCCACG CCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACTGTATG CACGAACCCCCCGTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATC GTCTTGAGTCCAACCCGGAAAGACATGCAAAAGCACCACTGGCAGCAG CCACTGGTAATTGATTTAGAGGAGTTAGTCTTGAAGTCATGCGCCGGTT AAGGCTAAACTGAAAGGACAAGTTTTGGTGACTGCGCTCCTCCAAGCCA GTTACCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACCTTCGAAAAACCG CCCTGCAAGGCGGTTTTTTCGTTTTCAGAGCAAGAGATTACGCGCAGAC CAAAACGATCTCAAGAAGATCATCTTATTAATCAGATAAAATATTTCTA GATTTCAGTGCAATTTATCTCTTCAAATGTAGCACCTGAAGTCAGCCCCA TACGATATAAGTTGTAATTCTCATGTTTGACAGCT
pDule2_*Mj* (SM16)	AACTCAGAATAAGAAATGAGGCCGCTCATGGCGTTCTGTTGCCCGTCTC ACCGGTGAAAAGAAAAACAACCCTGGCGCCGCTTCTTTGAGCGAACGAT CAAAAATAAGTGGCGCCCCATCAAAAAAATATTCTCAACATAAAAAACT TTGTGTAATACTTGTAACGCTGAGTTTACGCTTTGAGGAATCCCCCATGG ACGAATTTGAAATGATAAAGAGAAACACATCTGAAATTATCAGCGAGG AAGAGTTAAGAGAGGTTTTAAAAAAAGATGAAAAATCTGCTGGGATAG GTTTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTCCAAATAAA AAAGATGATTGATTTACAAAATGCTGGATTTGATATAATTATACAGTTG GCTGATTTACACGCCTATTTAAACCAGAAAGGAGAGTTGGATGAGATTA GAAAAATAGGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTTAA AGGCAAAATATGTTTATGGAAGTGAATCTGATCTTGATAAGGATTATAC ACTGAATGTCTATAGATTGGCTTTAAAAACTACCTTAAAAAGAGCAAGA AGGAGTATGGAACTTATAGCAAGAGAGGATGAAAATCCAAAGGTTGCT GAAGTTATCTATCCAATAATGCAGGTTAATTCTATTCATTATAATGGCGT TGATGTTGCAGTTGGAGGGATGGAGCAGAGAAAAATACACATGTTAGC AAGGGAGCTTTTACCAAAAAAGGTTGTTTGTATTCACAACCCTGTCTTA ACGGGTTTGGATGGAGAAGGAAAGATGAGTTCTTCAAAAGGGAATTTTA TAGCTGTTGATGACTCTCCAGAAGAGATTAGGGCTAAGATAAAGAAAGC ATACTGCCCAGCTGGAGTTGTTGAAGGAAATCCAATAATGGAGATAGCT AAATACTTCCTTGAATATCCTTTAACCATAAAAAGGCCAGAAAAATTTG GTGGAGATTTGACAGTTAATAGCTATGAGGAGTTAGAGAGTTTATTTAA AAATAAGGAATTGCATCCAATGGATTTAAAAAATGCTGTAGCTGAAGAA CTTATAAAGATTTTAGAGCCAATTAGAAAGAGATTATAACTGCAGTTTC AAACGGGTACCATATGGGAATTCGAAGCTTGGGCCCGAACAAAAACTC ATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATC ATTGAGTTTAAACGGTCTCCAGCTTGGCTGTTTTGGCGGATGAGAGAAG ATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAA AACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCAT GCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCT

CCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGC
TCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACG
CTCTCCTGAGTAGGACAAATCCGCCGGGAGCTGTCCCTCCTGTTCAGCT
ACTGACGGGGTGGTGCGTAACGGCAAAAGCACCGCCGGACATCAGCGC
TAGCGGAGTGTATACTGGCTTACTATGTTGGCACTGATGAGGGTGTCAG
TGAAGTGCTTCATGTGGCAGGAGAAAAAAGGCTGCACCGGTGCGTCAG
CAGAATATGTGATACAGGATATATTCCGCTTCCTCGCTCACTGACTCGCT
ACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGGGGC
GGAGATTTCCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGA
GGGCCGCGGCAAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCA
TCACGAAATCTGACGCTCAAATCAGTGGTGGCGAAACCCGACAGGACTA
TAAAGATACCAGGCGTTTCCCCCTGGCGGCTCCCTCGTGCGCTCTCCTGT
TCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTATGGCCGCGTTTGTC
TCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAAGCT
GGACTGTATGCACGAACCCCCCGTTCAGTCCGACCGCTGCGCCTTATCC
GGTAACTATCGTCTTGAGTCCAACCCGGAAAGACATGCAAAAGCACCAC
TGGCAGCAGCCACTGGTAATTGATTTAGAGGAGTTAGTCTTGAAGTCAT
GCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGGTGACTGCGCTCC
TCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACCTTC
GAAAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGCAAGAGATTA
CGCGCAGACCAAAACGATCTCAAGAAGATCATCTTATTAATCAGATAAA
ATATTTCTAGATTTCAGTGCAATTTATCTCTTCAAATGTAGCACCTGAAG
TCAGCCCCATACGATATAAGTTGTAATTCTCATGTTTGACAGCTTATCAT
CGATGCGACCGAGTGAGCTAGCTATTTGTTTATTTTTCTAAATACATTCA
AATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAAT
ATTGAAAAAGGAAGAGTATGAGGGAAGCGGTGATCGCCGAAGTATCGA
CTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGAACCGAC
GTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCCTGAAG
CCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTTGATG
AAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAAACTTCGGCTTC
CCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTG
CACGACGACATCATTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAAT
TTGGAGAATGGCAGCGCAATGACATTCTTGCAGGTATCTTCGAGCCAGC
CACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACAT
AGCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTC
CTGAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGGAA
CTCGCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTG
TCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATG
TCGCTGCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGT
CATACTTGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTG
GCCTCGCGCGCAGATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGCG
AGATCACCAAGGTAGTCGGCAAATAATGTCTAACAATTCGTTCAAGCCG
AGGGGCCGCAAGATCCGGCCACGATGACCCGGTCGTCGGTTCAGGGCA
GGGTCGTTAAATAGCCGCTTATGTCTATTGCTGGTTTACCGGGATCCTCT
ACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGC
TGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCAC
TTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCG
TGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCG
GCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGC
AGGAGTCGCATAAGGGAGAGCGTCGACCGATGCCCTTGAGAGCCTTCA
ACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGC
ACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAG
CGCTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGAT
GATCGGCCTGTCGCTTGCGGTATTCGGAATCTTGCACGCCCTCGCTCAAG
CCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCAT
TATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGCGTTC
GCGACGCGAGGCTGGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGG
CGGCATCGGGATGCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGAT
GACGACCATCAGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCC
TAACTTCGATCATTGGACCGCTGATCGTCACGGCGATTTATGCCGCCTCG
GCGAGCACATGGAACGGGTTGGCATGGATTGTAGGCGCCGCCCTATACC
TTGTCTGCCTCCCCGCGTTGCGTCGCGGTGCATGGAGCCGGGCCACCTC

	GACCTGAATGGAAGCCGGCGGCACCTCGCTAACGGATTCACCACTCCAA GAATTGGAGCCAATCAATTCTTGCGGAGAACTGTGAATGCGCAAACCAA CCCTTGGCAGAACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACG CGGCGCATCTCGGGCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCT CAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCAT AAGGGAGAGCGTCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGT TGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCA GTGCCAAGCTTAAAAAAAATCCTTAGCTTTCGCTAAGGATCTGCAGTGG TCCGGCGGGCCGGATTTGAACCAGCGCCATGCGGATTTAGAGTCCGCCG TTCTGCCCTGCTGAACTACCGCCGGGAATTCAGCGTTACAAGTATTACA CAAAGTTTTTTATGTTGAGAATATTTTTTTGATGGGGCGCCACTTATTTTT GATCGTTCGCTCAAAGAAGCGGCGCCAGGGTTGTTTTTCTTTTCACCAGT GAGACGGGCAACAGAACGCCATGAGCGGCCTCATTTCTTATTCTGAGTT ACAACAGTCCGCACCGCTGCCGGTAGCTCCTTCCGGTGGGCGCGGGGCA TGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTA GGACAGGTGCCGGCAGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCC ACCATACCCACGCCGAAACAAGCGCCCTGCACCATTATGTTCCGGATCT GCATCGCAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATT AACGAAGCGCTAACCGTTTTTATCATGCTCTGGGAGGCAGAATAAATGA TCATATCGTCAATTATTACCTCCACGGGGAGAGCCTGAGCAAACTGGCC TCAGGCATTTGAGAAGCACACGGTCACACTGCTTCCGGTAGTCAATAAA CCGGTAAACCAGCAATAGACATAAGCGGCTATTTAACGACCCTGCCCTG AACCGACGACCGGGTCGAATTTGCTTTCGAATTTCTGCCATTCATCCGCT TATTATCACTTATTCAGGCGTAGCAACCAGGCGTTTAAGGGCACCAATA ACTGCCTTAAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTTGACT GGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTG ACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTC TCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCG CGAGGCAGCAGATCAATTCGCGCGCGAAGGCGAAGCGGCATGCATAAT GTGCCTGTCAAATGGACGAAGCAGGGATTCTGCAAACCCTATGCTACTC CGTCAAGCCGTCAATTGTCTGATTCGTTACCAATTATGACAACTTGACGG CTACATCATTCACTTTTTCTTCACAACCGGCACGGAACTCGCTCGGGCTG GCCCCGGTGCATTTTTTAAATACCCGCGAGAAATAGAGTTGATCGTCAA AACCAACATTGCGACCGACGGTGGCGATAGGCATCCGGGTGGTGCTCAA AAGCAGCTTCGCCTGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACG CTAATCCCTAACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGAC AAGCAAACATGCTGTGCGACGCTGGCGATATCAAAATTGCTGTCTGCCA GGTGATCGCTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATCGG TGGATGGAGCGACTCGTTAATCGCTTCCATGCGCCGCAGTAACAATTGC TCAAGCAGATTTATCGCCAGCAGCTCCGAATAGCGCCCTTCCCCTTGCCC GGCGTTAATGATTTGCCCAAACAGGTCGCTGAAATGCGGCTGGTGCGCT TCATCCGGGCGAAAGAACCCCGTATTGGCAAATATTGACGGCCAGTTAA GCCATTCATGCCAGTAGGCGCGCGGACGAAAGTAAACCCACTGGTGATA CCATTCGCGAGCCTCCGGATGACGACCGTAGTGATGAATCTCTCCTGGC GGGAACAGCAAAATATCACCCGGTCGGCAAACAAATTCTCGTCCCTGAT TTTTCACCACCCCCTGACCGCGAATGGTGAGATTGAGAATATAACCTTTC ATTCCCAGCGGTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAA TCGGCGTTAAACCCGCCACCAGATGGGCATTAAACGAGTATCCCGGCAG CAGGGGATCATTTTGCGCTTCAGCCATACTTTTCATACTCCCGCCATTCA GAGAAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGT CTTTTACTGGCTCTTCTCGCTAACCAAACCGGTAACCCGGGACCAAGTTT ACTCATATATACGGACAGCGGTGCGGACTGTTGT
pBAD_sfGFP150TAG	AAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTT TACTGGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTTATTAAAAGCAT TCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAG TGTCTATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTC ACACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTAC CTGACGCTTTTTATCGCAACTCTCTACTGTTTCTCCATACCCGTTTTTTGG GCTAACAGGAGGAATTAACCATGGTTAGCAAAGGTGAAGAACTGTTTAC CGGCGTTGTGCCGATTCTGGTGGAACTGGATGGTGATGTGAATGGCCAT AAATTTAGCGTTCGTGGCGAAGGCGAAGGTGATGCGACCAACGGTAAA

CTGACCCTGAAATTTATTTGCACCACCGGTAAACTGCCGGTTCCGTGGCC
GACCCTGGTGACCACCCTGACCTATGGCGTTCAGTGCTTTAGCCGCTATC
CGGATCATATGAAACGCCATGATTTCTTTAAAAGCGCGATGCCGGAAGG
CTATGTGCAGGAACGTACCATTAGCTTCAAAGATGATGGCACCTATAAA
ACCCGTGCGGAAGTTAAATTTGAAGGCGATACCCTGGTGAACCGCATTG
AACTGAAAGGTATTGATTTTAAAGAAGATGGCAACATTCTGGGTCATAA
ACTGGAATATAATTTCAACAGCCATTAGGTGTATATTACCGCCGATAAA
CAGAAAAATGGCATCAAAGCGAACTTTAAAATCCGTCACAACGTGGAA
GATGGTAGCGTGCAGCTGGCGGATCATTATCAGCAGAATACCCCGATTG
GTGATGGCCCGGTGCTGCTGCCGGATAATCATTATCTGAGCACCCAGAG
CGTTCTGAGCAAAGATCCGAATGAAAAACGTGATCATATGGTGCTGCTG
GAATTTGTTACCGCCGCGGGCATTACCCACGGTATGGATGAACTGTATA
AAGGCAGCCACCATCATCATCACCATTAAAGCTCGAGATCTGCAGCTGG
TACCATATGGGAATTCGAAGCTTGGCTGTTTTGGCGGATGAGAGAAGAT
TTTCAGCCTGATACAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAA
CAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGC
CGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCC
CCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCT
CTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGC
AACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGC
ATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCT
ACAAACTCTTTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAG
AGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGC
ATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAG
ATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCT
CAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCA
ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGT
TGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAAT
GACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCA
TGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACAC
TGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACC
GCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGA
ACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGAT
GCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTA
CTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATA
AAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATT
GCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAG
CACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC
GGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGAT
AGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCA
TATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTA
GGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGT
TTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTC
TTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAAC
CACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCT
TTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTC
CTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAC
CGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGT
GGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGG
ATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCA
GCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCT
ATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGG
GGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGAC
TTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAA
AAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTT
TTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGT
ATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCG
AGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGT

ATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACT
CTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCC
GCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACA
CCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACA
GACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACC
GTCATCACCGAAACGCGCGAGGCAGCAGATCAATTCGCGCGCGAAGGC
GAAGCGGCATGCATAATGTGCCTGTCAAATGGACGAAGCAGGGATTCTG
CAAACCCTATGCTACTCCGTCAAGCCGTCAATTGTCTGATTCGTTACCAA
TTATGACAACTTGACGGCTACATCATTCACTTTTTCTTCACAACCGGCAC
GGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAATACCCGCGAGAA
ATAGAGTTGATCGTCAAAACCAACATTGCGACCGACGGTGGCGATAGGC
ATCCGGGTGGTGCTCAAAAGCAGCTTCGCCTGGCTGATACGTTGGTCCT
CGCGCCAGCTTAAGACGCTAATCCCTAACTGCTGGCGGAAAAGATGTGA
CAGACGCGACGGCGACAAGCAAACATGCTGTGCGACGCTGGCGATATC
AAAATTGCTGTCTGCCAGGTGATCGCTGATGTACTGACAAGCCTCGCGT
ACCCGATTATCCATCGGTGGATGGAGCGACTCGTTAATCGCTTCCATGC
GCCGCAGTAACAATTGCTCAAGCAGATTTATCGCCAGCAGCTCCGAATA
GCGCCCTTCCCCTTGCCCGGCGTTAATGATTTGCCCAAACAGGTCGCTGA
AATGCGGCTGGTGCGCTTCATCCGGGCGAAAGAACCCCGTATTGGCAAA
TATTGACGGCCAGTTAAGCCATTCATGCCAGTAGGCGCGCGGACGAAAG
TAAACCCACTGGTGATACCATTCGCGAGCCTCCGGATGACGACCGTAGT
GATGAATCTCTCCTGGCGGGAACAGCAAAATATCACCCGGTCGGCAAAC
AAATTCTCGTCCCTGATTTTTCACCACCCCCTGACCGCGAATGGTGAGAT
TGAGAATATAACCTTTCATTCCCAGCGGTCGGTCGATAAAAAAATCGAG
ATAACCGTTGGCCTCAATCGGCGTTAAACCCGCCACCAGATGGGCATTA
AACGAGTATCCCGGCAGCAGGGGATCATTTTGCGCTTCAGCCATACTTT
TCATACTCCCGCCATTCAGAG

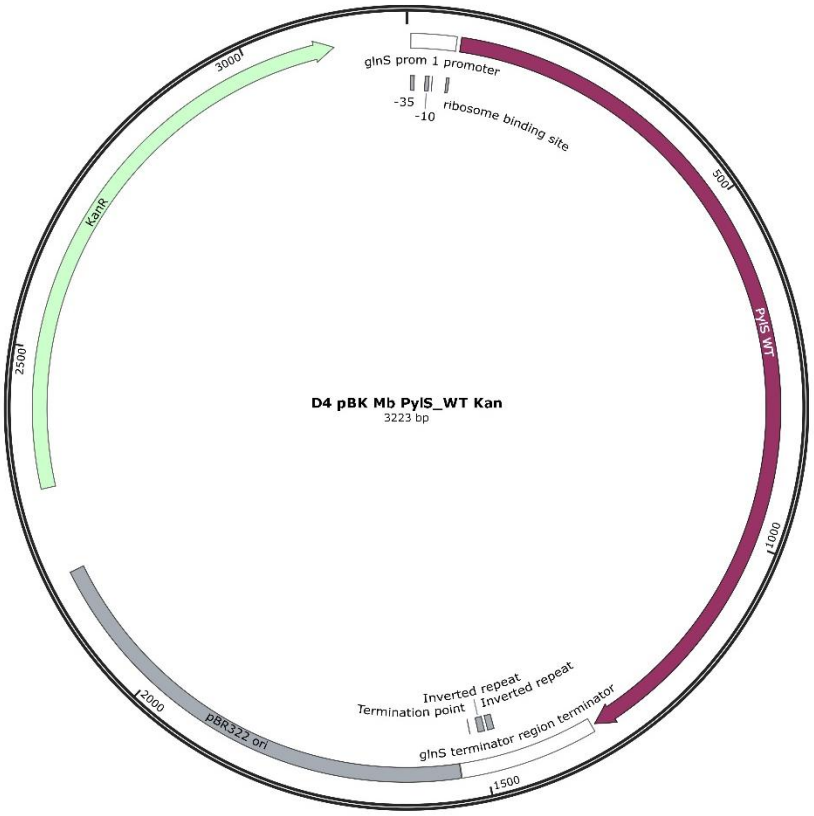


Figure II.1: Plasmid map of D4

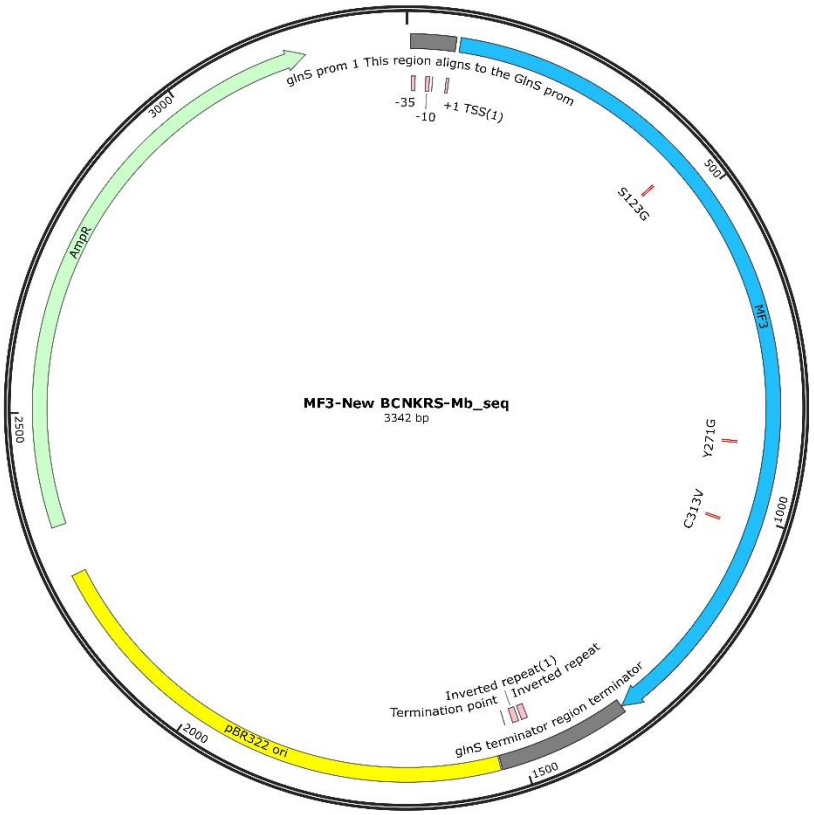


Figure II.2: Plasmid map of MF3

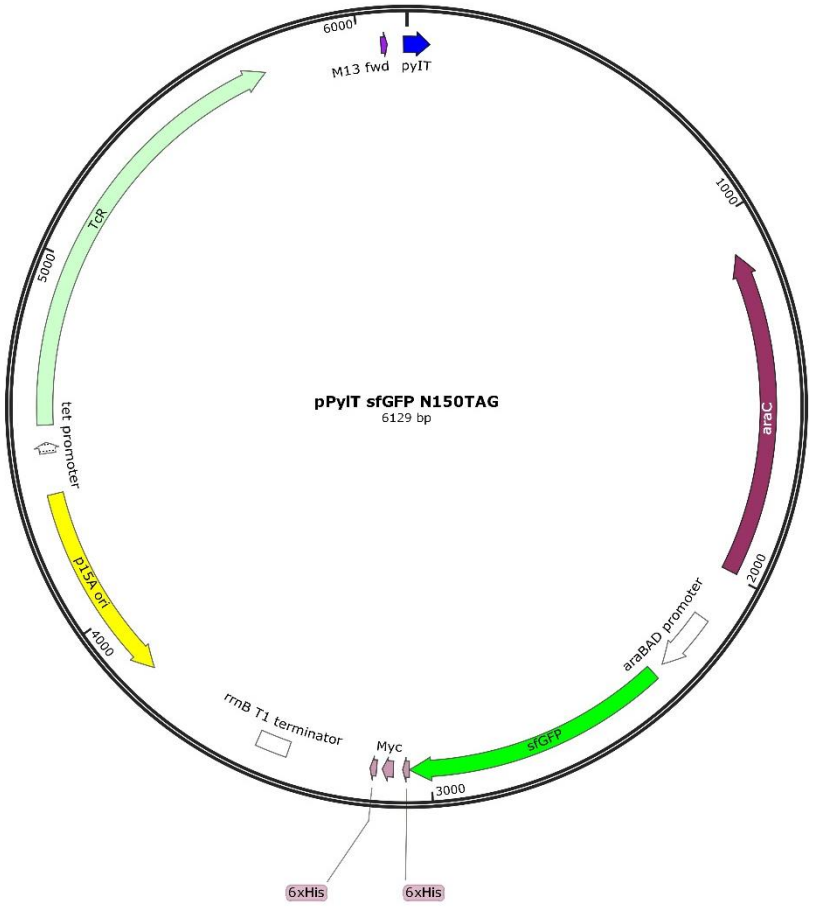


Figure II.3: Plasmid map of pPylt_sfGFPN150TAG

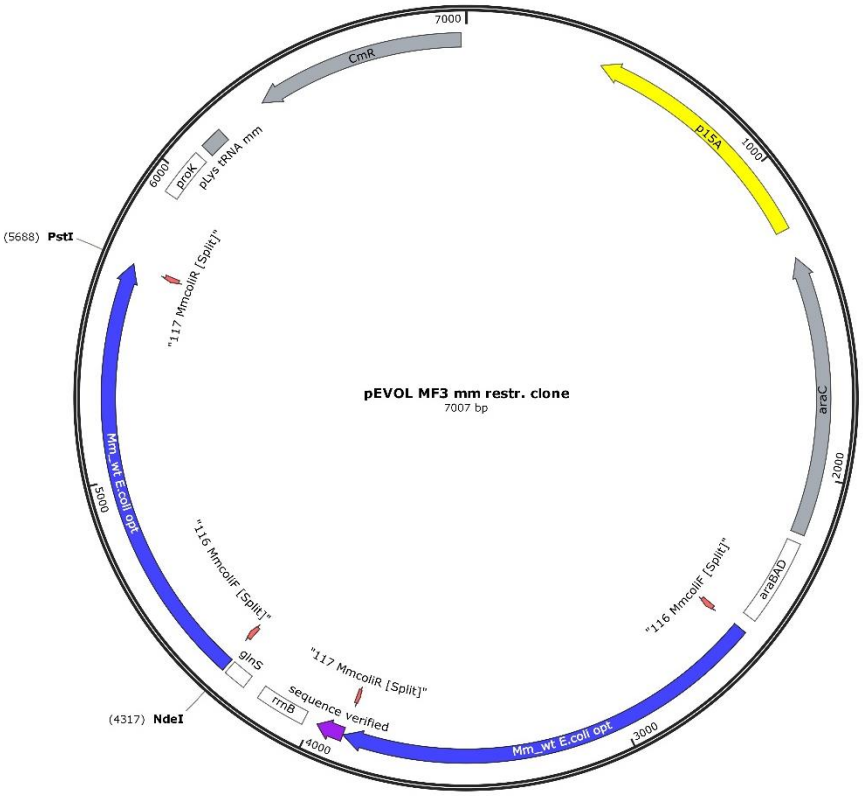


Figure II.4: Plasmid map of pEVOL_*Mm* MF3

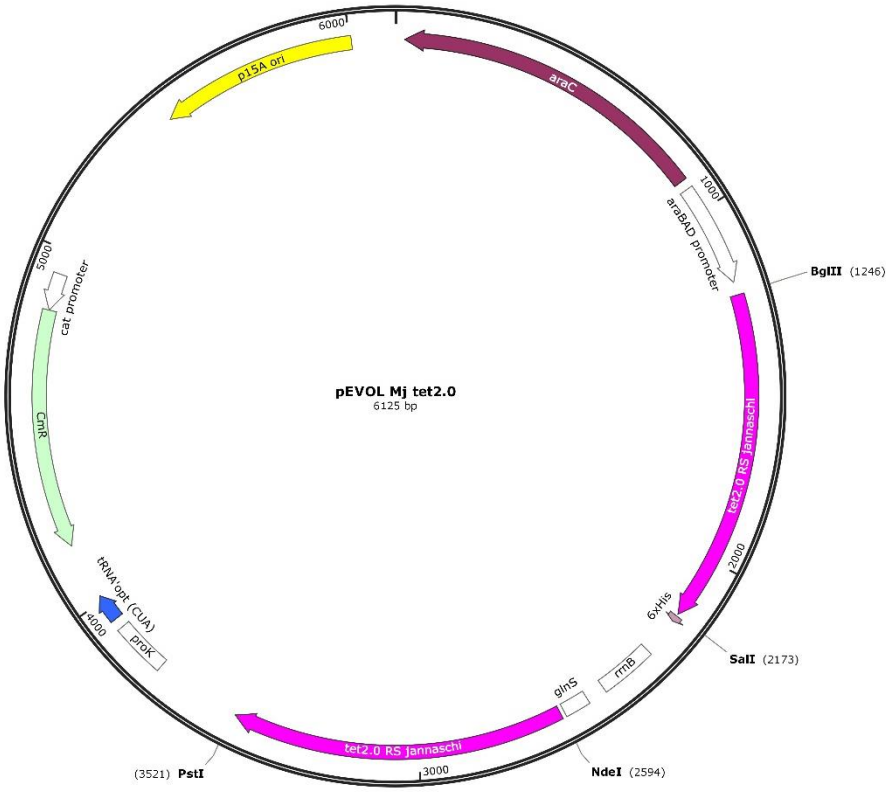


Figure II.5: Plasmid map of SM44

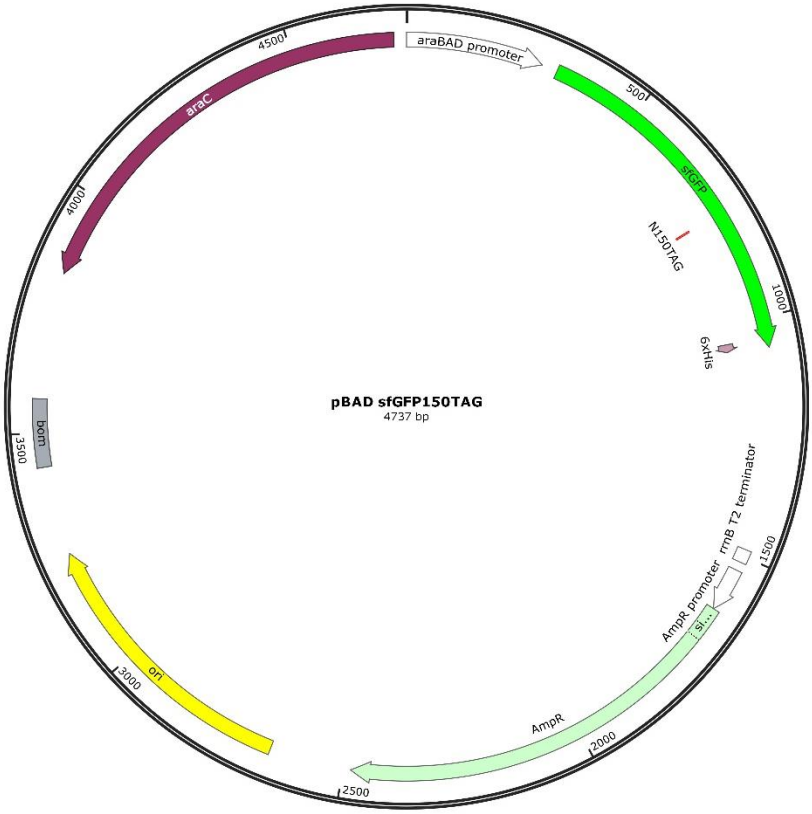


Figure II.6: Plasmid map of pBAD sfGFP150TAG

II.2.1.3 Used Primers

Quikchange primer 1-40 were designed using program Primer X hosted on Bioinformatics.org (http://www.bioinformatics.org/primerx/, last accessed 05.07.2017). Primers were designed to be 25-45 bp in length with the mutation site approx. in the middle and a T_m value ≥ 78 °C. The GC content was optimally > 40 % and C or G as a terminating base. QC primers 41-58 were designed using NEBaseChanger v 1.2.8 (NEB Inc., Ipswich, USA). Primer for restriction cloning, NEB HiFi DNA Assembly or sequencing were self-designed using SnapGene 3.1.4 (GSL Biotech LLC, Chicago, USA) in combination with NEB T_m Calculator v 1.9.13 (NEB Inc., Ipswich, USA). All primers were obtained from Sigma Aldrich (St. Louis, USA) by custom synthesis.

Table II.8: List of used QC Primer

#	Name	T_m [°C]	Sequence 5‘-3‘
QC1	Mb PylS N311A und C313A fw	55	CTGGAAGAATTCACCATGGTTGCCTTTGCCCAAATG GGCAGCGGCTGCAC
QC2	Mb PylS N311A und C313A rv	55	GTGCAGCCGCTGCCCATTTGGGCAAAGGCAACCAT GGTGAATTCTTCCAG
QC3	Mb PylS A267T fw	55	CCTGCGTCCGATGCTGACCCCGACCCTGTATAAC
QC4	Mb PylS A267T rv	55	GTTATACAGGGTCGGGGTCAGCATCGGACGCAGG
QC5	Mb PylS W382A fw	55	GGGCATTGATAAACCGGCGATTGGCGCGGGTTTTG

QC6	Mb PylS W382A rv	55	CAAAACCCGCGCCAATCGCCGGTTTATCAATGCCC
QC7	Mb PylS L274A fw	55	CCCCGACCCTGTATAACTATGCGCGTAAACTGGAT CGTATTC
QC8	Mb PylS L274A rv	55	GAATACGATCCAGTTTACGCGCATAGTTATACAGG GTCGGGG
QC9	TM3 qc L270I fw	55	GTCCGATGCTGGCCCCGACCATTTATAACTATCTGC GTAAACTG
QC10	TM3 qc L270I rv	55	CAGTTTACGCAGATAGTTATAAATGGTCGGGGCCA GCATCGGAC
QC11	TM3 qc Y271F fw	55	GCTGGCCCCGACCATTTTTAACTATCTGCGTAAAC
QC12	TM3 qc Y271F rv	55	GTTTACGCAGATAGTTAAAAATGGTCGGGGCCAGC
QC15	TM3 qc C313F fw	55	CACCATGGTTAACTTTTTTCAAATGGGCAGCGGCTG C
QC16	TM3 qc C313F rv	55	GCAGCCGCTGCCCATTTGAAAAAAGTTAACCATGG TG
QC17	Mb C313A fw	55	GAATTCACCATGGTTAACTTTGCGCAAATGGGCAG CGGCTGCAC
QC18	Mb C313A rv	55	GTGCAGCCGCTGCCCATTTGCGCAAAGTTAACCAT GGTGAATTC
QC19	Mb Y271A fw	55	GATGCTGGCCCCGACCCTGGCGAACTATCTGCGTA AACTG
QC20	Mb Y271A rv	55	CAGTTTACGCAGATAGTTCGCCAGGGTCGGGGCCA GCATC
QC21	Mb L274A fw	55	CCCGACCCTGTATAACTATGCGCGTAAACTGGATC GTATTC
QC22	Mb L274A rv	55	GAATACGATCCAGTTTACGCGCATAGTTATACAGG GTCGGG
QC23	SM24 C313V fw	55	GAATTCACCATGGTTAACTTTGTGCAAATGGGCAG CGGCTGCAC
QC24	SM24 C313V rv	55	GTGCAGCCGCTGCCCATTTGCACAAAGTTAACCAT GGTGAATTC
QC25	SM24 C313S fw	55	CACCATGGTTAACTTTTCGCAAATGGGCAGCGGCT G
QC26	SM24 C313S rv	55	CAGCCGCTGCCCATTTGCGAAAAGTTAACCATGGT G
QC27	SM23 I378L fw	55	CTGGATCGTGAATGGGGCCTGGATAAACCGTGGAT TGG
QC28	SM23 I378L rv	55	CCAATCCACGGTTTATCCAGGCCCCATTCACGATCC AG
QC29	D4 N311A C313S fw	55	CACCTGGAAGAATTCACCATGGTTGCGTTTTCGCAA ATGGGCAGCGGCTGCACCCGTG
QC30	D4 N311A C313S rv	55	CACGGGTGCAGCCGCTGCCCATTTGCGAAAACGCA ACCATGGTGAATTCTTCCAGGTG
QC31	D171 QC Y349W fw	55	GGGCGATAGCTGCATGGTGTGGGGCGATACCCTGG

			ATATTATG
QC32	D171 QC Y349W rv	55	CATAATATCCAGGGTATCGCCCCACACCATGCAGC TATCGCCC
QC35	TM3 qc L274G new fw	55	CCCCGACCCTGTATAACTATGGCCGTAAACTGGAT CGTATTC
QC36	TM3 qc L274G new fw	55	GAATACGATCCAGTTTACGGCCATAGTTATACAGG GTCGGGG
QC39	D4 QC Y349F fw	70.5	GCTGCATGGTGTTTGGCGATACCCTG
QC40	D4 QC Y349F rv	70.5	CAGGGTATCGCCAAACACCATGCAGC
QC41	Primer QC_F Kit S109F	58	TGGAAGTGAATTTCAGCTTGATAAG
QC42	Primer QC_R Kit S109F	58	TAAACATATTTTGCCTTTAACC
QC43	MF3 QC L274A_F	67	GGGGAACTATgccCGTAAACTGGATCGTATTCTG
QC44	MF3 QC L274A_R	67	AGGGTCGGAGCCAGCATC
QC45	MF3 QC V313F_F	59	GGTTAACTTTtttCAAATGGGCAGC
QC46	MF3 QC V313F_R	59	ATGGTGAATTCTTCCAGG
QC47	OmpC QC Y232TAG fw	66	CACCGCTGCTTAGATCGGTAACG
QC48	OmpC QC Y232TAG rv	66	TTCTGAGCATCAGTACGTTTGG
QC49	mmPylRS_Y306G fw	68	ACCGAATCTGGGCAATTATCTGCGCAAACTGGATC
QC50	mmPylRS_Y306G rv	68	GCCAGCATCGGACGCAGA
QC51	mmPylRS_C348V fw	59	GCTGAATTTTGTGCAGATGGGTAGCGG
QC52	mmPylRS_C348V rv	59	ATGGTAAATTCTTCCAGATG
QC53	QC TAG150N fw	56	CAACAGCCATAATGTGTATATTACC
QC54	QC TAG150N rv	56	AAATTATATTCCAGTTTATGACC
QC55	QC TAGN40TAG fw	67	TGATGCGACCTAGGGTAAACTGACCC
QC56	QC TAGN40TAG rv	67	CCTTCGCCTTCGCCACGA
QC57	QC TAGS3TAG fw	59	AACCATGGTTTAGAAAGGTGAAGAACTGTTTACC
QC58	QC TAGS3TAG rv	59	AATTCCTCCTGTTAGCCC

Table II.9: List of used primers and genestrings for restriction cloning and NEB HiFi DNA Assembly

#	Name	T_m [°C]	Sequence 5‘-3‘
1	R_backbone pDule	64	GCAAGGGAGCTTTTACCAAAA
2	F_backbone pDule	64	CGTCCATGGGGGATTCCTC
3	F_string M.jannschii	64	GAGGAATCCCCCATGGACG
4	R_string M.jannschii	64	TTTTGGTAAAAGCTCCCTTGC
GS5	Genestring Mj AzoF	-	GAGGAATCCCCCATGGACGAATTTGAAATGATAAAGAG AAACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAG AGGTTTTAAAAAAAGATGAAAAATCTGCTGGGATAGGT

			TTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTC CAAATAAAAAAGATGATTGATTTACAAAATGCTGGATT TGATATAATTATAGAATTGGCTGATTTACACGCCTATTT AAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATA GGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTT AAAGGCAAAATATGTTTATGGAAGTGAAGCGGAACTTG ATAAGGATTATACACTGAATGTCTATAGATTGGCTTTAA AAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTT ATAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGT TATCTATCCAATAATGCAGGTTAATGGCATTCATTATCA TGGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAA AAATACACATGTTAGCAAGGGAGCTTTTACCAAAA
GS6	Genestring Mj TetPyrA_1	-	GAGGAATCCCCCATGGACGAATTTGAAATGATAAAGAG AAACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAG AGGTTTTAAAAAAAGATGAAAAATCTGCTGGGATAGGT TTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTC CAAATAAAAAAGATGATTGATTTACAAAATGCTGGATT TGATATAATTATACAGTTGGCTGATTTAGCGGCCTATTT AAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATA GGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTT AAAGGCAAAATATGTTTATGGAAGTGAATCTCAGCTTG ATAAGGATTATACACTGAATGTCTATAGATTGGCTTTAA AAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTT ATAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGT TATCTATCCAATAATGGAAGTTAATTCTATTCATTATAA TGGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAA AAATACACATGTTAGCAAGGGAGCTTTTACCAAAA
GS7	Genestring Mj TetPyrA_1 ohne	-	GAGGAATCCCCCATGGACGAATTTGAAATGATAAAGAG AAACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAG AGGTTTTAAAAAAAGATGAAAAATCTGCTGGGATAGGT TTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTC CAAATAAAAAAGATGATTGATTTACAAAATGCTGGATT TGATATAATTATACAGTTGGCTGATTTAGCGGCCTATTT AAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATA GGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTT AAAGGCAAAATATGTTTATGGAAGTGAATCTCAGCTTG ATAAGGATTATACACTGAATGTCTATAGATTGGCTTTAA AAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTT ATAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGT TATCTATCCAATAATGCAGGTTAATTCTATTCATTATAA TGGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAA AAATACACATGTTAGCAAGGGAGCTTTTACCAAAA
GS8	Genestring Mj TetPyrA_2	-	GAGGAATCCCCCATGGACGAATTTGAAATGATAAAGAG AAACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAG AGGTTTTAAAAAAAGATGAAAAATCTGCTGGGATAGGT TTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTC CAAATAAAAAAGATGATTGATTTACAAAATGCTGGATT TGATATAATTATACAGTTGGCTGATTTAGCGGCCTATTT AAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATA GGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTT AAAGGCAAAATATGTTTATGGAAGTGAATCTCAGCTTG ATAAGGATTATACACTGAATGTCTATAGATTGGCTTTAA AAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTT ATAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGT TATCTATCCAATAATGGAAGTTAATTCTATTCATTATAA TGGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAA AAATACACATGTTAGCAAGGGAGCTTTTACCAAAA
GS9	Genestring Mj	-	GAGGAATCCCCCATGGACGAATTTGAAATGATAAAGAG AAACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAG

	TetPyrA_2 ohne		AGGTTTTAAAAAAAGATGAAAAATCTGCTGGGATAGGT TTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTC CAAATAAAAAAAGATGATTGATTTACAAAATGCTGGATT TGATATAATTATACAGTTGGCTGATTTAGCGGCCTATTT AAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATA GGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTT AAAGGCAAAATATGTTTATGGAAGTGAATTTCAGCTTG ATAAGGATTATACACTGAATGTCTATAGATTGGCTTTAA AAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTT ATAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGT TATCTATCCAATAATGCAGGTTAATTCTATTCATTATAA TGGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAA AAATACACATGTTAGCAAGGGAGCTTTTACCAAAA
GS10	Genestring Mj BiPyA	-	GAGGAATCCCCCATGGACGAATTTGAAATGATAAAGAG AAACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAG AGGTTTTAAAAAAAGATGAAAAATCTGCTGGCATAGGT TTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTC CAAATAAAAAAAGATGATTGATTTACAAAATGCTGGATT TGATATAATTATATATTTGGCTGATTTAGCGGCCTATTT AAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATA GGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTT AAAGGCAAAATATGTTTATGGAAGTGAATTTCAGCTTG ATAAGGATTATACACTGAATGTCTATAGATTGGCTTTAA AAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTT ATAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGT TATCTATCCAATAATGGAAGTTAATGGCTGGCATTATAG CGGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAA AAATACACATGTTAGCAAGGGAGCTTTTACCAAAA
GS11	Genestring Mj BPA	-	GAGGAATCCCCCATGGACGAATTTGAAATGATAAAGAG AAACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAG AGGTTTTAAAAAAAGATGAAAAATCTGCTGGCATAGGT TTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTC CAAATAAAAAAAGATGATTGATTTACAAAATGCTGGATT TGATATAATTATACTGTTGGCTGATTTACACGCCTATTT AAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATA GGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTT AAAGGCAAAATATGTTTATGGAAGTCCGTTTCAGCTTGA TAAGGATTATACACTGAATGTCTATAGATTGGCTTTAAA AACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTTA TAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGTT ATCTATCCAATAATGCAGGTTAATACCAGCCATTATGCG GGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAAA AATACACATGTTAGCAAGGGAGCTTTTACCAAAA
GS12	Genestring Mj oBNY	-	GAGGAATCCCCCATGGACGAATTTGAAATGATAAAGAG AAACACATCTGAAATTATCAGCGAGGAAGAGTTAAGAG AGGTTTTAAAAAAAGATGAAAAATCTGCTGGCATAGGT TTTGAACCAAGTGGTAAAATACATTTAGGGCATTATCTC CAAATAAAAAAAGATGATTGATTTACAAAATGCTGGATT TGATATAATTATAGGCTTGGCTGATTTAATGGCCTATTT AAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATA GGAGATTATAACAAAAAAGTTTTTGAAGCAATGGGGTT AAAGGCAAAATATGTTTATGGAAGTGAAGGCCAGCTTG ATAAGGATTATACACTGAATGTCTATAGATTGGCTTTAA AAACTACCTTAAAAAGAGCAAGAAGGAGTATGGAACTT ATAGCAAGAGAGGATGAAAATCCAAAGGTTGCTGAAGT TATCTATCCAATAATGCAGGTTAATAGCATGCATTATGA AGGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAA AAATACACATGTTAGCAAGGGAGCTTTTACCAAAA

13	D4 bb fw	67	CTGCTGAAAGTGATGCATGGCTTC
14	D4 bb rv	67	CCGGAATCAGAATCGGGC
15	D4 String fw	67	GCCCGATTCTGATTCCGG
16	D4 String rv	67	GAAGCCATGCATCACTTTCAGCAG
GS17	Genestring PylS HA1	-	TAAAAGCCCGATTCTGATTCCGGCGGAATATGTGGAAC GTGCGGGCATTAACAACGACACCGAACTGAGCAAACAA ATTTTCCGCGTGGATAAAAACCTGTGCCTGCGTCCGATG CTGGCCCCGACCCTGGCGAACTATGTGCGTAAACTGGA TCGTATTCTGCCGGGTCCGATCAAAATTTTTGAAGTGGG CCCGTGCTATCGCAAAGAAAGCGATGGCAAAGAACACC TGGAAGAATTCACCATGGTTAACTTTGTGCAATATGGCA GCGGCTGCACCCGTGAAAACCTGGAAGCGCTGATCAAA GAATTCCTGGATTATCTGGAAATCGACTTCGAAATTGTG GGCGATAGCTGCATGGTGTATGGCGATACCCTGGATATT ATGCATGGCGATCTGGAACTGAGCAGCGCGgtgGTGGGT CCGCGTAGCCTGGATCGTGAATGGGGCATTGATAAACC GTGGATTGGCGCGGGTTTTGGCCTGGAACGTCTGCTGAA AGTGATGCATGGCTTCAAAAA
GS18	Genestring PylS HA2	-	TAAAAGCCCGATTCTGATTCCGGCGGAATATGTGGAAC GTATGGGCATTAACAACGACACCGAACTGAGCAAACAA ATTTTCCGCGTGGATAAAAACCTGTGCCTGCGTCCGATG CTGGCCCCGACCCTGGCGAACTATGTGCGTAAACTGGA TCGTATTCTGCCGGGTCCGATCAAAATTTTTGAAGTGGG CCCGTGCTATCGCAAAGAAAGCGATGGCAAAGAACACC TGGAAGAATTCACCATGGTTAACTTTGTGCAATATGGCA GCGGCTGCACCCGTGAAAACCTGGAAGCGCTGATCAAA GAATTCCTGGATTATCTGGAAATCGACTTCGAAATTGTG GGCGATAGCTGCATGGTGTATGGCGATACCCTGGATATT ATGCATGGCGATCTGGAACTGAGCAGCGCGgtgGTGGGT CCGCGTAGCCTGGATCGTGAATGGGGCATTGATAAACC GTGGATTGGCGCGGGTTTTGGCCTGGAACGTCTGCTGAA AGTGATGCATGGCTTCAAAAA
GS19	Genestring PylS HA3	-	TAAAAGCCCGATTCTGATTCCGGCGGAATATGTGGAAC GTATGGGCATTAACAACGACACCGAACTGAGCAAACAA ATTTTCCGCGTGGATAAAAACCTGTGCCTGCGTCCGATG CTGGCCCCGACCCTGGCGAACTATGTGCGTAAACTGGA TCGTATTCTGCCGGGTCCGATCAAAATTTTTGAAGTGGG CCCGTGCTATCGCAAAGAAAGCGATGGCAAAGAACACC TGGAAGAATTCACCATGGTTAACTTTGTGCAATATGGCA GCGGCTGCACCCGTGAAAACCTGGAAGCGCTGATCAAA GAATTCCTGGATTATCTGGAAATCGACTTCGAAATTGTG GGCGATAGCTGCATGGTGTTTGGCGATACCCTGGATATT ATGCATGGCGATCTGGAACTGAGCAGCGCGgtgGTGGGT CCGCGTAGCCTGGATCGTGAATGGGGCATTGATAAACC GTGGATTGGCGCGGGTTTTGGCCTGGAACGTCTGCTGAA AGTGATGCATGGCTTCAAAAA
38	F_pET_Mj TyrS	59	ATCTCGACGCTCTCCCTTATGAGGAATCCCCCATGGACG
39	R_pET_Mj TyrS	59	AGTGCCACCTGACGTCTAATGCAGTTATAATCTCTTTCT AATTG
40	Primer Mj BglII	48	GGAATTAGATCTATGGACGAATTTGAAATG
41	Primer Mj SalI	48	GATGATGGTCGACTTATAATCTCTTTCTAATTG
42	Primer Mj NdeI	55	GGAATCCCATATGGACGAATTTGAAATG

43	Primer Mj PstI	55	TTGAAACTGCAGTTATAATCTCTTTCTAATTG
44	F_pET301_Backbone	66	TTAGACGTCAGGTGGCACT
45	R_pET301_Backbone	66	TAAGGGAGAGCGTCGAGAT
46	OmpC fw	66	ATGAAAGTTAAAGTACTGTCCCTCCTGG
47	OmpC rv	66	GAACTGGTAAACCAGACCCAG
48	pPylt bb_ompc fw	66	CCAGGAGGGACAGTACTTTAACTTTCATGGTTAATTCCT CCTGTTAGCC
49	pPylt bb_ompc rv	66	CTGGGTCTGGTTTACCAGTTCCATCATCATCATCATCAT TGAGTTTAAACG
50	Primer insert MF3 NdeI	56	AATCCCATATGGATAAAAAAC
51	Primer insert MF3 PstI	56	GAAACTGCAGTTACAGATTG
52	MW17 FP_pEVOL_backbone	62	GTCGACCATCATCATCATCA
53	MW37 R_BackboneI_pEVOL	62	AGATCTAATTCCTCCTGTTAGC
54	MW35 F_GA_PylS_overhang I	60	GCTAACAGGAGGAATTAGATCTATGGACAAAAAACCGC TGAA
55	Primer insert MF3 mm rv	60	TGATGATGATGATGGTCGACTTACAGATTGGTGCTAATA C

Table II.10: Used sequencing primer

#	Name	Sequence 5‘-3‘
S1	pD2 TyrRS fw	ATACTTGTAACGCTGAG
S2	pD2 TyrRS rv	CCGCCAAAACAGCCAAGCTGG
S3	pD2 bb Myc fw	AAACTCATCTCAGAAGAGG
S4	pD2 bb SmR fw	TGATCTGGCTATCTTGCTG
S5	pD2 tRNA fw	GTAAAACGACGGCCAGTGCC
S6	pD2 bb araC fw	CGAGGCAGCAGATCAATTC
S7	pD2 bb rv TyrRS rv	GATTCCTCAAAGCGTAAAC
S8	pD2 bb SmR rv	TTCAGGAACCGGATCAAAG
S9	pD2 bb tRNA rv	CATGGCGCTGGTTCAAATC
S10	pD2 bbfw2	CGACCGGGTCGAATTTGC
S11	pBK PylS fwd	AGTTGAAGGATCCTCGG
S12	pBK PylS rv	CCTACAAAAGCACGCAAAc
S13	pGFP pylt seq fw	GCGTCACACTTTGCTATG
S14	pYOBB2 seq rv (CAM)	CTTCCATGTCGGCAGAATGC

S15	Seq_PEVOL_ rrnB fw	TGTTTTGGCGGATGAGAG
S16	Seq_pEVOL_Mj#2 rv	TCATGTAGGCCTGATAAGCG
S17	Seq_pEVOL_CAM fw	ATCAACGGTGGTATATCCAG
S18	Seq_pEVOL_p15A ori fw	AAAGATACCAGGCGTTTCCC
S19	Seq_pEVOL_araBAD rv	TCCCGCTTTGTTACAGAATG
S20	Seq_pET301 fw	ATTCACCACCCTGAATTGAC
S21	Seq_pET301 rv	TGAAGCATTTATCAGGGTTA
S22	Seq_pEVOL_fw	CACTGCGTCTTTTACTGGC
S23	Seq_pEVOL_PylS_rv	ATCAGACCGCTTCTGCG
S27	Seq_pBK Pyls D1 fw	CGTCAGAGGATCCTCGG
S28	Seq_OmpCfw	CCATCTGGTGAAGGCTTTAC

II.2.1.4 Enzymes, proteins, buffers and other additives for PCR

Table II.11: List of used enzymes and proteins

Enzyme/Protein	Buffer	Supplier
AflIII restriction endonuclease	NEB 3.1 buffer	NEB Inc. (Ipswich, USA)
Antarctic phosphatase	10x Antarctic phosphatase reaction buffer	NEB Inc. (Ipswich, USA)
Anti-His_6-Peroxidase, mouse monoclonal antibody (50 U/mL)	-	Roche Applied Science (Penzberg, Germany)
BglII restriction endonuclease	10x NEB 3.1 buffer	NEB Inc. (Ipswich, USA)
BSA, protein standard (1 mg/mL)	0.15 M NaCl	Sigma Aldrich (St. Louis, USA)
DNAseI powder from bovine pancreas	-	Sigma Aldrich (St. Louis, USA)
Dpn1 restriction endonuclease	10x Cutsmart	Itzen lab, TUM (Munich)
HRP-conjugated Streptavidin	-	Thermo Fisher Scientific (Waltham, USA)
NEBuilder HiFi DNA Assembly Master Mix	-	NEB Inc. (Ipswich, USA)
NcoI restriction endonuclease	10x NEB 3.1 buffer	NEB Inc. (Ipswich, USA)
NcoI HF restriction endonuclease	10x Cutsmart	NEB Inc. (Ipswich, USA)
NdeI restriction endonuclease	10x Cutsmart	NEB Inc. (Ipswich, USA)
Penicillin Amidase (0.752 U/µL)	-	Sigma Aldrich (St. Louis, USA)
PstI HF restriction endonuclease	10x Cutsmart	NEB Inc. (Ipswich, USA)
PfuTurbo DNA polymerase	10x Pfu reaction buffer	Agilent Technologies (Santa Clara, USA)
PfuUltra II Fusion HS DNA polymerase	10x PfuUltra II reaction buffer	Agilent Technologies (Santa Clara, USA)

Q5 High Fidelity DNA polymerase	5x Q5 reaction buffer	NEB Inc. (Ipswich, USA)
SalI HF restriction endonuclease	10x NEB 3.1 buffer	NEB Inc. (Ipswich, USA)
Streptavidin, tetramethylrhodamine conjugate (2 mg/mL, 15 U/mg)	-	Thermo Fisher Scientific (Waltham, USA)
T4 DNA ligase	10x T4 DNA ligase buffer	NEB Inc. (Ipswich, USA)
T4 Polynucleotide Kinase	10x T4 DNA ligase buffer	NEB Inc. (Ipswich, USA)
XbaI restriction endonuclease	10x Cutsmart	NEB Inc. (Ipswich, USA)

Table II.12: List of used buffers and additives for PCR

Buffers / addivites for PCR	Supplier
10x Antarctic phosphatase reaction buffer	NEB Inc. (Ipswich, USA)
10x Cutsmart	NEB Inc. (Ipswich, USA)
DMSO	NEB Inc. (Ipswich, USA)
dNTP (10 mM each)	NEB Inc. (Ipswich, USA)
10x NEB 3.1 buffer	NEB Inc. (Ipswich, USA)
10x Pfu reaction buffer	Agilent Technologies (Santa Clara, USA)
10x PfuUltra II reaction buffer	Agilent Technologies (Santa Clara, USA)
5x Q5 enhancer	NEB Inc. (Ipswich, USA)
5x Q5 reaction buffer	NEB Inc. (Ipswich, USA)
10x T4 DNA ligase buffer (+ 10 mM ATP)	NEB Inc. (Ipswich, USA)

II.2.1.5 Marker and Standards

Table II.13: List of used marker and standards

Marker and Standard	Supplier
BluRay Prestained Protein Marker (10-180 kDa)	Jena Bioscience (Jena, Germany)
Color Prestained Protein Standard, Broad Range (11–245 kDa)	NEB Inc. (Ipswich, USA)
Protein Marker III (6.5-200 kDa)	Applichem (Darmstadt, Germany)
QuickLoad purple 100 bp DNA Ladder	NEB Inc. (Ipswich, USA)
500 bp DNA Ladder (0.5 – 5 kb)	Jena Bioscience (Jena, Germany)
High Range DNA Ladder (0.5-10 kb)	Jena Bioscience (Jena, Germany)

II.2.1.6 Kits

Table II.14: List of used kits

Kit	Supplier
HiSpeed Plasmid Midi Kit	Qiagen (Hilden, Germany)
HiSpeed Plasmid Maxi Kit	Qiagen (Hilden, Germany)
GenElute™ Bacterial Genomic DNA Kits	Sigma Aldrich (St. Louis, USA)
Monarch PCR DNA Cleanup Kit	NEB Inc. (Ipswich, USA)
Monarch DNA Gel Extraction Kit	NEB Inc. (Ipswich, USA)
peqGOLD Plasmid Miniprep Kit (Classic Line)	VWR Life Science Competence Center (Erlangen, Germany)

II.2.1.7 Antibiotics

Antibiotics were dissolved, sterile filtered with a 0.2 µm membrane filter, aliquoted and stored at -20 °C. To avoid degradation of the antibiotics addition to the respective solutions (eq. media, agar, etc.) took place after cooling to approx. 50 °C.

Table II.15: List of used antibiotics

Antibiotic	Stock solution [mg/mL]	Working concentration [µg/mL]	Solvent
Ampicillin (Amp)	100	100	MQ-H_2O
Chloramphenicol (Cam)	50	50	EtOH
Kanamycin (Kan)	50	50	MQ-H_2O
Spectinomycin (Sm)	50	50	MQ-H_2O
Tetracycline (Tcn)	17	17	70 % EtOH/MQ-H_2O

II.2.1.8 Buffers, stock solutions, liquid and solid media compositions

Solutions, buffers and media were prepared with demineralized water if not indicated otherwise and media was autoclaved at 121 °C for 20 min. Thermally sensitive compounds were dissolved in MQ water and sterile filtered with a 0.2 µm membrane filter. The pH was adjusted using HCl and NaOH and refers to rt, if not mentioned otherwise.

Table II.16: List of used stock solutions

Solution	Amount/Concentration	Storage	Comment
5 % Agarose	5 % (w/v)	rt	autoclaved
L-Arabinose	20 % (w/v)	4 °C	sterile filtered
5 % Aspartate, pH 7.5	5 % (w/v)	rt	sterile filtered
APS	10 % (w/v)	-20 °C	sterile filtered

Biotin	1 mg/mL	4 °C	sterile filtered
$CaCl_2$ (for trace metal mix)	2 M	rt	sterile filtered
$CaCl_2$	1 M	4 °C	autoclaved
$CaCl_2$ (for competent cells)	100 mM	4 °C	autoclaved
Citric acid	100 mM	Rt	sterile filtered
$CoCl_2$	200 mM	4 °C	sterile filtered
$CuCl_2$	200 mM	4 °C	sterile filtered
$FeCl_3$	100 mM	4 °C	sterile filtered
Glucose	40 % (w/v)	rt	sterile filtered
10 % Glycerol	10 % (v/v)	4 °C	autoclaved
H_3BO_3	200 mM	4 °C	sterile filtered
Imidazole, pH 8	5 M	4 °C	-
India Ink	-	rt	autoclaved
IPTG	1 M	-20 °C	sterile filtered
Leucine, pH 7.5	4 mg/mL	rt	autoclaved
β-Mercaptoethanol	1 M		sterile filtered
$MgSO_4$	1 M	rt	autoclaved
$MnCl_2$	1 M	4 °C	sterile filtered
NaCl	5 M	rt	autoclaved
NaOH	1 M	rt	sterile filtered
NaOH	100 mM	rt	sterile filtered
Na_2HPO_4	200 mM	rt	autoclaved
Na_2MoO_4	200 mM	4 °C	sterile filtered
Na_2SeO_3	200 mM	4 °C	sterile filtered
$NiCl_2$	200 mM	4 °C	sterile filtered
Nicotinamide	1 M	rt	sterile filtered
PMSF in DMSO	200 mM	4 °C	sterile filtered
Thiamine	1 mg/mL	rt	sterile filtered
TFA	1 M	rt	sterile filtered
$ZnSO_4$	1 M	4 °C	sterile filtered

Table II.17: List of unnatural amino acid stock solutions

Unnatural amino acid (uAA)	Stock solution [mM]	Working concentration [mM]	Solvent
BCNK	50	2	100 mM NaOH/DMSO (1:1)
BocK	100	2	100 mM NaOH
CpK	100	2	100 mM NaOH
TetA	100	2	200 mM TFA (H_2O)
TetF	100	0.5-4	200 mM TFA (H_2O)
IpTetF	100	0.5-2	200 mM TFA (H_2O)
TetPyrA	100	2	200 mM TFA (H_2O)
TetK	100	0.5-2	200 mM TFA (H_2O)
mTetK	100	0.5-2	200 mM TFA (DMSO)
K-mTetK	100	2	H_2O
pTetK	100	2	200 mM TFA (DMSO)
K-pTetK	100	2	H_2O
mTet2K	100	2	200 mM TFA (DMSO)
HmTetK	100	0.5-2	200 mM TFA (DMSO)
PhTetK	100	2	200 mM TFA (DMSO)
PacK	100	2	300 mM TFA (H_2O)
AOPacK	100	2	200 mM TFA (H_2O)

Table II.18: List of compound stock solutions

Compound	Stock solution [mM]	Working concentration [mM]	Solvent
Azidoacetic acid (AzAcOH)	100	5-15	ACN
BCN-TAMRA	2	0.002	DMSO
BCN-methanol	100		MeOH
CpdSL	100	0.1-1	DMSO
Cp-PROXYL	100	0.2	DMSO
Cp-TEMPO	100	0.2	DMSO
DimethylTet	100		ACN
DiphenylTet	100		DMSO
DipyridylTet	50		DMSO
PheMeTet	100	0.5	ACN
photo-DMBO amine	50-100		DMSO
photo-DMBO-BCN	2		DMSO

TCO-TAMRA	2	0.002	DMSO
mTet-TAMRA	2	0.002	DMSO
Tet-TEMPO	100	0.2	DMSO
Tet-PROXYL	100	0.2	DMSO
TM94	10	0.1-5	MeOH
$TM94_{dec}$	20		ACN
TM127	10		ACN
$TM127_{dec}$	10		ACN
TM168 (TM94-TAMRA)	2		DMSO
TM176 (TM94-PEG-TAMRA)	2		DMSO
TM176dec	2		DMSO
TM179 (TM94-PEG-Biotin)	2		DMSO
TM275 (TM94-PEG-sulfoCy5)	2		DMSO
TM275dec	2		DMSO

Table II.19: List of used buffers, media and agar compositions

Buffer	Amount / Concentration	Components	Storage	Comment
25x 18 Amino acid mix	5 g	Glutamic acid sodium salt	4 °C	sterile filtered
	5 g	Aspartic acid		
	5 g	Lysine-HCl		
	5 g	Arginine-HCl		
	5 g	Histidine-HCl-H2O		
	5 g	Alanine		
	5 g	Proline		
	5 g	Glycine		
	5 g	Threonine		
	5 g	Serine		
	5 g	Glutamine		
	5 g	Asparagine-H2O		
	5 g	Valine		
	5 g	Leucine		
	5 g	Isoleucine		
	5 g	Phenylalanine		
	5 g	Tryptophan		
	5 g	Methionine		
	fill to 1 L	MQ H_2O		
Citrate-phosphate buffer, pH 6	36.85 mL	0.1 M citric acid	rt	-
	63.15 mL	0.2 M Na_2HPO_4		
Citrate-phosphate buffer,	48.5 mL	0.1 M citric acid	rt	-

pH 5	51.5 mL	0.2 M Na_2HPO_4		
Citrate-phosphate buffer, pH 4	61.45 mL 38.55 mL	0.1 M citric acid 0.2 M Na_2HPO_4	rt	-
Citrate-phosphate buffer, pH 3	79.45 mL 20.55 mL	0.1 M citric acid 0.2 M Na_2HPO_4	rt	-
Citrate-phosphate buffer, pH 2	95 mL 5 mL	0.1 M citric acid 0.2 M Na_2HPO_4	rt	-
Citrate-phosphate buffer, pH 1	100 mL	0.1 M citric acid	rt	-
10x DNA loading buffer	40 % (w/v) 0.15 % (w/v) 0.05 % (w/v) 0.05 % (w/v)	Sucrose Orange G Xylene cyanol Bromophenol blue	-20 °C	-
Elution buffer	20 mM 300 mM 300 mM	Tris HCl pH 8 Imidazole, pH 8 NaCl	4 °C	Always prepare fresh
LB medium, pH 7	10 g 5 g 5 g fill to 1 L	Tryptone Yeast extract NaCl H_2O	rt	autoclaved
LB agar	2.5 g 1.25 g 1.25 g 3.75 g fill to 250 mL	Tryptone Yeast extract NaCl Agar agar H_2O	rt	autoclaved
Lysis buffer, pH 8	20 mM 30 mM 300 mM 1 mM 0.1 mg/mL	Tris HCl pH 8 Imidazole, pH 8 NaCl PMSF Dnase I	-	Always prepare fresh
50x M	1.25 M 1.25 M 2.5 M 0.25 M	Na_2HPO_4 KH_2PO_4 NH_4Cl Na_2SO_4	rt	autoclaved
5x M9 salts	42.5 g 15 g 0.4 g 5 g fill to 1 L	$Na_2HPO_4 \cdot 2H_2O$ KH_2PO_4 NaCl NH_4Cl H_2O	rt	autoclaved
10x PBS, pH 7.4	100 mM 18 mM 0.4 M	$Na_2HPO_4 \cdot 2H_2O$ KH_2PO_4 NaCl	4 °C	-

	27 mM	KCl		
1x PBS, pH 7.4	10 mM	$Na_2HPO_4 \cdot 2H_2O$	4 °C	-
	0.4 mM	KH_2PO_4		
	137 mM	NaCl		
	2.7 mM	KCl		
4x Resolving buffer, pH 8.8	1.5 M	Tris HCl pH 8.8	rt	-
	0.4 % (w/v)	SDS		
20x SDS-MES Running buffer, pH 7.3	1 M	MES	rt	Light sensitive. Store dark
	1 M	Tris base		
	2 % (w/v)	SDS		
	20 mM	EDTA		
4x SDS-Loading buffer	40 % (v/v)	Glycerol	-20 °C	-
	240 mM	Tris/HCl pH 6.8		
	8 % (w/v)	SDS		
	0.04 % (w/v)	Bromophenol blue		
	5 % (v/v)	β-Mercaptoethanol		
1x SDS-Loading buffer	25 % (v/v)	4x SDS Loading buffer	-20 °C	-
	75 % (v/v)	MQ H_2O		
SOB medium, pH 7	20 g	Tryptone	rt	autoclaved
	5 g	Yeast extract		
	0.96 g	$MgCl_2$		
	0.5 g	NaCl		
	0.186 g	KCl		
	fill to 1 L	H_2O		
SOC medium, pH 7	20 g	Tryptone	-20 °C	Add 20 mM Glucose after autoclaving
	5 g	Yeast extract		
	0.96 g	$MgCl_2$		
	0.5 g	NaCl		
	0.186 g	KCl		
	20 mM	Glucose		
	fill to 1 L	H_2O		
4x Stacking buffer, pH 6.8	0.5 M	Tris HCl pH 6.8	rt	-
	0.4 % (w/v)	SDS		
Streptavidin-TAMRA buffer	5 mM	NaN_3	-	-
	1x	PBS, pH 7.4		
50x TAE-buffer, pH 8.5	2 M	Tris	rt	-
	1 M	Acetic acid		
	50 mM	EDTA		
1x TAE-buffer, pH 8.5	40 mM	Tris	rt	-
	20 mM	Acetic acid		
	1 mM	EDTA		
10x TBS, pH 7.6	200 mM	Tris base	4 °C	-

	1.5 M	NaCl		
1x TBST, pH 7.6	20 mM	Tris base	4 °C	-
	150 mM	NaCl		
	0.1 % (v/v)	Tween 20		
5000x Trace metal mix	20 mM	$CaCl_2$	4 °C	sterile filtered
	10 mM	$MnCl_2$		
	10 mM	$ZnSO_4$		
	2 mM	$CoCl_2$		
	2 mM	$CuCl_2$		
	2 mM	$NiCl_2$		
	2 mM	Na_2MoO_4		
	2 mM	Na_2SeO_3		
	2 mM	H_3BO_3		
	50 mM	$FeCl_3$		
2 M Tris-HCl, pH 8.8	2 M	Tris base	4 °C	-
1 M Tris-HCl, pH 8	1 M	Tris base	4 °C	-
1 M Tris-HCl, pH 6.8	1 M	Tris base	4 °C	-
5x wash buffer	100 mM	Tris HCl pH 8	4 °C	-
	150 mM	Imidazole, pH 8		
	1.5 M	NaCl		
5x wash buffer w/o imidazole	100 mM	Tris HCl pH 8	4 °C	-
	1.5 M	NaCl		
1x wash buffer	20 mM	Tris HCl pH 8	4 °C	-
	30 mM	Imidazole, pH 8		
	300 mM	NaCl		
1x wash buffer w/o imidazole	20 mM	Tris HCl pH 8	4 °C	-
	300 mM	NaCl		
WB blocking buffer	5 % (w/v)	Skim milk powder		
		1x TBST, pH 7.6		
WB incubation buffer	1 % (w/v)	Skim milk powder		-
		1x TBST, pH 7.6		
1x WB Transfer buffer, pH 8.9 - 9.4	48 mM	Tris base	rt	-
	39 mM	Glycine		
	1,3 mM = 0.0375 % (w/v)	SDS		
	20 % (v/v)	MeOH		
2xYT medium, pH 7	16 g	Tryptone	rt	autoclaved
	10 g	Yeast extract		
	5 g	NaCl		
	fill to 1 L	H_2O		

II.2.1.9 Consumables and commercial solutions

Table II.20: List of used consumables

Consumables/Solutions	Supplier
Acrylamide/Bis Solution 37.5:1 (40 % w/v)	Serva Electrophoresis GmbH (Heidelberg, Germany)
Amersham ECL Prime Western Blotting Detection Reagent	GE Healthcare Life Science (Marlborough, USA)
Amersham Protran Nitrocellulose Blotting Membrane, 0.2 µm	GE Healthcare Life Science (Marlborough, USA)
Amicon Ultra 0.5 mL Centrifugal filters, 3,000 NMWL	Merck KGaA (Darmstadt, Germany)
Amicon Ultra 0.5 mL Centrifugal filters, 10,000 NMWL	Merck KGaA (Darmstadt, Germany)
Bolt™ Empty Mini Gel Cassettes	Thermo Fisher Scientific (Waltham, USA)
iBlot 2 NC Regular Top & Bottom Stack	Thermo Fisher Scientific (Waltham, USA)
Instant*Blue* coomassie stain	Expedeon Ltd. (Cambridgeshire, UK)
Bradford Reagent for 0.1-1.4 mg/mL protein	Sigma Aldrich (St. Louis, USA)
Coverslips for microscopy, ø 12 mm	Carl Roth GmbH (Karlsruhe, Germany)
DNA stain Clear G	Serva Electrophoresis GmbH (Heidelberg, Germany)
Electroporation cuvettes, gap width 0.2 cm	Sigma Aldrich (St. Louis, USA)
High Density Nickel Agarose	Jena Bioscience (Jena, Germany)
LC-MS: • Vials (1.5 mL) with PP screw cap ND9, 6 mm hole, silicon white/PTFE red septum, 55° shore A, 1.0 mm (548-0911) • Micro insert 0.1 mL 31 x 3.6 mm (548-0006)	VWR International GmbH (Darmstadt, Gemany)
Microplate, 96 well, PS, F-Bottom (chimney well), µCLEAR®, black, medium binding (655096) for fluorescence measurements	Greiner Bio-One (Kremsmünster, Austria)
Microplate, 96 well, UV-Star®, COC, F-Bottom (chimney well), µCLEAR®, black (655809) for absorption measurements	Greiner Bio-One (Kremsmünster, Austria)
Microscopy: • Coverslips, ø 12 mm • Mounting Medium: Shandon Immu-Mount • Slides, ground edges 90 °	Carl Roth GmbH (Karlsruhe, Germany) Thermo Fisher Scientific (Waltham, USA) VWR International GmbH (Darmstadt, Gemany)
Plastic consumables	Sarstedt (Nümbrecht, Germany)
Quick Coomassie Stain	Serva Electrophoresis GmbH (Heidelberg, Germany)
Rotilabo Blotting paper, thickness 1.5 mm	Carl Roth GmbH + Co KG. (Karlsruhe, Germany)
Tips for pipettes	Starlab GmbH (Hamburg)
UV cuvettes, micro	Brand GmbH + CO KG (Wertheim, Germany)

VIEWseal™, sealing film, transparent (676070)	Greiner Bio-One (Kremsmünster, Austria)

II.2.1.10 Equipment

Table II.21: List of used equipment

Equipment	Supplier
AccuBlock™ Digital Dry Bath D1110-230V	Labnet International Inc. (Edison, USA)
Analog Vortex Mixer	VWR International GmbH (Darmstadt, Gemany)
Bio Imaging System GeneGenius (DNA)	SynGene (Cambridge, UK)
BioPhotometer 6131 for OD_{600} measurements	Eppendorf (Hamburg, Germany)
Centrifuges: • Eppendorf 5804 R • Eppendorf 5424 • Heraeus Multifuge X1R • Heraeus Multifuge X3R • VWR Micro Star 17	 Eppendorf (Hamburg, Germany) Eppendorf (Hamburg, Germany) Thermo Fisher Scientific (Waltham, USA) Thermo Fisher Scientific (Waltham, USA) VWR International GmbH (Darmstadt, Gemany)
Concentrator 5301	Eppendorf (Hamburg, Germany)
Forma™ 905 -86°C Upright Ultra-Low Temperature Freezer	Thermo Fisher Scientific (Waltham, USA)
HPLC: LC20-AT	Shimadzu (Kyoto, Japan)
iBlot2	Thermo Fisher Scientific (Waltham, USA)
Electroporator MicroPulser	Bio-Rad Laboratories Inc. (Hercules, USA)
Fluorometer Fluoromax-4 with Temperature Controller LFI-3751 (5A, 40 W)	Horiba Jobin Yvon (Edison, USA) Wavelength Electronics (Bozeman, USA)
ImageQuant LAS 4000 and LAS 4000mini	GE Healthcare Life Science (Marlborough, USA)
Incubator INCU-Line IL 115	VWR International GmbH (Darmstadt, Gemany)
Incubation shaker Infors HT Multitron Standard	Infors AG (Bottmingen, Germany)
LC-MS: 1260 Infinity 6130 Single Quadrupole	Agilent Technolgies (Santa Clara, USA)
Lyophilizer Christ Alpha 2-4 LDplus	Martin Christ Gefriertrocknungsanlagen GmbH (Osterode am Harz, Germany)
Mini Gel Tank	Thermo Fisher Scientific (Waltham, USA)
Nanophotometer P330	Implen GmbH (Munich, Germany)
NMR: AVHD300, AVHD400, AVHD500	Bruker (Billerica, USA)
PCR Cycler: T100 Thermal Cycler	Bio-Rad Laboratories Inc. (Hercules, USA)
Pipettes	Eppendorf (Hamburg, Germany)
Precision scale ABJ220-4NM	Kern & Sohn GmbH (Balingen, Germany)
Rotavapor: • Rotavapor R-300, R-215	 Büchi Labortechnik AG (Flawil, Switzerland) KNF Neuberger GmbH (Freiburg, Germany)

• Membrane pump SC 920	
PerfectBlue™ Gelsystem Mini S, Mini L and Midi S	VWR Life Science Competence Center (Erlangen, Germany)
PerfectBlue™ 'Semi-Dry' Electro Blotter Sedec M	VWR Life Science Competence Center (Erlangen, Germany)
PowerPac™ Power Supply Basic and HC (for SDS Page)	Bio-Rad Laboratories Inc. (Hercules, USA)
Power Supply EV231 (for DNA electrophoresis)	Consort bvba (Turnhout, Belgium)
SONOPLUS Homogenisator HD2070	BANDELIN electronic GmbH & Co. KG (Berlin, Germany)
Syringe pump LA30	Landgraf Laborsysteme GmbH (Langenhagen, Germany)
Tecan Spark 10M Microplate reader	Tecan Group Ltd (Männedorf, Switzerland)
Thermo-Incubation shaker Thriller	VWR Life Science Competence Center (Erlangen, Germany)
Tumbling roll mixer RM5	Ingenieurbüro CAT M.Zipperer GmbH (Ballrechten-Dottingen, Germany)
UV Lamp (handheld), UVGL-58, 1x 6 W, 254/365 nm	UVP (Cambridge, UK)
UV Lamp: VL-215.L, 2x 15 W, 365 nm	Vilber Lourmat (Marne-la-Vallée Cedex 3, France)
Varian Cary 50 UV-Vis Spectrophotometer	Agilent Technolgies (Santa Clara, USA)
WB Imager Fusion Pulse 6	Vilber Lourmat (Marne-la-Vallée Cedex 3, France)

II.2.1.11 Software and Databases

Table II.22: List of used Software and Databases

Software	Supplier
Adobe Illustrator CS4 14.0.0	Adobe (San José, USA)
ChemDraw Professional 16.0.1.4 (77)	Perkin Elmer Informatics Inc. (Waltham, USA)
Converter Weight to Molar Quantity (http://www.molbiol.ru/ger/scripts/01_04.htmL, last accessed 19.02.19)	www.molbiol.ru/ger
EvolutionCapt Pulse 6 17.03	Vilber Lourmat (Marne-la-Vallée Cedex 3, France)
ExPASy – Compute pI/MW (https://web.expasy.org/compute_pi/, last accessed 20.01.19)	SIB Swiss Institute of Bioinformatics (Lausanne, Switzerland)
ExPASy – ProtParam (https://web.expasy.org/protparam/, last accessed 20.01.19)	SIB Swiss Institute of Bioinformatics (Lausanne, Switzerland)
FluorEssence v2.5.2.0 with Origin v.7.5878	Horiba Jobin Yvon (Edison, USA)
GeneSnap v7.12.06	SynGene (Cambridge, United Kingdom)
ImageReader LAS-4000	

MestReNova v11.0.4-18998	Mestrelab Research, S.L. (Santiago de Compostela, Spain)
NEBaseChanger v1.2.8 (https://nebasechanger.neb.com/, last accessed 24.05.18)	NEB Inc. (Ipswich, USA)
NEBcloner v1.3.8 (https://nebcloner.neb.com/#!/redigest, last accessed 05.12.18)	NEB Inc. (Ipswich, USA)
NEB T_m Calculator v1.9.13 (https://tmcalculator.neb.com/#!/main, last accessed 04.12.18)	NEB Inc. (Ipswich, USA)
OriginPro 2017G b9.4.1.354	OriginLab Corporation (Northampton, USA)
OpenLab CDS ChemStation Edition C.01.07.SR3 [465]	Agilent Technologies (Santa Clara, USA)
PrimerX (http://www.bioinformatics.org/primerx/, last accessed 07.07.2017)	Bioinformatics.org
Promega BioMath Calculator (https://www.promega.com/a/apps/biomath/, last accessed 21.11.18)	Promega (Madison, USA)
PyMOL v1.8.6.1	Schrödinger, LLC (Cambridge, USA)
SciFinder	American Chemical Society (Washington, DC, USA)
SnapGene 3.1.4	GSL Biotech LLC (Chicago, USA)
SnapGene Viewer 3.3.4	GSL Biotech LLC (Chicago, USA)
SparkControl v2.1	Tecan Group Ltd (Männedorf, Switzerland)

II.2.2. Molecular biological methods

II.2.2.1 Isolation genomic and plasmid DNA of *E. coli*

For the isolation of plasmid DNA from 6-8 mL overnight cultures the peqGOLD Plasmid Miniprep Kit was used according to the manufacturer's protocol. The DNA was eluted with 50 µL MQ water instead of TE buffer, as indicated in the protocol. The column was incubated with the water for 1 min before elution. Isolated plasmids were stored at -20 °C.

Qiagen HiSpeed Plasmid Midi and Maxi Kits were used to isolate plasmid DNA according to the manufacturer's protocol from 50-150 mL or 150-250 mL cultures, respectively. Elution was again performed with MQ water instead of TE buffer in a volume of 750 µL.

E. coli DH10B genomic DNA was isolated from 1.5 mL overnight cultures utilizing the GenElute™ Bacterial Genomic DNA Kits according to the manufacturer's protocol. Elution was performed with 200 µL MQ water.

II.2.2.2 Agarose Gel electrophoresis

Separation of DNA samples agarose gel electrophoresis was used. The negatively charged DNA migrates according to its size and conformation in an electric current towards the anode. This can be used to examine a PCR reaction mixture for products or to analyze an analytic or preparative DNA digest.

Gels were cast by dissolving agarose in 1x TAE buffer to obtain a 1 % suspension, which clears upon heating. DNA stain Clear G (4 µL per 100 mL gel) was added to the agarose solution and the gel directly poured into the gel chamber of the PerfectBlue systems and fitted with a suitable comb. Polymerization was complete after 20-30 min and the gel transferred into the electrophoresis chamber, containing 1x TAE buffer, which was also used to overlay the gel. DNA samples were prepared by mixing with 10x DNA loading buffer, before loading into the wells of gel. A DNA standard was used depending on the size of DNA, which was analyzed.

Analytical gels were loaded with 3-5 µL PCR mixture and run at 120 V for 20-25 min. DNA bands were visualized by excitation of the fluorescence of DNA stain Clear G with UV light using the Bio Imaging System GeneGenius.

Preparative scale gels were fitted with combs that formed larger wells, which allowed loading of volumes from 50 µL (e.g. PCR mixture) to 100 µL (e.g. restriction digest). There, gel slices containing DNA bands were cut out with a scalpel and purified with the Monarch DNA Gel Extraction Kit according to the manufacturer's protocol. Since DNA purified by this kit sometimes contained a lot of salt components, it was further purified using Monarch PCR DNA Cleanup Kit for PCR applications. In both cases DNA was eluted with MQ water.

II.2.2.3 Determination of DNA concentration

DNA concentration was determined on the Nanophotometer P330 using absorbance of nucleic acids at 260 nm, as well as DNA purity. The ratio of absorbance of 260 nm to 280 nm gives information about protein contamination, as protein concentration is determined by absorbance at 280 nm. A 260/280 ratio of ≥ 1.8 describes a pure DNA probe, smaller values report protein contamination. Purity is further assessed by measuring the ratio of absorbance of 260 nm to 230 nm, which informs about possible contamination with EDTA, carbohydrates, phenols or other compounds absorbing at 230 nm. There values ≥ 2.0 display pure DNA. MQ H_2O was used as blank.

II.2.2.4 Polymerase Chain Reaction (PCR)

Double-stranded DNA fragments, for example for cloning, were amplified in a polymerase chain reaction (PCR) using Q5·HF DNA polymerase, which exhibits the reaction at a velocity of 20-30 s/kb with an error rate of $1.4x10^{-6}$ mutations per base. The pipetting scheme and the cycling program are displayed in Table II.23 and Table II.24.

Table II.23: Pipetting scheme for Q5 PCR reaction

Total volume	25 µL	50 µL
5x Q5 reaction buffer	5	10
Q5 enhancer	5	10
dNTP (10 mM)	0.5	1
Forward primer (10 µM)	1.25	2.5
Reverse primer (10 µM)	1.25	2.5
Template (50 ng/µL)	0.5	1
Q5·HF DNA polymerase	0.25	0.5
MQ H_2O (nuclease free)	11.25	22.5

Table II.24: PCR cycling program for Q5 polymerase

Cycle step	Temperature	Time	# of cycles
Initial denaturation	98 °C	2 min	1x
Denaturation	98 °C	30 s	32-35x, 25x**
Annealing	Tm [°C]*	30 s	
Elongation	72 °C	30 s/kb	
Final elongation	72 °C	5 or 10 min***	1x
Storage	12 °C	∞	1x

* Tm values were calculated using NEB Tm Calculator v 1.9.13 (NEB Inc., Ipswich, USA)

** # of cycles was depending on DNA size; 35 cycles for DNA < 6000 bp, 32 cycles for ≥ 6000 bp; 25 cycles were used for screening PCR

*** Final elongation was chosen based on elongation time: t ≤ 2 min (5 min), t ≥ 2 min (10 min)

PCR mixtures were analyzed via agarose gel electrophoresis for the presence of desired products. Samples were purified using Monarch PCR DNA Cleanup Kit. If unspecific bands were visible preparative agarose gel electrophoresis was performed, bands cut out and purified with the Monarch DNA Gel Extraction Kit.

II.2.2.5. Site-directed mutagenesis by quickchange mutagenesis

To introduce small, site-specific mutations into a double-stranded DNA sequence site-directed mutagenesis was performed. Until QC primer 40, program Primer X was used to design primers. Polymerases Pfu Turbo or Pfu Ultra II were used for the PCR amplification.

Table II.25: Pipetting scheme for quikchange mutagenesis reaction using Pfu Polymerases

Total volume	50 µL
10x Pfu reaction buffer	5
DMSO (5 %)	2.5
dNTP (40 mM)	2.5
Forward primer (10 µM)	1
Reverse primer (10 µM)	1
Template (25 ng/µL)	1
Pfu Polymerase	1
MQ H_2O (nuclease free)	36

Table II.26: PCR cycling program for Pfu polymerases

Cycle step	Temperature	Time	# of cycles
Initial denaturation	95 °C	30 s	1x
Denaturation	95 °C	30 s	20x
Annealing	55 °C	1 min	
Elongation	68 °C	1 min/kb + 2 min	
Storage	12 °C	∞	1x

To destroy parental plasmid, 1 µL DpnI was added to the reaction mixture after PCR reaction and digested at 37 °C for 4 h. Then 5 µL digested PCR mix were transformed into chemical competent DH10B cells and plated on LB agar with corresponding antibiotics. Plasmid DNA of several resulting clones was isolated and sent for sequencing by GATC with respective sequencing primers.

QC primers 41-58 were designed using NEBaseChanger and PCR reactions were performed with Q5 DNA polymerase. Reactions were conducted according to the pipetting scheme in Table II.23 and the cycling program in Table II.27.

Table II.27: PCR cycling program for site-directed mutagenesis with Q5 polymerase

Cycle step	Temperature	Time	# of cycles
Initial denaturation	98 °C	30 s	1x
Denaturation	98 °C	15 s	25x
Annealing	Tm [°C]	20 s	
Elongation	72 °C	30 s/kb	
Final elongation	72 °C	5 min	1x
Storage	12 °C	∞	1x

After PCR amplification a KLD (Kinase, Ligase, DpnI) reaction including phosphorylation, ligation and template removal was carried out for 1 h at room temperature after the following pipetting scheme:

Table II.28: Pipetting scheme for KLD reaction

Total volume	10 µL
T4 Polynucleotide kinase	0.5
T4 DNA ligase	0.5
Dpn1	0.5
10x T4 DNA ligase buffer	1
PCR product	2
MQ H_2O (nuclease free)	5.5

5 µL KLD reaction were transformed into chemical competent DH10B cells and plated on LB agar with corresponding antibiotics. Plasmid DNA of several resulting clones was isolated and sent for sequencing by GATC with respective sequencing primers.

II.2.2.6 Restriction cloning

Restriction endonucleases are enzymes that are able to cut double-stranded DNA at specific small palindromic sequences to yield blunt or sticky DNA ends. This was used to introduce fragments of DNA into vectors or analyze cloning products.

For an analytical restriction digest two uniquely cutting restriction enzymes were combined that resulted in bands of specific sizes. The test digest was usually carried out with 400-700 ng (Table II.29) of DNA for 1.5 h at 37 °C in an incubator. Then the whole sample was analyzed via agarose gel electrophoresis.

Table II.29: Pipetting scheme for restriction digest

Total volume	**25 µL**	**50 µL**	**100** µL
DNA	1 µg in x µL	2.5 µg in x µL	5 µg in x µL
Restriction enzyme 1 (10 U/µL)	1 µL	2.5 µL	5 µL
Restriction enzyme 2 (10 U/µL)	1 µL	2.5 µL	5 µL
10x Cutsmart or NEB 3.1 buffer	2.5 µL	5 µL	10 µL
MQ H_2O (nuclease free)	20.5 - x µL	40 - x µL	80 - x µL

Unidirectional restriction cloning was achieved by combination of restriction enzymes that produce non-compatible ends. Therefore, 5 µg vector was linearized by double restriction digest for 1.5 h at 37 °C in an incubator and afterwards separated and purified via preparative agarose gel electrophoresis and purification by Monarch DNA Gel Extraction Kit, followed by Monarch PCR DNA Cleanup Kit if necessary. The linearized vector was then dephosphorylated overnight at 37 °C with Antarctic phosphatase, followed by purification by Monarch PCR DNA Cleanup Kit.

Table II.30: Dephosphorylation of linearized vectors

Total volume	**30 µL**
Linearized vector DNA	16 µL
Antarctic phosphatase (5 U/µL)	1 µL
10x Antarctic phosphatase buffer	2 µL
MQ H_2O (nuclease free)	9 µL

DNA inserts were amplified by Q5 PCR and thereby introducing restriction sites via overhangs if necessary. Double digest was usually conducted with 2.5 µg amplified insert for 1.5 h at 37 °C in an incubator, which was subsequently purified with the Monarch PCR DNA Cleanup Kit.

The Promega BioMath calculator was used to determine the concentration of double-stranded DNA in pmol/µl, which applies the following equation:

$$\text{pmol (DNA)} = \text{µg (DNA)} \times \frac{\text{pmol}}{660\ \text{pg}} \times \frac{10^6\ \text{pg}}{1\ \text{µg}} \times \frac{1}{\text{N}}$$

N = number of nucleotides
660 pg/pmol = the average molecular weight of a nucleotide pair

Linearized vector and insert were recombined by ligation of compatible the sticky ends with T4 DNA ligase at 16 °C overnight. DNA ratio between vector and insert was dependent on the size of the fragments. The smaller the insert size, the higher the amount in the ligation reaction. The pipetting scheme for the DNA ratios usually used is shown in Table II.31.

Table II.31 Pipetting scheme for ligation

DNA ratio vector:insert	1:3	1:5	1:7
Amount of linearized vector	100 ng in x µL	100 ng in x µL	100 ng in x µL
Amount of insert	y µL	y µL	y µL
T4 DNA ligase	1 µL	1 µL	1 µL
10x T4 DNA ligase buffer	2 µL	2 µL	2 µL
MQ-H_2O	17 µL - x µL - y µL	17 µL - x µL - y µL	17 µL - x µL - y µL
Total volume	20 µL	20 µL	20 µL

x = volume of vector; y = volume of insert

5 µL of the ligation was transformed into 100 µL chemical competent *E.coli* DH10B cells. 150 µL of the cell suspension were plated onto LB agar with corresponding antibiotics, then the cells centrifuged at 4200 rpm for 3 minutes, the pellet resuspended in 150 µL, which were spread on a second LB agar plate with corresponding antibiotics.

II.2.2.7 NEBuilder HiFi DNA Assembly cloning

Insertion of fragments via HiFi DNA assembly cloning was simulated using SnapGene. Primers were designed with an overlap on the 5'-end of around 18-25 bp to the next fragment and a T_m value for assembly of the overhangs > 50 °C. Fragments for assembly were amplified with a Q5 PCR and purified with the Monarch PCR DNA Cleanup Kit. The Promega BioMath calculator was used to determine the concentration of double-stranded DNA in pmol/µl.

Table II.32: Pipetting scheme for HiFi DNA assembly

	2-3 fragment assembly	
Total DNA content	0.1 pmol in x µL	0.1 pmol in x µL
DNA ratio vector: insert	1:2	1:3
Master-Mix (2x)	5 µL	5 µL
MQ-H_2O	5 µL - x µL	5 µL - x µL
Total volume	10 µL	10 µL

x = total volume of fragments

The reaction volume should contain at least 50 ng of vector. For smaller inserts (< 200 bp) the DNA ratio of vector to insert can be adjusted to 1:5.

The reaction was incubated at 50 °C in a PCR cycler (lid temperature 105 °C) for 75 min. 5 µL of the assembly was transformed into 100 µL chemical competent *E. coli* DH10B cells. 150 µL of the cell suspension were plated onto LB agar with corresponding antibiotics, the cells centrifuged, the pellet resuspended in 150 µL, which were spread on a second LB agar plate with corresponding antibiotics.

II.2.2.8 Sequencing

For sequencing of DNA the LIGHTRUN service of GATC Biotech (now part of Eurofins Genomics GmbH, Ebersberg, Germany) was used. Therefore, 5 µL of purified template (plasmid 80-100 ng/µL, PCR product 20-80 ng/µL) was mixed with 5 µL sequencing primer (5 µM) in a 1.5 mL Eppendorf tube, which was labeled with a barcode. The sequencing data could be downloaded from the GATC Biotech homepage on the next day.

II.2.3 Microbiological and protein biochemical methods

II.2.3.1 Solid bacterial culture

To obtain single colonies commercially available *E.coli* competent cells were streaked with an inoculation loop on LB agar plates containing the corresponding antibiotics if necessary. Otherwise 50-100 µL of cell suspension were spread-out equally on a LB agar plate. After permeation into the agar the plates were cultivated in an incubator at 37 °C overnight. For storage plates were sealed with Parafilm and kept a 4 °C.

II.2.3.2. Liquid bacterial culture

Liquid bacterial cultures of *E. coli* were prepared in LB or 2xYT medium with the corresponding antibiotics if necessary. The culture volumes were dependent on its purpose. Pre- or overnight cultures were performed in 50 mL medium in autoclaved 100 mL Erlenmeyer flasks covered with aluminum foil. Cultures for plasmid isolation contained 6-8 mL medium and test expressions were carried out in 4 mL medium. Both were performed in 14 mL semi-sealed cell culture tubes. Protein expression varied from 10-200 mL medium and was performed in Erlenmeyer flasks. All cultures were incubated at 37 °C in an incubation shaker at 200 rpm overnight, if not indicated otherwise. Cell density was determined in 1 cm semi-micro-cuvettes by measuring the optical density at 600 nm (OD_{600}) after blanking with the used medium.

II.2.3.3 Preparation of competent cells

II.2.3.3.1 Chemical competent *E. coli* cells

Overnight cultures were diluted in 100 mL fresh 2xYT medium without or with corresponding antibiotics to an OD_{600} of 0.05 and incubated at 37 °C at 200 rpm to an OD_{600} of 0.5-0.6. Then, cells were cooled on ice for 30 min before being centrifuged at 4000 g and 4 °C for 10-15 min. The supernatant was discarded and the pellet resuspended in 100 mL ice-cold 100 mM $CaCl_2$ solution and incubated on ice for 30 min. The cells were centrifuged at 4000 g and 4 °C for 10-15 min and the supernatant discarded. The pellet was resuspended in 10 mL ice-cold 100 mM $CaCl_2$ solution, 2 mL 100 % glycerol added and the cells aliquoted in 110 µL in precooled 1.5 mL Eppendorf tubes. The aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

II.2.3.3.2 Electro-competent *E. coli* cells

Overnight cultures were diluted in 500 mL fresh 2xYT medium without or with corresponding antibiotics to an OD_{600} of 0.05 and incubated at 37 °C at 200 rpm to an OD_{600} of 0.5-0.6. Then, cells were cooled on ice for 30 min before being centrifuged at 4000 g and 4 °C for 10-15 min. The supernatant was discarded and the pellet resuspended in 500 mL ice-cold 10 % glycerol and centrifuged at 4000 g and 4 °C for 10-15 min. The supernatant was discarded and the pellet again resuspended in 500 mL ice-cold 10 % glycerol and centrifuged at 4000 g and 4 °C for 10-15 min. The supernatant was discarded and then the pellet resuspended in 50 mL ice-cold 10 % glycerol and centrifuged at 4000 g and 4 °C for 10-15 min. The final pellet was resuspended in 2-3 mL 10 % glycerol and aliquoted in 110 µL in precooled 1.5 mL Eppendorf tubes. The aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

II.2.3.4 Transformation of *E. coli* cells

II.2.3.4.1 Transformation by electroporation

An aliquot of electro-competent cells was thawed on ice for 5-10 min and then 1 µL plasmid (or 1 µL of each plasmid for a co-transformation) was added to the cells, which were transferred after incubation on ice for 5 min to precooled electroporation cuvettes with a gap of 0.2 cm and electroporated at 2.0 kV. The cells were directly rescued in 900 µL SOC medium at room temperature and incubated at 37 °C and 200 rpm for 1 h. The resulting cell suspension was either used to inoculate 50 mL of overnight culture or 50-100 µL were spread on LB agar plates with corresponding antibiotics for solid cultures.

II.2.3.4.2 Chemical transformation by heat shock

Chemical competent cells were thawed on ice for 5-10 min, then 1 µL plasmid (or 1 µL of each plasmid for a co-transformation) was added to the cells and incubated on ice for 10 min. Cells were heat shocked in a water bath with 42 °C for 45 s, followed by incubation on ice for 5 min. Afterwards 900 µL SOC medium were added at room temperature to the Eppendorf

tube and the cells incubated at 37 °C and 200 rpm for 1 h. The resulting cell suspension was either used to inoculate 50 mL of overnight culture or 50-100 µL were spread on LB agar plates with corresponding antibiotics for solid cultures.

II.2.3.4.3 Test for transformation efficiency

The transformation efficiency of the cells, which stands for the number of colony forming units (cfu) growing per 1 µg of transformed plasmid DNA, was tested with flash frozen aliquots that were stored at -80 °C for 24 h. Therefore a single and a co-transformation was performed and 100 µL of the 1 mL cell suspension in SOC was equally plated on an LB agar plates with corresponding antibiotics, which were incubated at 37 °C overnight. After incubation at 37 °C overnight the resulting colonies were counted to determine the number of cfu per plate which corresponds to 100 µL of the transformed suspension. This number was multiplied by 10 to calculate the number of cfu per mL, which was divided by the amount of used DNA [ng] to determine the number of cfu per µg (DNA). Freshly prepared cells were plated on LB agar plates containing Kan, Tet, Amp or Cam and incubated at 37 °C overnight to check for contamination with plasmid DNA.

II.2.3.5 Directed evolution of PylRS synthetases

II.2.3.5.1 General information

A directed evolution approach was chosen for the selection of mutant Pyrrolysyl-tRNA synthetases for new unnatural amino acids. Therefore, libraries were used, where up to five different positions in the active site of the wild type PylRS were randomized by saturation mutagenesis. Libraries containing the following mutations were used:

Table II.33: List of used libraries

Library	AR	mutations
AB2	Kan	**A267T**, L270, Y271, N311, C313, Y349
AB3	Kan	L270, Y271, **L274A** N311, C313, Y349
ABshuffle	Kan	Recombined version of PylRS variants with known uAA specificities and a fixed error-prone *M. barkeri* PylRS wildtype proportion
Libfw	Amp	A267, Y271, L274,C313, M315
Nicholas RS5	Kan	Y271, N311, Y349, V366, W382

After co-transformation for the positive selection step the statistical significance of the coverage of each library member after was calculated by solving the following equation with 90 % confidence:

$$N = \frac{\ln(1 - \text{confidence})}{\ln\left(1 - \left(\frac{1}{\text{theoretical size}}\right)\right)}$$

N = Number of clones needed to reach the aimed confidence
Confidence = % of statistical confidence, e.g. 0.9 (90 %)
Theoretical size = Theoretical amount of members within the library

The theoretical size of a library with five randomized positions was calculated to 32^5 = 3.3x 10^7 library members. For a confidence of 90 % a 2.3-fold coverage of 7.7x 10^7 has to be reached.

Table II.34: List of used evolution reporter plasmids

Plasmid	bb	Nucleic acid	AR
D65	pREP	CAT111TAG	Tcn
D7L	pYOBB2	Barnase3TAG,45TAG	Cam
pPyltsfGFP150TAG	pPylt	sfGFP150TAG	Tcn
pPyltsfGFP150TAG CAT	pPylt	sfGFP150TAG CAT111TAG	Tcn

The positive selection plasmid D65 with a size of approx. 10.3 kb encodes a tetracycline resistance cassette (Tcn), a constitutively expressed mutant *M. barkeri* Pyl-$tRNA_{CUA}$ and a constitutively expressed chloramphenicol acetyltransferase with a TAG codon at position 111. D7L (approx. 7.4 kb) is the negative selection plasmid containing a chloramphenicol resistance cassette, a constitutively expressed mutant *M. barkeri* Pyl-$tRNA_{CUA}$ and a L-arabinose inducible barnase gene with TAG codons in positions 3 and 45. The dual reporter plasmid pPyltsfGFP150TAG CAT for the read-out has a size of approx. 7.7 kb and encodes a tetracycline resistance cassette, a constitutively expressed mutant *M. barkeri* Pyl-$tRNA_{CUA}$, a constitutively expressed chloramphenicol acetyltransferase with a TAG codon at position 111 and a L-arabinose inducible gene for the expression of sfGFP150TAG. The single reporter plasmid pPyltsfGFP150TAG with a size of approx. 6.2 kb and differs to the dual reporter plasmid pPyltsfGFP150TAG CAT in the chloramphenicol acetyltransferase with a TAG codon at position 111, which is not encoded here.

Selection plasmids D65, D7L and sfGFP CAT were chemically transformed into DH10B cells, plated onto LB agar with the corresponding antibiotic, incubated at 37 °C overnight and then stored at 4 °C up to 2-3 months for further use.

II.2.3.5.2 Positive selection

A DH10B_D65 single clone was used to inoculate 50 mL 2xYT media (+ Tcn) for an overnight culture at 37 °C. The overnight culture was diluted in fresh 100 mL 2xYT medium (+Tcn) to an OD_{600} of 0.05 and grown at 37 °C to an OD_{600} of 0.4 - 0.5, before preparing electro-competent DH10B_D65 cells. Therefore, cells were pelleted in 50 mL falcons and the pellets washed twice with 50 ml ice cold 10 % glycerol each. Subsequently, the cells were pooled into one 50 mL falcon and washed with 50 mL ice cold 10 % glycerol. The final pellet was resuspended in 100 µL 10 % glycerol to yield approx. 200 µL electro-competent DH10B_D65 cells.

1 µg library DNA was transformed into 4x 50 µL freshly prepared electro-competent DH10B_D65 cells by electroporation with the standard settings, rescued with 950 µL SOC each and incubated at 37 °C for 1 h, followed by inoculation of 1 L 2xYT (+ Tcn and Ab(library)). The cells were incubated at 37 °C in an incubation shaker at 200 rpm. After 15 min incubation 1 mL medium was taken and used to test the transformation efficiency and coverage by dilution of the transformed library. The remaining 1 L culture was incubated overnight at 37 °C. For the dilution series, 100 µL were directly plated on LB agar (+ Tcn and Ab(library)), which corresponds to a dilution of 10^{-4}. 100 µL were diluted with 900 µL 2xYT and then 100 µL plated on LB agar (≙10^{-5}). This was continued until a dilution of 10^{-8} was achieved. The plates were incubated at 37 °C overnight. The number of clones on the 10^{-7} and 10^{-8} dilution plates were counted the next day and used for determination of transformation efficiency.

If the transformation efficiency yielded a number of clones > $8x10^7$ the 1 L 2xYT overnight culture was diluted to an OD_{600} of 0.05 in 100 ml fresh 2xYT (+ Tcn and Ab(library)). Cells were incubated until an OD_{600} of 0.6, then aliquoted into 10 mL cultures and 2 mM uAA was added besides one control sample without uAA. Cultures were grown for 3 h at 37 °C and 200 rpm and then the OD_{600} determined, which is supposed to be around 2-3. Finally, 600 µL of the 10 mL cell suspension was plated on one 200 mL positive selection plate ().

Table II.35: Pipetting scheme for positive selection plates

Volume	Components	Comments
40 mL	5x M9 salts	Always prepare fresh: heat 5 % agarose until liquid and cool in water bath to 50 °C. Mix all other ingredients and warm in water bath to 50 °C. Combine and directly pour plates.
200 µ	1 M MgSO4	
60 µ	1 M $CaCl_2$	
mL	40 % Glucose	
2 mL	leucine (4 mg/mL)	
40 µL	Trace metal mix	
200 µL	Biotin	
200 µL	Thiamin	
2 mL	uAA (100 mM)	
240 µL	Cam	
200 µL	Tet	
200 µL	AB (depending on library)	
100 mL	5 % agarose	
fill to 200 mL	MQ H2O	

After 48 h incubation at 37 °C clones were harvested by scraping them off the agarose with 50 mL 2xYT (+ Tcn and Ab(library)). The resulting cell suspension was incubated at 37 °C for 2 h in an incubation shaker, before isolating the DNA by midiprep with the Qiagen HiSpeed Plasmid Midi Kit.

5 µg isolated DNA was separated by agarose gel electrophoresis at 145 V for 1.75 h. The DNA band at around 3.3 kb, which corresponds to the library DNA, extracted using the Monarch DNA Gel Extraction Kit.

II.2.3.5.3 Negative selection

Electro-competent DH10B_D7L cells were prepared according to the description in chapter II.2.3.5.2.

100 ng purified library DNA from the positive selection was transformed into 2x 50 µL freshly prepared electro-competent DH10B_D7L cells by electroporation with the standard settings, rescued with 950 µL SOB each and incubated at 37 °C for 1 h. The cells were combined and 600 µL plated onto one 200 mL negative selection plate and 1.4 mL on a second 200 mL negative selection plate. The plates were incubated at 37 °C overnight, followed by harvesting the cells and isolation and purification of library DNA described in chapter II.2.3.5.2.

Table II.36: Pipetting scheme for negative selection plates

Volume	Components	Comments
200 mL	LB agar	Cool agar before addition of antibiotics and arabinose
200 µL	Cam	
200 µL	AB (depending on library)	
2 mL	20% Arabinose	

II.2.3.5.4 Read-out

Electro-competent DH10B_sfGFP CAT cells were prepared according to the description in chapter II.2.3.5.2.

100 ng purified library DNA from the negative selection was transformed into 2x 50 µL freshly prepared electro-competent DH10B_sfGFP CAT cells by electroporation with the standard settings, rescued with 950 µL SOB each and incubated at 37 °C for 1 h. Cells were pelleted, half of the supernatant discarded, followed by resuspension and plating on one 200 mL autoinduction plate (+ Tcn, Cam and Ab(library)). As positive control, 1 µL D4 plasmid was transformed into 50 µL cells and plated on auto-induction plates containing BocK. The plates were incubated at 37 °C for 48-60 h. Green colonies, due to the expression of sfGFP, were picked and used to inoculate 1 mL non-inducing medium (+ Tcn and Ab(library)) in a 96 well megablock. 6 clones from the positive control were also picked.

Table II.37: Pipetting schemes for liquid and solid media used for the read-out

Medium	**Volume**	**Components**	**Comment**
Autoinduction plates	10 mL	5% Aspartate	Always prepare fresh: heat 5 % agarose until liquid and cool in water bath to 50 °C. Mix all other
	10 mL	10% Glycerol	

	8 mL	25x 18 AA mix	ingredients and warm in water bath to 50 °C. Combine and directly pour plates.
	4 mL	50x M	
	2 mL	Leucine (4 mg/mL)	
	500 µL	20% Arabinose	
	400 µL	1 M MgSO4	
	50 µL	40% Glucose	
	40 µL	Trace metal mix	
	200 µL	Biotin	
	200 µL	Thiamin	
	800 µL	India Ink	
	2 mL	uAA (100 mM)	
	200 µL	Tet	
	200 µL	AB (depending on library)	
	100 mL	5% agarose	
	fill to 200 mL	MQ H_2O	
Non-inducing medium	5 mL	5% Aspartate	always prepare fresh
	4 mL	25x 18 AA mix	
	2 mL	50x M	
	1 mL	leucine (4 mg/mL)	
	200 µL	1 M MgSO4	
	0.4 mL	40% Glucose	
	20 µL	Trace metal mix	
	1 mL	uAA (100 mM)	
	100 µL	AB1	
	100 µL	AB2	
	fill to 100 mL	MQ H2O	
Autoinducing medium	5 mL	5% Aspartate	always prepare fresh
	5 mL	10% Glycerol	
	4 mL	25x 18 AA mix	
	2 mL	50x M	
	1 mL	leucine (4 mg/mL)	
	250 µL	20% Arabinose	
	200 µL	1M MgSO4	
	125 µL	40% Glucose	
	20 µL	Trace metal mix	
	1 mL	uAA (100 mM)	
	100 µL	AB1	
	100 µL	AB2	
	fill to 100 mL	MQ H_2O	

After 24 h 50 µL non-inducing culture was used to inoculate two 96 well megablocks, one containing 1 ml autoinducing medium (+ Tcn and Ab(library)) with 2 mM uAA and one without uAA. After 36-48 h 150 µL were transferred into 96 well microplates and sfGFP fluorescence and OD_{600} were determined on a Tecan Spark 10M Microplate reader. The plate was shaken (amplitude 3.5) during measurements and GFP fluorescence excited at 480 nm

and measured at 527 nm. Fluorescence was normalized to the OD_{600} and then the fluorescence of the resulting values of each well incubated with the uAA divided by the resulting values of each matching well incubated without the uAA.

Clones with ratio values ≥ 3 were inoculated into 50 mL 2xYT (+ Tcn and Ab(library)) for overnight cultures and used the next day to express sfGFP150uAA. After His-tag purification protein mass was determined by LC-MS analysis. If incorporation of the uAA displayed the correct mass, DNA of the corresponding clones was isolated and separated by agarose gel electrophoresis. Plasmid DNA which encodes the PylRS gene sequence (approx. 3.3 kb) was extracted and sent for sequencing to identify mutations of the evolved PylRS mutants.

II.2.3.6 Protein expression via amber suppression

Overnight cultures of *E. coli* cells co-transformed with plasmids containing genes for tRNA synthetase mutants, tRNA and proteins of interest were diluted in 50 mL fresh 2xYT or autoinducing medium with corresponding antibiotics to an OD_{600} of 0.05 and incubated at 37 °C at 200 rpm until OD_{600} of 0.3. For test expressions, 4 mL of culture was aliquoted into 14 mL semi-sealed cell culture tubes and supplied with the corresponding amount of uAA stock solution besides one negative control sample without uAA and grown to OD_{600} of 0.6. There, protein expression performed in 2xYT medium was induced with a final concentration of 0.02 % arabinose or 1 mM IPTG, depending on promoters of the used plasmids. Autoinduction medium autoinduced protein expression after consumption of the containing glucose in the media, which correlates with an OD_{600} of about 0.7. Protein expression was performed in volumes from 50-200 mL with the corresponding amount of uAA stock solution overnight without a negative control sample, but otherwise performed like the test expressions.

II.2.3.7 Cell harvest and lysis of *E. coli*

For test expressions, 1 mL samples were taken and centrifuged for 5 min at full speed, the supernatant discarded and the pellets frozen for SDS-PAGE analysis.
Expression cultures were harvested in 50 mL sterile falcons and centrifuged at 4000 g and 4 °C for 10-15 min. After discarding the supernatant the pellet was either flash frozen in liquid nitrogen and stored at -20 °C or directly resuspended in lysis buffer (20 mL per 1 g pellet). The cells were disrupted using sonification. Therefore, the tip of the sonotrode was dipped into the ice-cooled cell suspension without touching the walls of the falcon. The cells were sonicated (80 % amplitude, 50 % cycle) for 5 min with pauses for cooling after 1.5 and 3 min. The lysed cell suspension was centrifuged at 14000 g and 4 °C for 20-30 min. The cleared lysate was purified via nickel agarose beads.

II.2.3.8 His-tagged protein purification via nickel agarose

0.2-0.5 mL high density nickel agarose slurry in 20 % EtOH was added to a 50 mL falcon tube containing 20 mL 1x wash buffer. The falcon was inverted several times and then centrifuged at 1000 g and 4 °C for 10 min. The supernatant was discarded and washed once more. After centrifugation at 1000 g and 4 °C for 10 min the supernatant was discarded and

the cleared lysate added to the washed nickel agarose beads. The resulting suspension was incubated at 4 °C for 1 h. Then, the slurry was applied to an empty plastic column and the flow through collected. The beads were washed with 20 CV 1x wash buffer and then the protein eluted in 0.4-1 mL fractions with elution buffer. All fractions were analyzed by SDS – PAGE. Pure fractions were combined and concentrated and washed with 1x wash buffer without imidazole using amicons with the corresponding cutoff.

II.2.3.9 Determination of protein concentration

II.2.3.9.1 using Nanophotometer

The concentration of the purified protein was determined by measurement of absorbance at 280 nm on the Nanophotometer. Absorbance at 280 nm correlates to the amino acid composition, especially the presence of tryptophan or tyrosine residues in the protein. The molar excitation coefficient is therefore protein-specific and was determined for each protein using ProtParam. It gave a value of 18910 $M^{-1}cm^{-1}$ for sfGFP and 15470 $M^{-1}cm^{-1}$ for myoglobin. Ubiquitin does not contain any tryptophans; therefore protein concentration could not be determined using absorbance measurements on the Nanophotometer. There, concentration was determined using Bradford assay.

II.2.3.9.2 using Bradford assay

Determination of protein concentration with Bradford reagent is based on the formation of a complex between protein and the dye Brilliant Blue G. The complex causes a shift in the absorption maximum of the dye from 465 to 595 nm. The absorption is proportional to the protein concentration. The assay was performed after protocol A in the Technical Bulletin (B6916, Sigma Aldrich (St. Louis, USA) https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/Bulletin/b6916bul.pdf, last accessed 26.02.19) of the commercially available Bradford reagent by mixing 1.5 mL Bradford Reagent and 0.05 mL sample directly in a cuvette and measuring absorbance at 595 nm after 15 min incubation at room temperature. Protein concentration was calculated from a calibration curve generated by measuring varying concentrations of BSA (0, 0.25, 0.5, 0.75 and 1 mg/mL in 1x wash buffer) as protein standard in Bradford reagent.

II.2.3.10 SDS-PAGE electrophoresis

Discontinuous sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was used to analyze pellets from test expressions and protein samples. The unfolded, reduced and negatively charged proteins are separated in a vertical polyacrylamide gel in an electric field due to their size. Smaller proteins run faster in the electric field than bigger ones. The SDS gels were prepared according to the following table:

Table II.38: Pipetting plan for the preparation of SDS gels

Stock solution	Stacking gel (4 gels)	15 % Resolving gel (4 gels)
40 % w/v Acryl amide/Bis Solution 37.5:1	1	10.05
4x resolving gel buffer [mL]	-	6.69
4x stacking gel buffer [mL]	3.5	-
ddH_2O [mL]	6.5	10.05
APS (10 %) [µL]	100	267.75
TEMED [µL]	10	26.78

Polyacrylamide gels were cast in empty Bolt mini gel cassettes, owing a size of 8 cm x 8 cm gel size with 1 mm thickness. APS initiates polymerization and TEMED functions as a catalyst, therefore they were added directly before gels were poured into the cassettes. First, the resolving gel mixture was prepared, poured into the Bolt cassettes and overlaid with isopropanol until polymerization was complete. Then, isopropanol was removed and the stacking gel mixture poured onto the resolving gel and fitted with a 10 or 12-well comb. Polymerized SDS gels were stored in wet paper at 4 °C.

Pellets from test expressions were resuspended in 1x SDS loading buffer (100 µL per OD_{600} = 1) and heated to 95 °C for 10 min, followed by centrifugation at full speed for 10 min. 10 µL of the sample were loaded onto the SDS gel, next to 5-8 µL of marker. Protein samples were prepared in 4x SDS loading buffer and heated to 95 °C for 2-10 min. There the concentration of protein was adjusted to loading 1-2 µg per lane.

Protein samples were separated in a Mini gel tank at 110-200 V (mini gels). Then, the Bolt cassettes were opened and the gels either stained with Quick Coomassie Stain or Instant*Blue* stain used for a WB analysis. Coomassie stained gels were destained with deionized water.

II.2.3.11 Western Blot analysis

For a more sensitive detection of His-tagged proteins or proteins labeled with biotinylated probes protein samples were analyzed with Western blot. Therefore, proteins separated by SDS-PAGE were transferred onto a nitrocellulose (NC) membrane using iBlot 2 Dry Blotting system (program P0 for 7 min: 20 V for 1 min, followed by 23 V for 4 min and 25 V for 2 min) and the corresponding iBlot 2 NC Regular Top & Bottom Stacks. Alternatively, a PerfectBlu semi-dry electro blotter Sedec M was used for transfer and the stacks self-

assembled onto the anode from 3 pieces of blotting paper (1.5 mm) soaked in WB transfer buffer, NC membrane, gel and another 3 pieces of blotting paper (1.5 mm) soaked in WB transfer buffer. The transfer was performed 100 mA per gel for 1.5 h, followed by blocking the membrane in a 50 ml falcon in 5 % skimmed milk powder in TBS-T (WB blocking buffer) for 1 h at room temperature. 1 ml of WB blocking buffer was diluted with 5 mL TBS-T to obtain the incubation buffer, where anti-His_6-peroxidase was added in a 1:5000 dilution. HRP-conjugated streptavidin was used 1:5000 diluted in incubation buffer bind to biotinylated proteins. Alternatively streptavidin conjugated TAMRA was used (1:5000 in TBST-T + 0.05 % SDS). Antibodies or streptavidin were either allowed to bind proteins on the membrane for 1 h at room temperature or at 4 °C overnight. The membrane was washed 3x for 15 min with TBS-T to remove excess antibody or streptavidin. To visualize the protein bands with chemiluminescence, membranes were treated with 1 mL of a freshly prepared 1:1 solution of solutions A (peroxide) and B (luminol) of the ECL Prime Western blot detection reagent. Chemiluminescence was imaged using WB Imager Fusion Pulse 6 with the corresponding program EvolutionCapt Pulse. TAMRA fluorescence was excited at 520 nm and captured using an ImageQuant LAS 4000 with a 575DF20 Cy3 filter.

II.2.3.12 *In vitro* labeling of proteins

II.2.3.12.1 Labeling of purified proteins for LC-MS *in vitro*

II.2.3.12.1.1 Labeling of POI_uAA

His-tag purified POI_uAA was diluted with MQ H_2O to a final concentration of 10 µM in 50 µL in a MS Vial with 0.1 mL micro insert. Small molecules probes were added in a 10 to 100-fold excess and 1-2 µL protein sample analyzed via LC-MS (C4, 5-55 % B in 5 min) at certain time points (e.g. 30 min, 1, 2, 5 and 24 h).

II.2.3.12.1.2 Labeling of POI_uAA with photo-DMBO and other cyclopropenone-caged compounds

His-tag purified POI_uAA was diluted in MQ H_2O to a final concentration of 10 µM in 50 µL in a MS Vial with 0.1 mL micro insert. photo-DMBO (or other cyclopropenone-caged) compounds were added in a 10 to 100-fold excess, the iEDDAC reaction induced at 365 nm with UV Lamp VL-215.L for 10 min and then followed over time. Therefore, 1- µL protein sample was analyzed via LC-MS (C4, 5-55 % B in 5 min) at certain time points (e.g. 20, 35 min, 1, 2, 5 and 24 h).

II.2.3.12.2 Labeling of proteins in *vitro* for SDS in gel fluorescence imaging

Purified sfGFP-N150mTetK was diluted in MQ water to a final concentration of 10 µM in a volume of 100 µL in a 1.5 mL Eppendorf tube. BCN-TAMRA or photo-DMBO-Cy5 (final concentration 100 µM) were added in 10-fold excess and the reaction followed over time. Therefore, 8 µL samples were taken at the following time points: 0, 30 s, 1, 2, 5, 10, 30, 60 and 120 min and the fluorophore quenched with 4 µL dipyridyl-1,2,4,5-tetrazine (50 mM). 3 µL 4x SDS loading buffer were added, the samples cooked at 95 °C for 10 min and the whole samples loaded onto an SDS-PAGE. After separation labeling was visualized by

fluorescence using an ImageQuant LAS 4000 (TAMRA: ex. 520 nm, 575DF20 Cy3 filter; Cy5: ex. 630 nm, R670 Cy5 filter), followed by staining with Quick Coomassie Stain (Serva) and destaining in deionized water.

II.2.3.13 Labeling of proteins in lysate

II.2.3.13.1 Lysate labeling with BCN- or TCO-TAMRA

5 mL of *E. coli* culture expressing POI (sfGFP-NXTAG-His6, X = 40 or 150; Myo-S4TAG-His6 or Ub-K6TAG-His6) without uAA, with BocK, BCNK, TetK or mTetK were harvested, washed 3x 10 % (v/v) DMSO/PBS and then sonicated in 5 mL lysis buffer (20 mM Tris pH 8.0, 30 mM imidazole pH 8.0, 300 mM NaCl, 0.175 mg/mL PMSF, 0.1 mg/mL DNase I and one cOmplete^{TM} protease inhibitor tablet (Roche)). The lysed cells were centrifuged (20 min, 14.000 × g, 4 °C) and the cleared lysate was flash frozen and stored at -80 °C for labeling experiments.
Cleared *E.coli* lysate (adjusted to the same OD_{600}) was thawed on ice and 50 µL taken for labeling experiments. BCN- or TCO-TAMRA were added to a final concentration of 2-4 µM and the reaction performed at 37 °C for 1 h at 350 rpm or samples were taken at different time points for time-dependent labeling: 15 min, 30 min, 45 min and 60 min. 16 µL 4x SDS loading buffer were added and the samples heated at 95 °C for 10 min. 14 µL samples were analyzed by a 15 % SDS PAGE. Labeling was visualized by excitation of TAMRA in-gel fluorescence at 520 nm utilizing an ImageQuant LAS 4000 with a 575DF20 Cy3 filter, followed by staining with Quick Coomassie Stain and destaining in deionized water.

II.2.3.13.2 sfGFP-N40 lysate labeling with photo-DMBO-Cy5

5 mL of *E. coli* culture expressing sfGFP-N40TAG without uAA, with BocK, BCNK or mTetK were harvested, washed 3x 10 % (v/v) DMSO/PBS and then sonicated in 5 mL lysis buffer (20 mM Tris pH 8.0, 30 mM imidazole pH 8.0, 300 mM NaCl, 0.175 mg/mL PMSF, 0.1 mg/mL DNase I and one cOmplete^{TM} protease inhibitor tablet (Roche)). The lysed cells were centrifuged (20 min, 14.000 × g, 4 °C) and the cleared lysate was flash frozen and stored at - 80 °C for labeling experiments.

Cleared *E. coli* lysate (adjusted to the same OD_{600}) was thawed on ice. 45 µL lysate was mixed with 20 % DMSO (final concentration) and photo-DMBO-Cy5 and DMBO-Cy5 added to a final concentration of 10 µM. The reaction was induced at 365 nm with UV Lamp VL-215.L for 30 min and then incubated at 37 °C for 1 h at 350 rpm. The reaction was quenched with 5 µL BCN-OH (100 mM) for 15 min and then 18 µL 4x SDS loading buffer were added and the samples heated at 95 °C for 10 min. 15 µL samples were analyzed by a 15 % SDS PAGE. Labeling was visualized by fluorescence using an ImageQuant LAS 4000 (ex. 630 nm, R670 Cy5 filter), followed by staining with Quick Coomassie Stain and destaining in deionized water.

II.2.3.14 Labeling of proteins *in cellulo*

II.2.3.14.1 Labeling of OmpC-Y232mTetK on live *E. coli* with BCN-TAMRA

E. coli cultures expressing OmpC-Y232TAG without uAA, with BocK or mTetK were harvested in 0.5-1 mL samples (pellet resuspended in 1 mL ≙ OD_{600} = 2), pelleted at 4000 g for 3 min and the pellet washed 3x with 1 mL 10 % (v/v) DMSO/PBS buffer (resuspension followed by centrifuged at 4000 g for 3 min). The pellets were resuspended in 100 µL 20 % (v/v) DMSO/PBS, 5 µM BCN-TAMRA added and reacted at 37 °C for 60 min (or samples taken at different time points: 15 min, 30 min, 45 min and 60 min). Samples were diluted with 1 mL 10 % (v/v) DMSO/PBS buffer and pelleted at 4000 g for 3 min and washed 2x with 10 % (v/v) DMSO/PBS buffer. The reaction was quenched for 30 min by resuspending the pellets with 10 µL dipyridyl-1,2,4,5-tetrazine (50 mM) diluted in 40 µL PBS. The pellets were then washed 3x with 10 % DMBO/PBS and the samples resuspended in 1x SDS loading buffer (100 µL for a pellet of 1 mL OD_{600} = 1), cooked for 10 min at 95 °C and centrifuged (17.000 x g, 10 min). 15 µL samples was separated via SDS-PAGE and labeling was visualized by excitation of TAMRA fluorescence at 520 nm utilizing an ImageQuant LAS 4000 with a 575DF20 Cy3 filter, followed by staining with Quick Coomassie Stain and destaining in deionized water.

II.2.3.14.2 Labeling of OmpC-232mTetK on live *E. coli* with photo-DMBO-Cy5

E. coli cultures expressing OmpC-Y232TAG without uAA, with BocK, BCNK or mTetK were harvested in 0.5-1 mL samples (pellet resuspended in 1 mL ≙ OD_{600} = 2), pelleted at 4000 g for 3 min and the pellet washed 3x with 1 mL 10 % (v/v) DMSO/PBS buffer (resuspension followed by centrifuged at 4000 g for 3 min). The pellets were resuspended in 100 µL 20 % (v/v) DMSO/PBS and 5 µM photo-DMBO-Cy5 or DMBO-Cy5 added. The reaction was induced at 365 nm with UV Lamp VL-215.L for 30 min on ice and then incubated at 37 °C and 350 rpm for 60 min. Samples were diluted with 1 mL 10 % (v/v) DMSO/PBS buffer and pelleted at 4000 g for 3 min and washed 2x with 10 % (v/v) DMSO/PBS buffer. The reaction was quenched for 30 min by resuspending the pellet in 10 µL dipyridyl-1,2,4,5-tetrazine (50 mM) diluted in 40 µL PBS, followed by washing the pellets 3x with 10 % DMBO/PBS.

For confocal microscopy 10 µL culture with an OD_{600} = 1 were mixed with 10 µL 1 % (w/v) agarose in MQ water and 10 µL mixture mounted between a coverslip and a microscopy slide.

For SDS-PAGE fluorescence imaging pellets were resuspended in 1x SDS (100 µL for a pellet of 1 mL OD_{600} = 1), cooked for 10 min at 95 °C and centrifuged for 10 min at 17.000 x g. 15 µL samples were separated via SDS-PAGE. Labeling was visualized by excitation of Cy5 fluorescence at 630 nm and utilizing an ImageQuant LAS 4000 with a R670 Cy5 filter, followed by staining with Quick Coomassie Stain and destaining in deionized water.

II.2.3.14.2 Sequential labeling of OmpC232mTetK on live *E. coli* with photo-DMBO-BCN, followed by Tet-Cy5

E. coli cultures expressing OmpC-Y232mTetK were harvested in 0.5-1 mL samples (pellet resuspended in 1 mL ≙ OD_{600} = 2), pelleted at 4000 g for 3 min and the pellet washed 3x with 1 mL 10 % (v/v) DMSO/PBS buffer (resuspension followed by centrifuged at 4000 g for 3 min). The pellets were resuspended in 100 µL 20 % (v/v) DMSO/PBS and 50 µM photo-DMBO-BCN or no compound added. The reaction was carried out at 37 °C and 350 rpm for 60 min. Samples were diluted with 1 mL 10 % (v/v) DMSO/PBS buffer and pelleted at 4000 g for 3 min and washed 2x with 10 % (v/v) DMSO/PBS buffer. The reaction was quenched for 30 min by resuspending the pellets in 10 µL dipyridyl-1,2,4,5-tetrazine (10 mM) diluted in 40 µL PBS. Pellets were then washed 3x with 10 % DMBO/PBS and resuspended in 100 µL 20 % (v/v) DMSO/PBS and 5 µM Tet-Cy5 added. The reaction was induced at 365 nm with UV Lamp VL-215.L for 60 min on ice and then incubated at 37 °C and 350 rpm for 60 min. Samples were diluted with 1 mL 10 % (v/v) DMSO/PBS buffer and pelleted at 4000 g for 3 min and washed 2x with 10 % (v/v) DMSO/PBS buffer. Reaction was quenched by resuspending the pellets in 5 µL BCN-alcohol (100 mM) diluted in 45 µL PBS for 30 min, then washed 3x with 10 % DMBO/PBS.

For confocal microscopy 10 µL culture with an OD_{600} = 1 were mixed with 10 µL 1 % (w/v) agarose in MQ water and 10 µL mixture mounted between a coverslip and a microscopy slide.

For SDS-PAGE fluorescence imaging pellets were resuspended in 1x SDS (100 µL for a pellet of 1 mL OD_{600} = 1), cooked for 10 min at 95 °C and centrifuged for 10 min at 17.000 x g. 15 µL samples separated via SDS-PAGE. Labeling was visualized by excitation of Cy5 fluorescence at 630 nm and utilizing an ImageQuant LAS 4000 with a R670 Cy5 filter, followed by staining with Quick Coomassie Stain and destaining in deionized water.

II.2.3.15 Kinetic measurements

II.2.3.15.1 Fluorimeter

Rate constants k between sfGFP150mTetK and BCN-alcohol or DMBO-NH_2 were measured under pseudo first order conditions with an excess of BCN-alcohol (10, 15, 20, 30, 50 or 75-fold) or DMBO-NH_2 (10, 15, 20, 30 or 50-fold) in PBS by following exponential increase of GFP fluorescence over time on a Horiba Jobin Yvon Fluoromax-4 in a Quartz cuvette with a stirrer at 25 °C. The final protein concentration of purified sfGF150mTetK was 10 µM in a final volume of 1 mL. GFP fluorescence was excited at 488 nm (slit 1 nm) and measured at 506 nm (slit 2 nm). Data points were taken every 0.25-1 s for 300-5000 s and measured as triplicates. The observed k' was plotted against the concentration of BCN-alcohol or DMBO-NH_2 to obtain the rate constant k from the slope of the plot. All data processing was performed using Kaleidagraph software (Synergy Software; Reading, UK).

II.2.3.15.2 Tecan Reader

Rate constants k between sfGFP150mTetK and Cp or Nor-OH, as well as sfGFP150TetF and DMBO-NH_2 or Cp were measured under pseudo first order conditions with an excess of cyclopropene (10, 25, 50, 75 or 100-fold), norbornenol (100, 250, 500, 750 or 1000-fold) or DMBO-NH_2 (10, 25, 40, 60 and 75-fold) in water or PBS by following exponential increase of GFP fluorescence over time on a Tecan Spark 10M Microplate reader. The final protein concentration of purified sfGF150mTetK or sfGFP150TetF was 10 µM in a final volume of 100 µL in a 96 well plate. The plate was shaken (amplitude 3.5) during measurements and GFP fluorescence excited at 480 nm and measured at 527 nm. Data points were taken as triplicates per well every 60 s for 6-8 h at 25 °C. Biological replicas were measured in triplicates and the mean of the observed k' was plotted against the concentration of cyclopropane, norbornene or DMBO-NH_2 to obtain the rate constant k from the slope of the plot. All data processing was performed using Kaleidagraph software (Synergy Software; Reading, UK), with the exception of the kinetic measurements between sfGFP150TetF and Cp, which were processed using Origin software.

III. REFERENCES

III. References

1. Baslé, E.; Joubert, N.; Pucheault, M., Protein Chemical Modification on Endogenous Amino Acids. *Chemistry & Biology* **2010,** *17* (3), 213-227.

2. Giepmans, B. N.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y., The fluorescent toolbox for assessing protein location and function. *Science* **2006,** *312* (5771), 217-24.

3. (a) Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A., Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiological Reviews* **2010,** *90* (3), 1103-1163; (b) Tsien, R. Y., THE GREEN FLUORESCENT PROTEIN. *Annual Review of Biochemistry* **1998,** *67* (1), 509-544.

4. Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J., Crystal Structure of the Aequorea victoria Green Fluorescent Protein. *Science* **1996,** *273* (5280), 1392.

5. Shaner, N. C., Rainbow palette of GFP derived fluorescent proteins. https://www.conncoll.edu/ccacad/zimmer/GFP-ww/tsien.html, last accessed July 11th, 2019, 2004.

6. (a) Remington, S. J., Fluorescent proteins: maturation, photochemistry and photophysics. *Current Opinion in Structural Biology* **2006,** *16* (6), 714-721; (b) Shaner, N. C.; Patterson, G. H.; Davidson, M. W., Advances in fluorescent protein technology. *Journal of Cell Science* **2007,** *120* (24), 4247.

7. (a) Chudakov, D. M.; Lukyanov, S.; Lukyanov, K. A., Fluorescent proteins as a toolkit for in vivo imaging. *Trends in Biotechnology* **2005,** *23* (12), 605-613; (b) Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y., A guide to choosing fluorescent proteins. *Nature Methods* **2005,** *2*, 905; (c) Verkhusha, V. V.; Lukyanov, K. A., The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nature Biotechnology* **2004,** *22*, 289.

8. (a) Fernández-Suárez, M.; Ting, A. Y., Fluorescent probes for super-resolution imaging in living cells. *Nature Reviews Molecular Cell Biology* **2008,** *9*, 929; (b) Lippincott-Schwartz, J.; Patterson, G. H., Photoactivatable fluorescent proteins for diffraction-limited and super-resolution imaging. *Trends in Cell Biology* **2009,** *19* (11), 555-565.

9. (a) Giraldez, T.; Hughes, T. E.; Sigworth, F. J., Generation of Functional Fluorescent BK Channels by Random Insertion of GFP Variants. *The Journal of General Physiology* **2005,** *126* (5), 429; (b) Choi, J. H.; Choi, J.-I.; Lee, S. Y., Display of Proteins on the Surface of Escherichia coli by C-Terminal Deletion Fusion to the Salmonella typhimurium OmpC. *Journal of Microbiology and Biotechnology* **2005,** *15* (1), 141-146.

10. Griffin, B. A.; Adams, S. R.; Tsien, R. Y., Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* **1998,** *281* (5374), 269.

11. (a) Andresen, M.; Schmitz-Salue, R.; Jakobs, S., Short Tetracysteine Tags to β-Tubulin Demonstrate the Significance of Small Labels for Live Cell Imaging. *Molecular Biology of the Cell* **2004,** *15* (12), 5616-5622; (b) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y., New Biarsenical Ligands

and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. *Journal of the American Chemical Society* **2002,** *124* (21), 6063-6076.

12. Gaietta, G.; Deerinck, T. J.; Adams, S. R.; Bouwer, J.; Tour, O.; Laird, D. W.; Sosinsky, G. E.; Tsien, R. Y.; Ellisman, M. H., Multicolor and Electron Microscopic Imaging of Connexin Trafficking. *Science* **2002,** *296* (5567), 503.

13. (a) Luedtke, N. W.; Dexter, R. J.; Fried, D. B.; Schepartz, A., Surveying polypeptide and protein domain conformation and association with FlAsH and ReAsH. *Nature Chemical Biology* **2007,** *3*, 779; (b) Scheck, R. A.; Lowder, M. A.; Appelbaum, J. S.; Schepartz, A., Bipartite Tetracysteine Display Reveals Allosteric Control of Ligand-Specific EGFR Activation. *ACS Chemical Biology* **2012,** *7* (8), 1367-1376.

14. Stroffekova, K.; Proenza, C.; Beam, K. G., The protein-labeling reagent FLASH-EDT2 binds not only to CCXXCC motifs but also non-specifically to endogenous cysteine-rich proteins. *Pflügers Archiv* **2001,** *442* (6), 859-866.

15. Kim, K. K.; Escobedo, J. O.; St. Luce, N. N.; Rusin, O.; Wong, D.; Strongin, R. M., Postcolumn HPLC Detection of Mono- and Oligosaccharides with a Chemosensor. *Organic Letters* **2003,** *5* (26), 5007-5010.

16. Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A., Selective Recognition of Protein Tetraserine Motifs with a Cell-Permeable, Pro-fluorescent Bis-boronic Acid. *Journal of the American Chemical Society* **2009,** *131* (2), 438-439.

17. (a) Juillerat, A.; Gronemeyer, T.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K., Directed Evolution of O6-Alkylguanine-DNA Alkyltransferase for Efficient Labeling of Fusion Proteins with Small Molecules In Vivo. *Chemistry & Biology* **2003,** *10* (4), 313-317; (b) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K., A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature Biotechnology* **2002,** *21*, 86.

18. (a) Jones, S. A.; Shim, S.-H.; He, J.; Zhuang, X., Fast, three-dimensional super-resolution imaging of live cells. *Nature Methods* **2011,** *8*, 499; (b) Klein, T.; Löschberger, A.; Proppert, S.; Wolter, S.; van de Linde, S.; Sauer, M., Live-cell dSTORM with SNAP-tag fusion proteins. *Nature Methods* **2010,** *8*, 7; (c) Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R., Jr.; Kindermann, M.; Beaufils, F.; Johnsson, K., An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chemistry & Biology* **2008,** *15* (2), 128-136.

19. Sun, X.; Zhang, A.; Baker, B.; Sun, L.; Howard, A.; Buswell, J.; Maurel, D.; Masharina, A.; Johnsson, K.; Noren, C. J.; Xu, M.-Q.; Corrêa Jr, I. R., Development of SNAP-Tag Fluorogenic Probes for Wash-Free Fluorescence Imaging. *ChemBioChem* **2011,** *12* (14), 2217-2226.

20. Reck-Peterson, S. L.; Yildiz, A.; Carter, A. P.; Gennerich, A.; Zhang, N.; Vale, R. D., Single-Molecule Analysis of Dynein Processivity and Stepping Behavior. *Cell* **2006,** *126* (2), 335-348.

21. Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V., HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chemical Biology* **2008,** *3* (6), 373-382.

22. Miller, L. W.; Cai, Y.; Sheetz, M. P.; Cornish, V. W., In vivo protein labeling with trimethoprim conjugates: a flexible chemical tag. *Nature Methods* **2005,** *2*, 255.

23. Crivat, G.; Taraska, J. W., Imaging proteins inside cells with fluorescent tags. *Trends in Biotechnology* **2012,** *30* (1), 8-16.

24. Hinner, M. J.; Johnsson, K., How to obtain labeled proteins and what to do with them. *Current Opinion in Biotechnology* **2010,** *21* (6), 766-776.

25. (a) George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K., Specific Labeling of Cell Surface Proteins with Chemically Diverse Compounds. *Journal of the American Chemical Society* **2004,** *126* (29), 8896-8897; (b) Yin, J.; Liu, F.; Li, X.; Walsh, C. T., Labeling Proteins with Small Molecules by Site-Specific Posttranslational Modification. *Journal of the American Chemical Society* **2004,** *126* (25), 7754-7755.

26. Zhou, Z.; Koglin, A.; Wang, Y.; McMahon, A. P.; Walsh, C. T., An Eight Residue Fragment of an Acyl Carrier Protein Suffices for Post-Translational Introduction of Fluorescent Pantetheinyl Arms in Protein Modification in vitro and in vivo. *Journal of the American Chemical Society* **2008,** *130* (30), 9925-9930.

27. Green, N. M., Avidin. In *Advances in Protein Chemistry*, Anfinsen, C. B.; Edsall, J. T.; Richards, F. M., Eds. Academic Press: 1975; Vol. 29, pp 85-133.

28. (a) Beckett, D.; Kovaleva, E.; Schatz, P. J., A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Science* **1999,** *8* (4), 921-929; (b) Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y., Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nature Methods* **2005,** *2*, 99.

29. (a) Chandler, C. S.; Ballard, F. J., Multiple biotin-containing proteins in 3T3-L1 cells. *Biochemical Journal* **1986,** *237* (1), 123; (b) Kirkeby, S.; Moe, D.; Bog-Hansen, T. C.; van Noorden, C. J., Biotin carboxylases in mitochondria and the cytosol from skeletal and cardiac muscle as detected by avidin binding. *Histochemistry* **1993,** *100* (6), 415-21.

30. Yao, J. Z.; Uttamapinant, C.; Poloukhtine, A.; Baskin, J. M.; Codelli, J. A.; Sletten, E. M.; Bertozzi, C. R.; Popik, V. V.; Ting, A. Y., Fluorophore Targeting to Cellular Proteins via Enzyme-Mediated Azide Ligation and Strain-Promoted Cycloaddition. *Journal of the American Chemical Society* **2012,** *134* (8), 3720-3728.

31. Uttamapinant, C.; White, K. A.; Baruah, H.; Thompson, S.; Fernández-Suárez, M.; Puthenveetil, S.; Ting, A. Y., A fluorophore ligase for site-specific protein labeling inside living cells. *Proceedings of the National Academy of Sciences* **2010,** *107* (24), 10914.

32. Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y., Diels–Alder Cycloaddition for Fluorophore Targeting to Specific Proteins inside Living Cells. *Journal of the American Chemical Society* **2012,** *134* (2), 792-795.

33. Baalmann, M.; Ziegler, M. J.; Werther, P.; Wilhelm, J.; Wombacher, R., Enzymatic and Site-Specific Ligation of Minimal Size Tetrazines and Triazines to Proteins for Bioconjugation and Live-Cell Imaging. *Bioconjugate Chemistry* **2019**.

34. Best, M.; Degen, A.; Baalmann, M.; Schmidt, T. T.; Wombacher, R., Two-Step Protein Labeling by Using Lipoic Acid Ligase with Norbornene Substrates and Subsequent Inverse-Electron Demand Diels–Alder Reaction. *ChemBioChem* **2015,** *16* (8), 1158-1162.

35. Puthenveetil, S.; Liu, D. S.; White, K. A.; Thompson, S.; Ting, A. Y., Yeast Display Evolution of a Kinetically Efficient 13-Amino Acid Substrate for Lipoic Acid Ligase. *Journal of the American Chemical Society* **2009,** *131* (45), 16430-16438.

36. Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y., Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. *Nature Biotechnology* **2007,** *25*, 1483.

37. Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O., Staphylococcus aureus Sortase, an Enzyme that Anchors Surface Proteins to the Cell Wall. *Science* **1999,** *285* (5428), 760.

38. Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L., Sortagging: a versatile method for protein labeling. *Nature Chemical Biology* **2007,** *3*, 707.

39. (a) Lang, K.; Chin, J. W., Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chemical Reviews* **2014,** *114* (9), 4764-4806; (b) Lang, K.; Chin, J. W., Bioorthogonal Reactions for Labeling Proteins. *ACS Chemical Biology* **2014,** *9* (1), 16-20.

40. Berg, J. M.; Tymoczko, J. L.; Stryer, L., Synthese und Prozessierung von RNA. In *Stryer Biochemie*, 7 ed.; Springer Spektrum Akademischer Verlag: Heidelberg, 2013; pp 862-871.

41. (a) Khorana, H. G., Nucleic acid synthesis in the study of the genetic code. In *Nobel lectures, Physiology or Medicine 1963-1970*, Elsevier Publishing Company: Amsterdam, 1972; (b) Holley, R. W., Alanine transfer RNA. In *Nobel lectures, Physiology or Medicine 1963-1970*, Elsevier Publishing Company: Amsterdam, 1972; (c) Nirenberg, M. W., The genetic code. In *Nobel lectures, Physiology or Medicine 1963-1970*, Elsevier Publishing Company: Amsterdam, 1972.

42. Berg, J. M.; Tymoczko, J. L.; Stryer, L., Proteinsynthese. In *Stryer Biochemie*, 7 ed.; Springer Spektrum Akademischer Verlag: Heidelberg, 2013; pp 899-929.

43. Lodish, H.; Berk, A.; Matsudaira, P.; Kaiser, C. A.; Krieger, M.; Scott, M. P.; Zipursky, L.; Darnell, J., Basic Molecular Genetic Mechanisms In *Molecular Cell Biology*, 5the edition ed.; W.H.Freeman & Co Ltd: 2003; pp 101-145.

44. (a) Srinivasan, G.; James, C. M.; Krzycki, J. A., Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **2002,** *296* (5572), 1459-62; (b) Bo¨ck, A.; Forchhammer, K.; Heider, J.; Baron, C., Selenoprotein synthesis: an expansion of the genetic code. *Trends in Biochemical Sciences* **1991,** *16*, 463-467.

45. Davis, L.; Chin, J. W., Designer proteins: applications of genetic code expansion in cell biology. *Nature reviews. Molecular cell biology* **2012,** *13* (3), 168-82.

46. (a) Dawson, P. E.; Kent, S. B., Synthesis of native proteins by chemical ligation. *Annu Rev Biochem* **2000,** *69*, 923-60; (b) Kimmerlin, T.; Seebach, D., '100 years of peptide synthesis': ligation methods for peptide and protein synthesis with applications to beta-peptide assemblies. *The journal of peptide research : official journal of the American Peptide Society* **2005,** *65* (2), 229-60.

47. (a) van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A., Efficient Incorporation of Unsaturated Methionine Analogues into Proteins in Vivo. *Journal of the American Chemical Society* **2000,** *122* (7), 1282-1288; (b) van Hest, J. C. M.; Tirrell, D. A., Efficient introduction of alkene functionality into proteins in vivo. *FEBS Letters* **1998,** *428* (1-2), 68-70; (c) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R., Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proceedings of the National Academy of Sciences* **2002,** *99* (1), 19; (d) Cowie, D. B.; Cohen, G. N., Biosynthesis by Escherichia coli of active altered proteins containing selenium instead of sulfur. *Biochimica et biophysica acta* **1957,** *26* (2), 252-61.

48. (a) Tang, Y.; Tirrell, D. A., Attenuation of the editing activity of the Escherichia coli leucyl-tRNA synthetase allows incorporation of novel amino acids into proteins in vivo. *Biochemistry* **2002,** *41* (34), 10635-45; (b) Kast, P.; Hennecke, H., Amino acid substrate specificity of Escherichia coli phenylalanyl-tRNA synthetase altered by distinct mutations. *J Mol Biol* **1991,** *222* (1), 99-124; (c) Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A., Biosynthesis of Proteins Incorporating a Versatile Set of Phenylalanine Analogues. *ChemBioChem* **2002,** *3* (2-3), 235-237.

49. Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G., A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **1989,** *244* (4901), 182.

50. Heckler, T. G.; Chang, L. H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M., T4 RNA ligase mediated preparation of novel "chemically misacylated" tRNAPhes. *Biochemistry* **1984,** *23* (7), 1468-1473.

51. Xie, J.; Schultz, P. G., An expanding genetic code. *Methods (San Diego, Calif.)* **2005,** *36* (3), 227-38.

52. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G., Expanding the genetic code of Escherichia coli. *Science* **2001,** *292* (5516), 498-500.

53. Krzycki, J. A., The direct genetic encoding of pyrrolysine. *Current Opinion in Microbiology* **2005,** *8* (6), 706-712.

54. (a) Willis, J. C. W.; Chin, J. W., Mutually orthogonal pyrrolysyl-tRNA synthetase/tRNA pairs. *Nature Chemistry* **2018,** *10* (8), 831-837; (b) Beránek, V.; Willis, J. C. W.; Chin, J. W., An Evolved Methanomethylophilus alvus Pyrrolysyl-tRNA Synthetase/tRNA Pair Is Highly Active and Orthogonal in Mammalian Cells. *Biochemistry* **2019,** *58* (5), 387-390.

55. Dower, W. J.; Miller, J. F.; Ragsdale, C. W., High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res* **1988,** *16* (13), 6127-45.

56. (a) Liu, D. R.; Schultz, P. G., Progress toward the evolution of an organism with an expanded genetic code. *Proceedings of the National Academy of Sciences* **1999,** *96* (9), 4780; (b) Seitchik, J. L.; Peeler, J. C.; Taylor, M. T.; Blackman, M. L.; Rhoads, T. W.; Cooley, R. B.; Refakis, C.; Fox, J. M.; Mehl, R. A., Genetically encoded tetrazine amino acid directs rapid site-specific in vivo bioorthogonal ligation with trans-cyclooctenes. *J Am Chem Soc* **2012,** *134* (6), 2898-901.

57. Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G., Designing logical codon reassignment – Expanding the chemistry in biology. *Chemical Science* **2015,** *6* (1), 50-69.

58. (a) Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W., Genetic Encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. *J Am Chem Soc* **2012,** *134* (25), 10317-20; (b) Uttamapinant, C.; Howe, J. D.; Lang, K.; Beranek, V.; Davis, L.; Mahesh, M.; Barry, N. P.; Chin, J. W., Genetic code expansion enables live-cell and super-resolution imaging of site-specifically labeled cellular proteins. *J Am Chem Soc* **2015,** *137* (14), 4602-5.

59. (a) Evans, E. G. B.; Millhauser, G. L., Genetic Incorporation of the Unnatural Amino Acid p-Acetyl Phenylalanine into Proteins for Site-Directed Spin Labeling. *Methods in enzymology* **2015,** *563*, 503-527; (b) Furter, R., Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. *Protein science : a publication of the Protein Society* **1998,** *7* (2), 419-426.

60. (a) Fleissner, M. R.; Brustad, E. M.; Kálai, T.; Altenbach, C.; Cascio, D.; Peters, F. B.; Hideg, K.; Peuker, S.; Schultz, P. G.; Hubbell, W. L., Site-directed spin labeling of a genetically encoded unnatural amino acid. *Proceedings of the National Academy of Sciences* **2009,** *106* (51), 21637-21642; (b) Schmidt, M. J.; Borbas, J.; Drescher, M.; Summerer, D., A Genetically Encoded Spin Label for Electron Paramagnetic Resonance Distance Measurements. *Journal of the American Chemical Society* **2014,** *136* (4), 1238-1241.

61. (a) Grunbeck, A.; Huber, T.; Sachdev, P.; Sakmar, T. P., Mapping the ligand-binding site on a G protein-coupled receptor (GPCR) using genetically encoded photocrosslinkers. *Biochemistry* **2011,** *50* (17), 3411-3413; (b) Grunbeck, A.; Sakmar, T. P., Probing G protein-coupled receptor-ligand interactions with targeted photoactivatable cross-linkers. *Biochemistry* **2013,** *52* (48), 8625-32.

62. Nguyen, T.-A.; Cigler, M.; Lang, K., Expanding the Genetic Code to Study Protein–Protein Interactions. *Angewandte Chemie International Edition* **2018,** *57* (44), 14350-14361.

63. Cigler, M.; Müller, T. G.; Horn-Ghetko, D.; von Wrisberg, M.-K.; Fottner, M.; Goody, R. S.; Itzen, A.; Müller, M. P.; Lang, K., Proximity-Triggered Covalent Stabilization of Low-Affinity Protein Complexes In Vitro and In Vivo. *Angewandte Chemie International Edition* **2017,** *56* (49), 15737-15741.

64. Li, J.; Chen, P. R., Development and application of bond cleavage reactions in bioorthogonal chemistry. *Nat Chem Biol* **2016,** *12* (3), 129-37.

65. (a) Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G., A facile system for encoding unnatural amino acids in mammalian cells. *Angewandte Chemie (International ed. in English)* **2009,** *48* (22), 4052-4055; (b) Deiters, A.; Groff, D.; Ryu, Y.; Xie, J.; Schultz, P. G., A Genetically Encoded Photocaged Tyrosine. *Angewandte Chemie International Edition* **2006,** *45* (17), 2728-2731; (c) Lemke, E. A.; Summerer, D.; Geierstanger, B. H.; Brittain, S. M.; Schultz, P. G., Control of protein phosphorylation with a genetically encoded photocaged amino acid. *Nature Chemical Biology* **2007,** *3*, 769; (d) Li, J.; Jia, S.; Chen, P. R., Diels-Alder reaction-triggered bioorthogonal protein decaging in living cells. *Nat Chem Biol* **2014,** *10* (12), 1003-5.

66. Reille-Seroussi, M.; Mayer, S. V.; Dörner, W.; Lang, K.; Mootz, H. D., Expanding the genetic code with a lysine derivative bearing an enzymatically removable phenylacetyl group. *Chemical Communications* **2019**.

67. Krishna, R. G.; Wold, F., Post-Translational Modifications of Proteins. In *Methods in Protein Sequence Analysis*, Imahori, K.; Sakiyama, F., Eds. Springer US: Boston, MA, 1993; pp 167-172.

68. (a) Rogerson, D. T.; Sachdeva, A.; Wang, K.; Haq, T.; Kazlauskaite, A.; Hancock, S. M.; Huguenin-Dezot, N.; Muqit, M. M. K.; Fry, A. M.; Bayliss, R.; Chin, J. W., Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nature chemical biology* **2015,** *11* (7), 496-503; (b) Zhang, M. S.; Brunner, S. F.; Huguenin-Dezot, N.; Liang, A. D.; Schmied, W. H.; Rogerson, D. T.; Chin, J. W., Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nature methods* **2017,** *14* (7), 729-736.

69. (a) Hoppmann, C.; Wong, A.; Yang, B.; Li, S.; Hunter, T.; Shokat, K. M.; Wang, L., Site-specific incorporation of phosphotyrosine using an expanded genetic code. *Nat Chem Biol* **2017,** *13* (8), 842-844; (b) Luo, X.; Fu, G.; Wang, R. E.; Zhu, X.; Zambaldo, C.; Liu, R.; Liu, T.; Lyu, X.; Du, J.; Xuan, W.; Yao, A.; Reed, S. A.; Kang, M.; Zhang, Y.; Guo, H.; Huang, C.; Yang, P.-Y.; Wilson, I. A.; Schultz, P. G.; Wang, F., Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nature chemical biology* **2017,** *13* (8), 845-849.

70. (a) Elsässer, S. J.; Ernst, R. J.; Walker, O. S.; Chin, J. W., Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nature Methods* **2016,** *13*, 158; (b) Kim, C. H.; Kang, M.; Kim, H. J.; Chatterjee, A.; Schultz, P. G., Site-specific incorporation of ε-N-crotonyllysine into histones. *Angewandte Chemie (International ed. in English)* **2012,** *51* (29), 7246-7249; (c) Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S., Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochemical and Biophysical Research Communications* **2008,** *371* (4), 818-822; (d) Wang, Y.-S.; Wu, B.; Wang, Z.; Huang, Y.; Wan, W.; Russell, W. K.; Pai, P.-J.; Moe, Y. N.; Russell, D. H.; Liu, W. R., A genetically encoded photocaged Nε-methyl-l-lysine. *Molecular BioSystems* **2010,** *6* (9), 1557-1560.

71. (a) Glickman, M. H.; Ciechanover, A., The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiological Reviews* **2002,** *82* (2), 373-428; (b) Mukhopadhyay, D.; Riezman, H., Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling. *Science* **2007,** *315* (5809), 201; (c) Pickart, C. M.; Eddins, M. J., Ubiquitin: structures, functions, mechanisms. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2004,** *1695* (1), 55-72.

72. (a) Stanley, M.; Virdee, S., Genetically Directed Production of Recombinant, Isosteric and Nonhydrolysable Ubiquitin Conjugates. *ChemBioChem* **2016,** *17* (15), 1472-1480; (b) Stanley, M.; Virdee, S., Genetically Directed Production of Recombinant, Isosteric and Nonhydrolysable Ubiquitin Conjugates. *Chembiochem* **2016,** *17* (15), 1472-80; (c) Virdee, S.; Kapadnis, P. B.; Elliott, T.; Lang, K.; Madrzak, J.; Nguyen, D. P.; Riechmann, L.; Chin, J. W., Traceless and Site-Specific Ubiquitination of Recombinant Proteins. *Journal of the American Chemical Society* **2011,** *133* (28), 10708-10711; (d) Virdee, S.; Ye, Y.; Nguyen, D. P.; Komander, D.; Chin, J. W., Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol* **2010,** *6* (10), 750-7; (e) Weikart, N. D.; Mootz, H. D., Generation of Site-Specific and Enzymatically Stable Conjugates of Recombinant Proteins with Ubiquitin-Like Modifiers by the CuI-Catalyzed Azide–Alkyne Cycloaddition. *ChemBioChem* **2010,** *11* (6), 774-777.

73. Fottner, M.; Brunner, A.-D.; Bittl, V.; Horn-Ghetko, D.; Jussupow, A.; Kaila, V. R. I.; Bremm, A.; Lang, K., Site-specific ubiquitylation and SUMOylation using genetic-code expansion and sortase. *Nature Chemical Biology* **2019,** *15* (3), 276-284.

74. Wang, K.; Neumann, H.; Peak-Chew, S. Y.; Chin, J. W., Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nature Biotechnology* **2007,** *25*, 770.

75. (a) Johnson, D. B.; Wang, C.; Xu, J.; Schultz, M. D.; Schmitz, R. J.; Ecker, J. R.; Wang, L., Release factor one is nonessential in Escherichia coli. *ACS Chem Biol* **2012,** *7* (8), 1337-44; (b) Johnson, D. B.; Xu, J.; Shen, Z.; Takimoto, J. K.; Schultz, M. D.; Schmitz, R. J.; Xiang, Z.; Ecker, J. R.; Briggs, S. P.; Wang, L., RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat Chem Biol* **2011,** *7* (11), 779-86; (c) Mukai, T.; Hayashi, A.; Iraha, F.; Sato, A.; Ohtake, K.; Yokoyama, S.; Sakamoto, K., Codon reassignment in the Escherichia coli genetic code. *Nucleic Acids Res* **2010,** *38* (22), 8188-95; (d) Ohtake, K.; Sato, A.; Mukai, T.; Hino, N.; Yokoyama, S.; Sakamoto, K., Efficient Decoding of the UAG Triplet as a Full-Fledged Sense Codon Enhances the Growth of a prfA-Deficient Strain of <span class="named-content genus-species" id="named-content-1">Escherichia coli</span&gt. *Journal of Bacteriology* **2012,** *194* (10), 2606.

76. Rackham, O.; Chin, J. W., A network of orthogonal ribosome·mRNA pairs. *Nature Chemical Biology* **2005,** *1* (3), 159-166.

77. Peister, A.; Goel, R. A.; Triman, K. L., Expanded Versions of the 16S and 23S Ribosomal RNA Mutation Databases (16SMDBexp and 23SMDBexp). *Nucleic Acids Research* **1998,** *26* (1), 280-284.

78. (a) Fredens, J.; Wang, K.; de la Torre, D.; Funke, L. F. H.; Robertson, W. E.; Christova, Y.; Chia, T.; Schmied, W. H.; Dunkelmann, D. L.; Beránek, V.; Uttamapinant, C.; Llamazares, A. G.; Elliott, T. S.; Chin, J. W., Total synthesis of Escherichia coli with a recoded genome. *Nature* **2019,** *569* (7757), 514-518; (b) Napolitano, M. G.; Landon, M.; Gregg, C. J.; Lajoie, M. J.; Govindarajan, L.; Mosberg, J. A.; Kuznetsov, G.; Goodman, D. B.; Vargas-Rodriguez, O.; Isaacs, F. J.; Söll, D.; Church, G. M., Emergent rules for codon choice elucidated by editing rare arginine codons in <em>Escherichia coli</em&gt. *Proceedings of the National Academy of Sciences* **2016,** *113* (38), E5588.

79. (a) Wan, W.; Huang, Y.; Wang, Z.; Russell, W. K.; Pai, P.-J.; Russell, D. H.; Liu, W. R., A Facile System for Genetic Incorporation of Two Different Noncanonical Amino Acids into One Protein in Escherichia coli. *Angewandte Chemie International Edition* **2010,** *49* (18), 3211-3214; (b) Wu, B.; Wang, Z.; Huang, Y.; Liu, W. R., Catalyst-Free and Site-Specific One-Pot Dual-Labeling of a Protein Directed by Two Genetically Incorporated Noncanonical Amino Acids. *ChemBioChem* **2012,** *13* (10), 1405-1408; (c) Xiao, H.; Chatterjee, A.; Choi, S.-h.; Bajjuri, K. M.; Sinha, S. C.; Schultz, P. G., Genetic Incorporation of Multiple Unnatural Amino Acids into Proteins in Mammalian Cells. *Angewandte Chemie International Edition* **2013,** *52* (52), 14080-14083.

80. Anderson, J. C.; Wu, N.; Santoro, S. W.; Lakshman, V.; King, D. S.; Schultz, P. G., An expanded genetic code with a functional quadruplet codon. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (20), 7566.

81. (a) Hohsaka, T.; Sisido, M., Incorporation of non-natural amino acids into proteins. *Current Opinion in Chemical Biology* **2002,** *6* (6), 809-815; (b) Murakami, H.; Hohsaka, T.; Ashizuka, Y.; Sisido, M., Site-Directed Incorporation of p-Nitrophenylalanine into Streptavidin and Site-to-Site Photoinduced Electron Transfer from a Pyrenyl Group to a Nitrophenyl Group on the Protein Framework. *Journal of the American Chemical Society* **1998,** *120* (30), 7520-7529; (c) Ohtsuki, T.; Manabe, T.; Sisido, M., Multiple incorporation of non-natural amino acids into a single protein using tRNAs with non-standard structures. *FEBS Letters* **2005,** *579* (30), 6769-6774.

82. (a) Monahan, S. L.; Lester, H. A.; Dougherty, D. A., Site-Specific Incorporation of Unnatural Amino Acids into Receptors Expressed in Mammalian Cells. *Chemistry & Biology* **2003,** *10* (6), 573-580; (b) Rodriguez, E. A.; Lester, H. A.; Dougherty, D. A., In vivo incorporation of multiple unnatural amino acids through nonsense and frameshift suppression. *Proceedings of the National Academy of Sciences* **2006,** *103* (23), 8650.

83. Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W., Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **2010,** *464*, 441.

84. Sachdeva, A.; Wang, K.; Elliott, T.; Chin, J. W., Concerted, Rapid, Quantitative, and Site-Specific Dual Labeling of Proteins. *Journal of the American Chemical Society* **2014,** *136* (22), 7785-7788.

85. (a) Fried, S. D.; Schmied, W. H.; Uttamapinant, C.; Chin, J. W., Ribosome Subunit Stapling for Orthogonal Translation in E. coli. *Angewandte Chemie International Edition* **2015,** *54* (43), 12791-12794; (b) Orelle, C.; Carlson, E. D.; Szal, T.; Florin, T.; Jewett, M. C.; Mankin, A. S., Protein synthesis by ribosomes with tethered subunits. *Nature* **2015,** *524*, 119.

86. Schmied, W. H.; Tnimov, Z.; Uttamapinant, C.; Rae, C. D.; Fried, S. D.; Chin, J. W., Controlling orthogonal ribosome subunit interactions enables evolution of new function. *Nature* **2018,** *564* (7736), 444-448.

87. (a) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M., Construction of modified ribosomes for incorporation of D-amino acids into proteins. *Biochemistry* **2006,** *45* (51), 15541-51; (b) Dedkova, L. M.; Hecht, S. M., Expanding the Scope of Protein Synthesis Using Modified Ribosomes. *Journal of the American Chemical Society* **2019**; (c) Maini, R.; Chowdhury, S. R.; Dedkova, L. M.; Roy, B.; Daskalova, S. M.; Paul, R.; Chen, S.; Hecht, S. M., Protein Synthesis with Ribosomes Selected for the Incorporation of beta-Amino Acids. *Biochemistry* **2015,** *54* (23), 3694-706; (d) Maini, R.; Nguyen, D. T.; Chen, S.; Dedkova, L. M.; Chowdhury, S. R.; Alcala-Torano, R.; Hecht, S. M., Incorporation of beta-amino acids into dihydrofolate reductase by ribosomes having modifications in the peptidyltransferase center. *Bioorganic & medicinal chemistry* **2013,** *21* (5), 1088-96.

88. Chen, S.; Ji, X.; Gao, M.; Dedkova, L. M.; Hecht, S. M., In Cellulo Synthesis of Proteins Containing a Fluorescent Oxazole Amino Acid. *Journal of the American Chemical Society* **2019,** *141* (14), 5597-5601.

89. McKay, Craig S.; Finn, M. G., Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. *Chemistry & Biology* **2014,** *21* (9), 1075-1101.

90. (a) Carell, T.; Vrabel, M., Bioorthogonal Chemistry—Introduction and Overview. *Topics in Current Chemistry* **2016,** *374* (1), 9; (b) Shih, H. W.; Kamber, D. N.; Prescher, J. A., Building better bioorthogonal reactions. *Curr Opin Chem Biol* **2014,** *21*, 103-11.

91. (a) Jencks, W. P., Studies on the Mechanism of Oxime and Semicarbazone Formation1. *Journal of the American Chemical Society* **1959,** *81* (2), 475-481; (b) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R., Engineering Chemical Reactivity on Cell Surfaces Through Oligosaccharide Biosynthesis. *Science* **1997,** *276* (5315), 1125; (c) Rideout, D., Self-assembling cytotoxins. *Science* **1986,** *233* (4763), 561; (d) Sander, E. G.; Jencks, W. P., Equilibria for additions to the carbonyl group. *Journal of the American Chemical Society* **1968,** *90* (22), 6154-6162.

92. Nauman, D. A.; Bertozzi, C. R., Kinetic parameters for small-molecule drug delivery by covalent cell surface targeting. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2001,** *1568* (2), 147-154.

93. (a) Luchansky, S. J.; Goon, S.; Bertozzi, C. R., Expanding the Diversity of Unnatural Cell-Surface Sialic Acids. *ChemBioChem* **2004,** *5* (3), 371-374; (b) Sadamoto, R.; Niikura, K.; Ueda, T.; Monde, K.; Fukuhara, N.; Nishimura, S.-I., Control of Bacteria Adhesion by Cell-Wall Engineering. *Journal of the American Chemical Society* **2004,** *126* (12), 3755-3761.

94. (a) Lim, R. K.; Lin, Q., Bioorthogonal chemistry: recent progress and future directions. *Chem Commun (Camb)* **2010,** *46* (10), 1589-600; (b) Prescher, J. A.; Bertozzi, C. R., Chemistry in living systems. *Nat Chem Biol* **2005,** *1* (1), 13-21.

95. (a) Dirksen, A.; Dawson, P. E., Rapid Oxime and Hydrazone Ligations with Aromatic Aldehydes for Biomolecular Labeling. *Bioconjugate Chemistry* **2008,** *19* (12), 2543-2548; (b) Dirksen, A.; Dirksen, S.; Hackeng, T. M.; Dawson, P. E., Nucleophilic Catalysis of Hydrazone Formation and Transimination: Implications for Dynamic Covalent Chemistry. *Journal of the American Chemical Society* **2006,** *128* (49), 15602-15603; (c) Dirksen, A.; Hackeng, T. M.; Dawson, P. E., Nucleophilic catalysis of oxime ligation. *Angew Chem Int Ed Engl* **2006,** *45* (45), 7581-4; (d) Dirksen, A.; Hackeng, T. M.; Dawson, P. E., Nucleophilic Catalysis of Oxime Ligation. *Angewandte Chemie International Edition* **2006,** *45* (45), 7581-7584.

96. (a) Rayo, J.; Amara, N.; Krief, P.; Meijler, M. M., Live Cell Labeling of Native Intracellular Bacterial Receptors Using Aniline-Catalyzed Oxime Ligation. *Journal of the American Chemical Society* **2011,** *133* (19), 7469-7475; (b) Zeng, Y.; Ramya, T. N. C.; Dirksen, A.; Dawson, P. E.; Paulson, J. C., High-efficiency labeling of sialylated glycoproteins on living cells. *Nature Methods* **2009,** *6*, 207.

97. Griffin, R. J., The medicinal chemistry of the azido group. *Progress in medicinal chemistry* **1994,** *31*, 121-232.

98. Luo, J.; Liu, Q.; Morihiro, K.; Deiters, A., Small-molecule control of protein function through Staudinger reduction. *Nature Chemistry* **2016,** *8*, 1027.

99. Pawlak, J. B.; Gential, G. P. P.; Ruckwardt, T. J.; Bremmers, J. S.; Meeuwenoord, N. J.; Ossendorp, F. A.; Overkleeft, H. S.; Filippov, D. V.; van Kasteren, S. I., Bioorthogonal Deprotection on the Dendritic Cell Surface for Chemical Control of Antigen Cross-Presentation. *Angewandte Chemie International Edition* **2015,** *54* (19), 5628-5631.

100. Fottner, M.; Brunner, A. D.; Bittl, V.; Horn-Ghetko, D.; Jussupow, A.; Kaila, V. R. I.; Bremm, A.; Lang, K., Site-specific ubiquitylation and SUMOylation using genetic-code expansion and sortase. *Nat Chem Biol* **2019,** *15* (3), 276-284.

101. Sletten, E. M.; Bertozzi, C. R., Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angewandte Chemie International Edition* **2009,** *48* (38), 6974-6998.

102. Saxon, E.; Armstrong, J. I.; Bertozzi, C. R., A "Traceless" Staudinger Ligation for the Chemoselective Synthesis of Amide Bonds. *Organic Letters* **2000,** *2* (14), 2141-2143.

103. Nilsson, B. L.; Kiessling, L. L.; Raines, R. T., Staudinger Ligation: A Peptide from a Thioester and Azide. *Organic Letters* **2000,** *2* (13), 1939-1941.

104. (a) Charron, G.; Zhang, M. M.; Yount, J. S.; Wilson, J.; Raghavan, A. S.; Shamir, E.; Hang, H. C., Robust Fluorescent Detection of Protein Fatty-Acylation with Chemical Reporters. *Journal of the American Chemical Society* **2009,** *131* (13), 4967-4975; (b) Hang, H. C.; Yu, C.; Kato, D. L.; Bertozzi, C. R., A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. *Proc Natl Acad Sci U S A* **2003,** *100* (25), 14846-51; (c) Hangauer, M. J.; Bertozzi, C. R., A FRET-based fluorogenic phosphine for live-cell imaging with the Staudinger ligation. *Angew Chem Int Ed Engl* **2008,** *47* (13), 2394-7; (d) Luchansky, S. J.; Argade, S.; Hayes, B. K.; Bertozzi, C. R., Metabolic Functionalization of Recombinant Glycoproteins. *Biochemistry* **2004,** *43* (38), 12358-12366; (e) Prescher, J. A.; Dube, D. H.; Bertozzi, C. R., Chemical remodelling of cell surfaces in living animals. *Nature* **2004,** *430* (7002), 873-7; (f) Saxon, E.; Bertozzi, C. R., Cell surface engineering by a modified Staudinger reaction. *Science* **2000,** *287* (5460), 2007-10; (g) Vocadlo, D. J.; Hang, H. C.; Kim, E. J.; Hanover, J. A.; Bertozzi, C. R., A chemical approach for identifying O-GlcNAc-modified proteins in cells. *Proc Natl Acad Sci U S A* **2003,** *100* (16), 9116-21; (h) Watzke, A.; Kohn, M.; Gutierrez-Rodriguez, M.; Wacker, R.; Schroder, H.; Breinbauer, R.; Kuhlmann, J.; Alexandrov, K.; Niemeyer, C. M.; Goody, R. S.; Waldmann, H., Site-selective protein immobilization by Staudinger ligation. *Angew Chem Int Ed Engl* **2006,** *45* (9), 1408-12.

105. Huisgen, R., 1,3-Dipolar Cycloadditions. Past and Future. *Angewandte Chemie International Edition in English* **1963,** *2* (10), 565-598.

106. (a) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angewandte Chemie International Edition* **2002,** *41* (14), 2596-2599; (b) Tornøe, C. W.; Christensen, C.; Meldal, M., Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *The Journal of Organic Chemistry* **2002,** *67* (9), 3057-3064.

107. Kolb, H. C.; Sharpless, K. B., The growing impact of click chemistry on drug discovery. *Drug discovery today* **2003,** *8* (24), 1128-37.

108. Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V., Copper(I)-Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates. *Journal of the American Chemical Society* **2005,** *127* (1), 210-216.

109. (a) Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R., A Comparative Study of Bioorthogonal Reactions with Azides. *ACS Chemical Biology* **2006,** *1* (10), 644-648;

(b) Speers, A. E.; Cravatt, B. F., Profiling enzyme activities in vivo using click chemistry methods. *Chem Biol* **2004,** *11* (4), 535-46.

110. (a) Jao, C. Y.; Salic, A., Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc Natl Acad Sci U S A* **2008,** *105* (41), 15779-84; (b) Salic, A.; Mitchison, T. J., A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc Natl Acad Sci U S A* **2008,** *105* (7), 2415-20; (c) Seo, T. S.; Bai, X.; Ruparel, H.; Li, Z.; Turro, N. J.; Ju, J., Photocleavable fluorescent nucleotides for DNA sequencing on a chip constructed by site-specific coupling chemistry. *Proc Natl Acad Sci U S A* **2004,** *101* (15), 5488-93.

111. (a) Jao, C. Y.; Roth, M.; Welti, R.; Salic, A., Metabolic labeling and direct imaging of choline phospholipids in vivo. *Proc Natl Acad Sci U S A* **2009,** *106* (36), 15332-7; (b) Neef, A. B.; Schultz, C., Selective fluorescence labeling of lipids in living cells. *Angew Chem Int Ed Engl* **2009,** *48* (8), 1498-500.

112. Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G., Bioconjugation by Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *Journal of the American Chemical Society* **2003,** *125* (11), 3192-3193.

113. (a) Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G., Adding Amino Acids with Novel Reactivity to the Genetic Code of Saccharomyces Cerevisiae. *Journal of the American Chemical Society* **2003,** *125* (39), 11782-11783; (b) Link, A. J.; Tirrell, D. A., Cell Surface Labeling of Escherichia coli via Copper(I)-Catalyzed [3+2] Cycloaddition. *Journal of the American Chemical Society* **2003,** *125* (37), 11164-11165; (c) Liu, C. C.; Schultz, P. G., Adding new chemistries to the genetic code. *Annu Rev Biochem* **2010,** *79*, 413-44; (d) Ngo, J. T.; Tirrell, D. A., Noncanonical Amino Acids in the Interrogation of Cellular Protein Synthesis. *Accounts of Chemical Research* **2011,** *44* (9), 677-685.

114. (a) Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G., Labeling Live Cells by Copper-Catalyzed Alkyne−Azide Click Chemistry. *Bioconjugate Chemistry* **2010,** *21* (10), 1912-1916; (b) Wolbers, F.; ter Braak, P.; Le Gac, S.; Luttge, R.; Andersson, H.; Vermes, I.; van den Berg, A., Viability study of HL60 cells in contact with commonly used microchip materials. *Electrophoresis* **2006,** *27* (24), 5073-80.

115. (a) Kennedy, D. C.; Lyn, R. K.; Pezacki, J. P., Cellular Lipid Metabolism Is Influenced by the Coordination Environment of Copper. *Journal of the American Chemical Society* **2009,** *131* (7), 2444-2445; (b) Link, A. J.; Vink, M. K. S.; Tirrell, D. A., Presentation and Detection of Azide Functionality in Bacterial Cell Surface Proteins. *Journal of the American Chemical Society* **2004,** *126* (34), 10598-10602; (c) Speers, A. E.; Adam, G. C.; Cravatt, B. F., Activity-Based Protein Profiling in Vivo Using a Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *Journal of the American Chemical Society* **2003,** *125* (16), 4686-4687.

116. (a) Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P., Increasing the efficacy of bioorthogonal click reactions for bioconjugation: a comparative study. *Angew Chem Int Ed Engl* **2011,** *50* (35), 8051-6; (b) Presolski, S. I.; Hong, V.; Cho, S.-H.; Finn, M. G., Tailored Ligand Acceleration of the Cu-Catalyzed Azide−Alkyne Cycloaddition Reaction: Practical and Mechanistic Implications. *Journal of the American Chemical Society* **2010,** *132* (41), 14570-14576; (c) Rodionov, V. O.; Presolski, S. I.; Díaz Díaz, D.; Fokin, V. V.; Finn, M. G.,

Ligand-Accelerated Cu-Catalyzed Azide−Alkyne Cycloaddition: A Mechanistic Report. *Journal of the American Chemical Society* **2007,** *129* (42), 12705-12712.

117. (a) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G., Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew Chem Int Ed Engl* **2009,** *48* (52), 9879-83; (b) Kennedy, D. C.; McKay, C. S.; Legault, M. C. B.; Danielson, D. C.; Blake, J. A.; Pegoraro, A. F.; Stolow, A.; Mester, Z.; Pezacki, J. P., Cellular Consequences of Copper Complexes Used To Catalyze Bioorthogonal Click Reactions. *Journal of the American Chemical Society* **2011,** *133* (44), 17993-18001; (c) Soriano del Amo, D.; Wang, W.; Jiang, H.; Besanceney, C.; Yan, A. C.; Levy, M.; Liu, Y.; Marlow, F. L.; Wu, P., Biocompatible Copper(I) Catalysts for in Vivo Imaging of Glycans. *Journal of the American Chemical Society* **2010,** *132* (47), 16893-16899.

118. (a) Brotherton, W. S.; Michaels, H. A.; Simmons, J. T.; Clark, R. J.; Dalal, N. S.; Zhu, L., Apparent Copper(II)-Accelerated Azide−Alkyne Cycloaddition. *Organic Letters* **2009,** *11* (21), 4954-4957; (b) Kuang, G.-C.; Michaels, H. A.; Simmons, J. T.; Clark, R. J.; Zhu, L., Chelation-Assisted, Copper(II)-Acetate-Accelerated Azide−Alkyne Cycloaddition. *The Journal of Organic Chemistry* **2010,** *75* (19), 6540-6548.

119. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *Journal of the American Chemical Society* **2004,** *126* (46), 15046-15047.

120. Wittig, G.; Krebs, A., Zur Existenz niedergliedriger Cycloalkine, I. *Chemische Berichte* **1961,** *94* (12), 3260-3275.

121. Bach, R. D., Ring Strain Energy in the Cyclooctyl System. The Effect of Strain Energy on [3 + 2] Cycloaddition Reactions with Azides. *Journal of the American Chemical Society* **2009,** *131* (14), 5233-5243.

122. Lin, F. L.; Hoyt, H. M.; van Halbeek, H.; Bergman, R. G.; Bertozzi, C. R., Mechanistic Investigation of the Staudinger Ligation. *Journal of the American Chemical Society* **2005,** *127* (8), 2686-2695.

123. Codelli, J. A.; Baskin, J. M.; Agard, N. J.; Bertozzi, C. R., Second-Generation Difluorinated Cyclooctynes for Copper-Free Click Chemistry. *Journal of the American Chemical Society* **2008,** *130* (34), 11486-11493.

124. (a) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R., Copper-free click chemistry for dynamic in vivo imaging. *Proceedings of the National Academy of Sciences* **2007,** *104* (43), 16793; (b) Chang, P. V.; Prescher, J. A.; Sletten, E. M.; Baskin, J. M.; Miller, I. A.; Agard, N. J.; Lo, A.; Bertozzi, C. R., Copper-free click chemistry in living animals. *Proceedings of the National Academy of Sciences* **2010,** *107* (5), 1821; (c) Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R., In Vivo Imaging of Membrane-Associated Glycans in Developing Zebrafish. *Science* **2008,** *320* (5876), 664.

125. Beatty, K. E.; Fisk, J. D.; Smart, B. P.; Lu, Y. Y.; Szychowski, J.; Hangauer, M. J.; Baskin, J. M.; Bertozzi, C. R.; Tirrell, D. A., Live-Cell Imaging of Cellular Proteins by a Strain-Promoted Azide–Alkyne Cycloaddition. *ChemBioChem* **2010,** *11* (15), 2092-2095.

126. (a) Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R., Rapid Cu-Free Click Chemistry with Readily Synthesized Biarylazacyclooctynones. *Journal of the American Chemical Society*

2010, *132* (11), 3688-3690; (b) Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G.-J., Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions. *Angewandte Chemie International Edition* **2008,** *47* (12), 2253-2255; (c) Mbua, N. E.; Guo, J.; Wolfert, M. A.; Steet, R.; Boons, G.-J., Strain-Promoted Alkyne–Azide Cycloadditions (SPAAC) Reveal New Features of Glycoconjugate Biosynthesis. *ChemBioChem* **2011,** *12* (12), 1912-1921.

127. (a) Debets, M. F.; van Berkel, S. S.; Schoffelen, S.; Rutjes, F. P. J. T.; van Hest, J. C. M.; van Delft, F. L., Aza-dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition. *Chemical Communications* **2010,** *46* (1), 97-99; (b) Gordon, C. G.; Mackey, J. L.; Jewett, J. C.; Sletten, E. M.; Houk, K. N.; Bertozzi, C. R., Reactivity of Biarylazacyclooctynones in Copper-Free Click Chemistry. *Journal of the American Chemical Society* **2012,** *134* (22), 9199-9208.

128. Gröst, C.; Berg, T., PYRROC: the first functionalized cycloalkyne that facilitates isomer-free generation of organic molecules by SPAAC. *Organic & Biomolecular Chemistry* **2015,** *13* (13), 3866-3870.

129. (a) McKay, C. S.; Blake, J. A.; Cheng, J.; Danielson, D. C.; Pezacki, J. P., Strain-promoted cycloadditions of cyclic nitrones with cyclooctynes for labeling human cancer cells. *Chemical Communications* **2011,** *47* (36), 10040-10042; (b) McKay, C. S.; Chigrinova, M.; Blake, J. A.; Pezacki, J. P., Kinetics studies of rapid strain-promoted [3 + 2]-cycloadditions of nitrones with biaryl-aza-cyclooctynone. *Organic & Biomolecular Chemistry* **2012,** *10* (15), 3066-3070; (c) McKay, C. S.; Moran, J.; Pezacki, J. P., Nitrones as dipoles for rapid strain-promoted 1,3-dipolar cycloadditions with cyclooctynes. *Chemical Communications* **2010,** *46* (6), 931-933.

130. (a) Gutsmiedl, K.; Wirges, C. T.; Ehmke, V.; Carell, T., Copper-Free “Click” Modification of DNA via Nitrile Oxide−Norbornene 1,3-Dipolar Cycloaddition. *Organic Letters* **2009,** *11* (11), 2405-2408; (b) Jawalekar, A. M.; Reubsaet, E.; Rutjes, F. P. J. T.; van Delft, F. L., Synthesis of isoxazoles by hypervalent iodine-induced cycloaddition of nitrile oxides to alkynes. *Chemical Communications* **2011,** *47* (11), 3198-3200; (c) Sanders, B. C.; Friscourt, F.; Ledin, P. A.; Mbua, N. E.; Arumugam, S.; Guo, J.; Boltje, T. J.; Popik, V. V.; Boons, G.-J., Metal-Free Sequential [3 + 2]-Dipolar Cycloadditions using Cyclooctynes and 1,3-Dipoles of Different Reactivity. *Journal of the American Chemical Society* **2011,** *133* (4), 949-957.

131. (a) Song, W.; Wang, Y.; Qu, J.; Lin, Q., Selective functionalization of a genetically encoded alkene-containing protein via "photoclick chemistry" in bacterial cells. *J Am Chem Soc* **2008,** *130* (30), 9654-5; (b) Yu, Z.; Ohulchanskyy, T. Y.; An, P.; Prasad, P. N.; Lin, Q., Fluorogenic, two-photon-triggered photoclick chemistry in live mammalian cells. *J Am Chem Soc* **2013,** *135* (45), 16766-9.

132. Lim, R. K. V.; Lin, Q., Azirine ligation: fast and selective protein conjugation via photoinduced azirine–alkene cycloaddition. *Chemical Communications* **2010,** *46* (42), 7993-7995.

133. Wallace, S.; Chin, J. W., Strain-promoted sydnone bicyclo-[6.1.0]-nonyne cycloaddition. *Chemical Science* **2014,** *5* (5), 1742-1744.

134. Andersen, K. A.; Aronoff, M. R.; McGrath, N. A.; Raines, R. T., Diazo Groups Endure Metabolism and Enable Chemoselectivity in Cellulo. *Journal of the American Chemical Society* **2015,** *137* (7), 2412-2415.

135. (a) Boutureira, O.; Bernardes, G. J. L., Advances in Chemical Protein Modification. *Chemical Reviews* **2015,** *115* (5), 2174-2195; (b) Chalker, J. M.; Wood, C. S. C.; Davis, B. G., A Convenient Catalyst for Aqueous and Protein Suzuki−Miyaura Cross-Coupling. *Journal of the American Chemical Society* **2009,** *131* (45), 16346-16347; (c) Lin, Y. A.; Boutureira, O.; Lercher, L.; Bhushan, B.; Paton, R. S.; Davis, B. G., Rapid Cross-Metathesis for Reversible Protein Modifications via Chemical Access to Se-Allyl-selenocysteine in Proteins. *Journal of the American Chemical Society* **2013,** *135* (33), 12156-12159; (d) Spicer, C. D.; Davis, B. G., Palladium-mediated site-selective Suzuki–Miyaura protein modification at genetically encoded aryl halides. *Chemical Communications* **2011,** *47* (6), 1698-1700; (e) Spicer, C. D.; Triemer, T.; Davis, B. G., Palladium-Mediated Cell-Surface Labeling. *Journal of the American Chemical Society* **2012,** *134* (2), 800-803; (f) Yang, M.; Yang, Y.; Chen, P. R., Transition-Metal-Catalyzed Bioorthogonal Cycloaddition Reactions. *Topics in Current Chemistry* **2015,** *374* (1), 2; (g) Li, J.; Chen, P. R., Development and application of bond cleavage reactions in bioorthogonal chemistry. *Nature Chemical Biology* **2016,** *12*, 129.

136. Li, J.; Lin, S.; Wang, J.; Jia, S.; Yang, M.; Hao, Z.; Zhang, X.; Chen, P. R., Ligand-Free Palladium-Mediated Site-Specific Protein Labeling Inside Gram-Negative Bacterial Pathogens. *Journal of the American Chemical Society* **2013,** *135* (19), 7330-7338.

137. Wang, J.; Cheng, B.; Li, J.; Zhang, Z.; Hong, W.; Chen, X.; Chen, P. R., Chemical Remodeling of Cell-Surface Sialic Acids through a Palladium-Triggered Bioorthogonal Elimination Reaction. *Angewandte Chemie International Edition* **2015,** *54* (18), 5364-5368.

138. (a) Mayer, S.; Lang, K., Tetrazines in Inverse-Electron-Demand Diels–Alder Cycloadditions and Their Use in Biology. *Synthesis* **2017,** *49* (04), 830-848; (b) Oliveira, B. L.; Guo, Z.; Bernardes, G. J. L., Inverse electron demand Diels–Alder reactions in chemical biology. *Chemical Society Reviews* **2017,** *46* (16), 4895-4950; (c) Šečkutė, J.; Devaraj, N. K., Expanding room for tetrazine ligations in the in vivo chemistry toolbox. *Current Opinion in Chemical Biology* **2013,** *17* (5), 761-767; (d) Wu, H.; Devaraj, N. K., Inverse Electron-Demand Diels–Alder Bioorthogonal Reactions. *Topics in Current Chemistry* **2015,** *374* (1), 3.

139. Diels, O.; Alder, K., Synthesen in der hydroaromatischen Reihe. *Justus Liebigs Annalen der Chemie* **1928,** *460* (1), 98-122.

140. Nicolaou, K. C.; Snyder, S. A.; Montagnon, T.; Vassilikogiannakis, G., The Diels–Alder Reaction in Total Synthesis. *Angewandte Chemie International Edition* **2002,** *41* (10), 1668-1698.

141. (a) Rideout, D. C.; Breslow, R., Hydrophobic acceleration of Diels-Alder reactions. *Journal of the American Chemical Society* **1980,** *102* (26), 7816-7817; (b) Otto, S.; Engberts, J. B. F. N., Diels–Alder reactions in water. *Pure and Applied Chemistry* **2000,** *72* (7), 1365-1372.

142. Bachmann, W. E.; Deno, N. C., The Diels--Alder Reaction of 1-Vinylnaphthalene with α,β- and α,β,γ,δ-Unsaturated Acids and Derivatives. *Journal of the American Chemical Society* **1949,** *71* (9), 3062-3072.

143. (a) Carboni, R. A.; Lindsey, R. V., Reactions of Tetrazines with Unsaturated Compounds. A New Synthesis of Pyridazines. *Journal of the American Chemical Society* **1959,** *81* (16), 4342-4346; (b) Müller, E.; Herrdegen, L., Einwirkung von wasserfreiem Hydrazin auf Nitrile. *Journal für Praktische Chemie* **1921,** *102* (4-7), 113-155.

144. Foster, R. A. A.; Willis, M. C., Tandem inverse-electron-demand hetero-/retro-Diels–Alder reactions for aromatic nitrogen heterocycle synthesis. *Chemical Society Reviews* **2013,** *42* (1), 63-76.

145. Knall, A.-C.; Slugovc, C., Inverse electron demand Diels-Alder (iEDDA)-initiated conjugation: a (high) potential click chemistry scheme. *Chemical Society Reviews* **2013,** *42* (12), 5131-5142.

146. Thalhammer, F.; Wallfahrer, U.; Sauer, J., Reaktivität einfacher offenkettiger und cyclischer dienophile bei Diels-Alder-reaktionen mit inversem elektronenbedarf. *Tetrahedron Letters* **1990,** *31* (47), 6851-6854.

147. (a) Blackman, M. L.; Royzen, M.; Fox, J. M., Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J Am Chem Soc* **2008,** *130* (41), 13518-9; (b) Karver, M. R.; Weissleder, R.; Hilderbrand, S. A., Synthesis and evaluation of a series of 1,2,4,5-tetrazines for bioorthogonal conjugation. *Bioconjug Chem* **2011,** *22* (11), 2263-70; (c) Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W., Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat Chem* **2012,** *4* (4), 298-304; (d) Wang, D.; Chen, W.; Zheng, Y.; Dai, C.; Wang, K.; Ke, B.; Wang, B., 3,6-Substituted-1,2,4,5-tetrazines: tuning reaction rates for staged labeling applications. *Organic & Biomolecular Chemistry* **2014,** *12* (23), 3950-3955; (e) Chen, W.; Wang, D.; Dai, C.; Hamelberg, D.; Wang, B., Clicking 1,2,4,5-tetrazine and cyclooctynes with tunable reaction rates. *Chemical Communications* **2012,** *48* (12), 1736-1738; (f) Kamber, D. N.; Nazarova, L. A.; Liang, Y.; Lopez, S. A.; Patterson, D. M.; Shih, H. W.; Houk, K. N.; Prescher, J. A., Isomeric cyclopropenes exhibit unique bioorthogonal reactivities. *J Am Chem Soc* **2013,** *135* (37), 13680-3; (g) Taylor, M. T.; Blackman, M. L.; Dmitrenko, O.; Fox, J. M., Design and Synthesis of Highly Reactive Dienophiles for the Tetrazine–trans-Cyclooctene Ligation. *Journal of the American Chemical Society* **2011,** *133* (25), 9646-9649.

148. (a) Fukui, K., Recognition of stereochemical paths by orbital interaction. *Accounts of Chemical Research* **1971,** *4* (2), 57-64; (b) Fukui, K., The Role of Frontier Orbitals in Chemical Reactions (Nobel Lecture). *Angewandte Chemie International Edition in English* **1982,** *21* (11), 801-809; (c) Boger, D. L., *Modern Organic Synthesis Lecture Notes*. TSRI Press: San Diego, CA, 1999.

149. Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A., Tetrazine-based cycloadditions: application to pretargeted live cell imaging. *Bioconjug Chem* **2008,** *19* (12), 2297-9.

150. Kronister, S.; Svatunek, D.; Denk, C.; Mikula, H., Acylation-Mediated 'Kinetic Turn-On' of 3-Amino-1,2,4,5-tetrazines. *Synlett* **2018,** *14* (10), 1297-1302.

151. (a) Kämpchen, T.; Massa, W.; Overheu, W.; Schmidt, R.; Seitz, G., Zur Kenntnis von Reaktionen des 1,2,4,5-Tetrazin-3,6-dicarbonsäure-dimethylesters mit Nucleophilen. *Chemische Berichte* **1982,** *115* (2), 683-694; (b) Blizzard, R. J.; Backus, D. R.; Brown, W.; Bazewicz, C. G.; Li, Y.; Mehl, R. A., Ideal Bioorthogonal Reactions Using A Site-Specifically Encoded Tetrazine Amino Acid. *J Am Chem Soc* **2015,** *137* (32), 10044-7.

152. (a) Lee, Y.-J.; Kurra, Y.; Yang, Y.; Torres-Kolbus, J.; Deiters, A.; Liu, W. R., Genetically encoded unstrained olefins for live cell labeling with tetrazine dyes. *Chemical Communications* **2014,** *50* (86), 13085-13088; (b) Shang, X.; Song, X.; Faller, C.; Lai, R.; Li, H.; Cerny, R.; Niu, W.; Guo, J., Fluorogenic protein labeling using a genetically encoded unstrained alkene. *Chemical Science* **2017,** *8* (2), 1141-1145.

153. (a) Eising, S.; Lelivelt, F.; Bonger, K. M., Vinylboronic Acids as Fast Reacting, Synthetically Accessible, and Stable Bioorthogonal Reactants in the Carboni–Lindsey Reaction. *Angewandte Chemie International Edition* **2016,** *55* (40), 12243-12247; (b) Vivat, J. F.; Adams, H.; Harrity, J. P. A., Ambient Temperature Nitrogen-Directed Difluoroalkynylborane Carboni−Lindsey Cycloaddition Reactions. *Organic Letters* **2010,** *12* (1), 160-163.

154. Walker, R.; Conrad, R. M.; Grubbs, R. H., The Living ROMP of trans-Cyclooctene. *Macromolecules* **2009,** *42* (3), 599-605.

155. Vrabel, M.; Kolle, P.; Brunner, K. M.; Gattner, M. J.; Lopez-Carrillo, V.; de Vivie-Riedle, R.; Carell, T., Norbornenes in inverse electron-demand Diels-Alder reactions. *Chemistry* **2013,** *19* (40), 13309-12.

156. Cserép, G. B.; Demeter, O.; Bätzner, E.; Kállay, M.; Wagenknecht, H.-A.; Kele, P., Synthesis and Evaluation of Nicotinic Acid Derived Tetrazines for Bioorthogonal Labeling. *Synthesis* **2015,** *47* (18), 2738-2744.

157. (a) Plass, T.; Milles, S.; Koehler, C.; Szymański, J.; Mueller, R.; Wießler, M.; Schultz, C.; Lemke, E. A., Amino Acids for Diels–Alder Reactions in Living Cells. *Angewandte Chemie International Edition* **2012,** *51* (17), 4166-4170; (b) Nikic, I.; Plass, T.; Schraidt, O.; Szymanski, J.; Briggs, J. A.; Schultz, C.; Lemke, E. A., Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy. *Angew Chem Int Ed Engl* **2014,** *53* (8), 2245-9; (c) Wagner, J. A.; Mercadante, D.; Nikic, I.; Lemke, E. A.; Grater, F., Origin of Orthogonality of Strain-Promoted Click Reactions. *Chemistry* **2015,** *21* (35), 12431-5.

158. Yang, J.; Liang, Y.; Seckute, J.; Houk, K. N.; Devaraj, N. K., Synthesis and reactivity comparisons of 1-methyl-3-substituted cyclopropene mini-tags for tetrazine bioorthogonal reactions. *Chemistry* **2014,** *20* (12), 3365-75.

159. Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J.; Rutjes, F. P.; van Hest, J. C.; Lefeber, D. J.; Friedl, P.; van Delft, F. L., Readily accessible bicyclononynes for bioorthogonal labeling and three-dimensional imaging of living cells. *Angew Chem Int Ed Engl* **2010,** *49* (49), 9422-5.

160. Selvaraj, R.; Fox, J. M., trans-Cyclooctene—a stable, voracious dienophile for bioorthogonal labeling. *Current Opinion in Chemical Biology* **2013,** *17* (5), 753-760.

161. (a) Schoch, J.; Staudt, M.; Samanta, A.; Wiessler, M.; Jaschke, A., Site-specific one-pot dual labeling of DNA by orthogonal cycloaddition chemistry. *Bioconjug Chem* **2012,** *23* (7), 1382-6; (b) Rossin, R.; van den Bosch, S. M.; ten Hoeve, W.; Carvelli, M.; Versteegen, R. M.; Lub, J.; Robillard, M. S., Highly Reactive trans-Cyclooctene Tags with Improved Stability for Diels–Alder Chemistry in Living Systems. *Bioconjugate Chemistry* **2013,** *24* (7), 1210-1217.

162. Darko, A.; Wallace, S.; Dmitrenko, O.; Machovina, M. M.; Mehl, R. A.; Chin, J. W.; Fox, J. M., Conformationally strained trans-cyclooctene with improved stability and excellent reactivity in tetrazine ligation. *Chemical Science* **2014,** *5* (10), 3770-3776.

163. Pinner, A., Ueber die Einwirkung von Hydrazin auf Imidoäther. *Berichte der deutschen chemischen Gesellschaft* **1893,** *26* (2), 2126-2135.

164. Pinner, A., Ueber die Einwirkung von Hydrazin auf die Imidoäther. *Berichte der deutschen chemischen Gesellschaft* **1897,** *30* (2), 1871-1890.

165. Lin, C.-H.; Lieber, E.; Horwitz, J. P., The Synthesis of sym-Diaminotetrazine1a,1b. *Journal of the American Chemical Society* **1954,** *76* (2), 427-430.

166. Coburn, M. D.; Buntain, G. A.; Harris, B. W.; Hiskey, M. A.; Lee, K. Y.; Ott, D. G., An improved synthesis of 3,6-diamino-1,2,4,5-tetrazine. II. From triaminoguanidine and 2,4-pentanedione. *Journal of Heterocyclic Chemistry* **1991,** *28* (8), 2049-2050.

167. Lang, S. A.; Johnson, B. D.; Cohen, E., Novel synthesis of unsymmetrically substituted s-tetrazines. *Journal of Heterocyclic Chemistry* **1975,** *12* (6), 1143-1153.

168. Grakauskas, V. A.; Tomasewski, A. J.; Horwitz, J. P., IV. Some 3,6-Unsymmetrically Disubstituted 1,2,4,5-Tetrazines1a,b,c. *Journal of the American Chemical Society* **1958,** *80* (12), 3155-3159.

169. M. Asselin, C.; C. Fraser, G.; K. Hall Jr, H.; Edward Lindsell, W.; B. Padias, A.; N. Preston, P., Synthesis and metallation of ferrocenylimines derived from ligating diaminoheteroarenes. *Journal of the Chemical Society, Dalton Transactions* **1997,** (20), 3765-3772.

170. Curtius, T.; Dedichen, G. M., 19. Synthesen von Benzolhydrazinen mittelst Hydrazinhydrat. *Journal für Praktische Chemie* **1894,** *50* (1), 241-274.

171. Junghahn, A.; Bunimowicz, J., Ueber die Einwirkung von Hydrazin auf Thiamide. *Berichte der deutschen chemischen Gesellschaft* **1902,** *35* (4), 3932-3940.

172. Stolleé, R., Über die Überführung von Hydrazinderivaten in heterocyklische Verbindungen. *Journal für Praktische Chemie* **1903,** *68* (1), 464-468.

173. Curtius, T.; Darapsky, A.; Müller, E., Über die Umwandlungsprodukte des Diazoessigesters unter dem Einfluß von Alkalien. *Berichte der deutschen chemischen Gesellschaft* **1908,** *41* (2), 3161-3172.

174. (a) Audebert, P.; Sadki, S.; Miomandre, F.; Clavier, G.; Claude Vernières, M.; Saoud, M.; Hapiot, P., Synthesis of new substituted tetrazines: electrochemical and spectroscopic properties. *New Journal of Chemistry* **2004,** *28* (3), 387-392; (b) Li, C.; Ge, H.; Yin, B.; She, M.; Liu, P.; Li, X.; Li, J., Novel 3,6-unsymmetrically disubstituted-1,2,4,5-tetrazines: S-induced one-pot synthesis, properties and theoretical study. *RSC Advances* **2015,** *5* (16), 12277-12286.

175. Abdel, N. O.; Kira, M. A.; Tolba, M. N., A direct synthesis of dihydrotetrazines. *Tetrahedron Letters* **1968,** *9* (35), 3871-3872.

176. Takimoto, H. H.; Denault, G. C., 3-Amino-s-tetrazines from the thermal decomposition of 4-amino-3-azido-s-triazoles. *Tetrahedron Letters* **1966,** *7* (44), 5369-5373.

177. Werbel, L. M.; McNamara, D. J.; Colbry, N. L.; Johnson, J. L.; Degnan, M. J.; Whitney, B., Synthesis and antimalarial effects of N,N-dialkyl-6-(substituted phenyl)-1,2,4,5-tetrazin-3-amines. *Journal of Heterocyclic Chemistry* **1979,** *16* (5), 881-894.

178. Meresz, O.; Foster-Verner, P. A., Synthesis of 3-monosubstituted s-tetrazines and their reactions with monosubstituted acetylenes. *Journal of the Chemical Society, Chemical Communications* **1972,** (16), 950-951.

179. (a) Esmail, R.; Kurzer, F., Heterocyclic compounds from urea derivatives. Part XXIII. Thiobenzoylated thiocarbonohydrazides and their cyclisation. *Journal of the Chemical Society, Perkin Transactions 1* **1975,** (18), 1787-1791; (b) Counotte-Potman, A.; Van der Plas, H. C.; Van Veldhuizen, B.; Landheer, C. A., Occurrence of the SN(ANRORC) mechanism in the hydrazination of 1,2,4,5-tetrazines. *The Journal of Organic Chemistry* **1981,** *46* (25), 5102-5109.

180. (a) Mangia, A.; Bortesi, F.; Amendola, U., 3-Monosubstituted and 3,6-Unsymmetrically disubstituted 1,2,4,5-Tetrazines. A general method of synthesis. *Journal of Heterocyclic Chemistry* **1977,** *14* (4), 587-593; (b) Johnson, J. L.; Whitney, B.; Werbel, L. M., Synthesis of 6-(arylthio)- and 6-[(arylmethyl)thio] -1,2,4,5-tetrazin-3-amines and n-phenyl- and n-(phenylmethyl)-1,2,4,5-tetrazine-3,6-diamines as potential antimalarial agents. *Journal of Heterocyclic Chemistry* **1980,** *17* (3), 501-506.

181. Sandström, J., The true Dithio-p-urazine and some related sym-tetrazines. *Acta Chemica Scandinavica* **1961,** *15*, 1575 - 1582.

182. Fields, S. C.; Parker, M. H.; Erickson, W. R., A Simple Route to Unsymmetrically Substituted 1,2,4,5-Tetrazines. *The Journal of Organic Chemistry* **1994,** *59* (26), 8284-8287.

183. (a) Chavez, D. E.; Hiskey, M. A., Synthesis of the bi-heterocyclic parent ring system 1,2,4-triazolo[4,3-b][1,2,4,5]tetrazine and some 3,6-disubstituted derivatives. *Journal of Heterocyclic Chemistry* **1998,** *35* (6), 1329-1332; (b) Chavez, D. E.; Hiskey, M. A., 1,2,4,5-tetrazine based energetic materials. *Journal of Energetic Materials* **1999,** *17* (4), 357-377.

184. Helm, M. D.; Plant, A.; Harrity, J. P., A novel approach to functionalised pyridazinone arrays. *Org Biomol Chem* **2006,** *4* (23), 4278-80.

185. Novák, Z.; Kotschy, A., First Cross-Coupling Reactions on Tetrazines. *Organic Letters* **2003,** *5* (19), 3495-3497.

186. Leconte, N.; Keromnes-Wuillaume, A.; Suzenet, F.; Guillaumet, G., Efficient Palladium-Catalyzed Synthesis of Unsymmetrical (Het)aryltetrazines. *Synlett* **2007,** *2007* (02), 0204-0210.

187. Yang, J.; Karver, M. R.; Li, W.; Sahu, S.; Devaraj, N. K., Metal-Catalyzed One-Pot Synthesis of Tetrazines Directly from Aliphatic Nitriles and Hydrazine. *Angewandte Chemie International Edition* **2012,** *51* (21), 5222-5225.

188. (a) Carlson, J. C.; Meimetis, L. G.; Hilderbrand, S. A.; Weissleder, R., BODIPY-tetrazine derivatives as superbright bioorthogonal turn-on probes. *Angew Chem Int Ed Engl* **2013,** *52* (27), 6917-20; (b) Denk, C.; Svatunek, D.; Mairinger, S.; Stanek, J.; Filip, T.;

Matscheko, D.; Kuntner, C.; Wanek, T.; Mikula, H., Design, Synthesis, and Evaluation of a Low-Molecular-Weight (11)C-Labeled Tetrazine for Pretargeted PET Imaging Applying Bioorthogonal in Vivo Click Chemistry. *Bioconjug Chem* **2016,** *27* (7), 1707-12; (c) Ehret, F.; Wu, H.; Alexander, S. C.; Devaraj, N. K., Electrochemical Control of Rapid Bioorthogonal Tetrazine Ligations for Selective Functionalization of Microelectrodes. *J Am Chem Soc* **2015,** *137* (28), 8876-9; (d) Meimetis, L. G.; Carlson, J. C.; Giedt, R. J.; Kohler, R. H.; Weissleder, R., Ultrafluorogenic coumarin-tetrazine probes for real-time biological imaging. *Angew Chem Int Ed Engl* **2014,** *53* (29), 7531-4.

189. Furuta, K.; Miwa, Y.; Iwanaga, K.; Yamamoto, H., Acyloxyborane: an activating device for carboxylic acids. *J Am Chem Soc* **1988,** *110* (18), 6254-5.

190. Royzen, M.; Yap, G. P.; Fox, J. M., A photochemical synthesis of functionalized trans-cyclooctenes driven by metal complexation. *J Am Chem Soc* **2008,** *130* (12), 3760-1.

191. (a) Petiniot, N.; Anciaux, A. J.; Noels, A. F.; Hubert, A. J.; Teyssié, P., Rhodium catalysed cyclopropenation of acetylenes. *Tetrahedron Letters* **1978,** *19* (14), 1239-1242; (b) Lou, Y.; Horikawa, M.; Kloster, R. A.; Hawryluk, N. A.; Corey, E. J., A new chiral Rh(II) catalyst for enantioselective [2 + 1]-cycloaddition. mechanistic implications and applications. *J Am Chem Soc* **2004,** *126* (29), 8916-8.

192. Patterson, D. M.; Nazarova, L. A.; Xie, B.; Kamber, D. N.; Prescher, J. A., Functionalized cyclopropenes as bioorthogonal chemical reporters. *J Am Chem Soc* **2012,** *134* (45), 18638-43.

193. Devaraj, N. K.; Upadhyay, R.; Haun, J. B.; Hilderbrand, S. A.; Weissleder, R., Fast and sensitive pretargeted labeling of cancer cells through a tetrazine/trans-cyclooctene cycloaddition. *Angew Chem Int Ed Engl* **2009,** *48* (38), 7013-6.

194. Devaraj, N. K.; Hilderbrand, S.; Upadhyay, R.; Mazitschek, R.; Weissleder, R., Bioorthogonal turn-on probes for imaging small molecules inside living cells. *Angew Chem Int Ed Engl* **2010,** *49* (16), 2869-72.

195. Wu, H.; Yang, J.; Seckute, J.; Devaraj, N. K., In situ synthesis of alkenyl tetrazines for highly fluorogenic bioorthogonal live-cell imaging probes. *Angew Chem Int Ed Engl* **2014,** *53* (23), 5805-9.

196. Yang, J.; Seckute, J.; Cole, C. M.; Devaraj, N. K., Live-cell imaging of cyclopropene tags with fluorogenic tetrazine cycloadditions. *Angew Chem Int Ed Engl* **2012,** *51* (30), 7476-9.

197. Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y., Diels-Alder cycloaddition for fluorophore targeting to specific proteins inside living cells. *J Am Chem Soc* **2012,** *134* (2), 792-5.

198. (a) Borrmann, A.; Milles, S.; Plass, T.; Dommerholt, J.; Verkade, J. M.; Wiessler, M.; Schultz, C.; van Hest, J. C.; van Delft, F. L.; Lemke, E. A., Genetic encoding of a bicyclo[6.1.0]nonyne-charged amino acid enables fast cellular protein imaging by metal-free ligation. *Chembiochem* **2012,** *13* (14), 2094-9; (b) Elliott, T. S.; Townsley, F. M.; Bianco, A.; Ernst, R. J.; Sachdeva, A.; Elsasser, S. J.; Davis, L.; Lang, K.; Pisa, R.; Greiss, S.; Lilley, K. S.; Chin, J. W., Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat Biotechnol* **2014,** *32* (5), 465-72; (c) Kaya, E.; Vrabel, M.; Deiml, C.; Prill, S.; Fluxa, V. S.; Carell, T., A genetically encoded norbornene amino acid

for the mild and selective modification of proteins in a copper-free click reaction. *Angew Chem Int Ed Engl* **2012,** *51* (18), 4466-9.

199. Tsai, Y.-H.; Essig, S.; James, J. R.; Lang, K.; Chin, J. W., Selective, rapid and optically switchable regulation of protein function in live mammalian cells. *Nature chemistry* **2015,** *7* (7), 554-561.

200. Brott, B. K.; Alessandrini, A.; Largaespada, D. A.; Copeland, N. G.; Jenkins, N. A.; Crews, C. M.; Erikson, R. L., MEK2 is a kinase related to MEK1 and is differentially expressed in murine tissues. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **1993,** *4* (11), 921-9.

201. Sebolt-Leopold, J. S.; Herrera, R., Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nature reviews. Cancer* **2004,** *4* (12), 937-47.

202. Rice, K. D.; Aay, N.; Anand, N. K.; Blazey, C. M.; Bowles, O. J.; Bussenius, J.; Costanzo, S.; Curtis, J. K.; Defina, S. C.; Dubenko, L.; Engst, S.; Joshi, A. A.; Kennedy, A. R.; Kim, A. I.; Koltun, E. S.; Lougheed, J. C.; Manalo, J.-C. L.; Martini, J.-F.; Nuss, J. M.; Peto, C. J.; Tsang, T. H.; Yu, P.; Johnston, S., Novel Carboxamide-Based Allosteric MEK Inhibitors: Discovery and Optimization Efforts toward XL518 (GDC-0973). *ACS Medicinal Chemistry Letters* **2012,** *3* (5), 416-421.

203. (a) Keliher, E. J.; Reiner, T.; Turetsky, A.; Hilderbrand, S. A.; Weissleder, R., High-yielding, two-step 18F labeling strategy for 18F-PARP1 inhibitors. *ChemMedChem* **2011,** *6* (3), 424-7; (b) Li, Z.; Cai, H.; Hassink, M.; Blackman, M. L.; Brown, R. C.; Conti, P. S.; Fox, J. M., Tetrazine-trans-cyclooctene ligation for the rapid construction of 18F labeled probes. *Chem Commun (Camb)* **2010,** *46* (42), 8043-5.

204. Wang, M.; Svatunek, D.; Rohlfing, K.; Liu, Y.; Wang, H.; Giglio, B.; Yuan, H.; Wu, Z.; Li, Z.; Fox, J., Conformationally Strained trans-Cyclooctene (sTCO) Enables the Rapid Construction of (18)F-PET Probes via Tetrazine Ligation. *Theranostics* **2016,** *6* (6), 887-95.

205. Knight, J. C.; Richter, S.; Wuest, M.; Way, J. D.; Wuest, F., Synthesis and evaluation of an 18F-labelled norbornene derivative for copper-free click chemistry reactions. *Org Biomol Chem* **2013,** *11* (23), 3817-25.

206. Denk, C.; Svatunek, D.; Filip, T.; Wanek, T.; Lumpi, D.; Frohlich, J.; Kuntner, C.; Mikula, H., Development of a (18) F-labeled tetrazine with favorable pharmacokinetics for bioorthogonal PET imaging. *Angew Chem Int Ed Engl* **2014,** *53* (36), 9655-9.

207. Herth, M. M.; Andersen, V. L.; Lehel, S.; Madsen, J.; Knudsen, G. M.; Kristensen, J. L., Development of a (11)C-labeled tetrazine for rapid tetrazine-trans-cyclooctene ligation. *Chem Commun (Camb)* **2013,** *49* (36), 3805-7.

208. Reiner, T.; Keliher, E. J.; Earley, S.; Marinelli, B.; Weissleder, R., Synthesis and in vivo imaging of a 18F-labeled PARP1 inhibitor using a chemically orthogonal scavenger-assisted high-performance method. *Angew Chem Int Ed Engl* **2011,** *50* (8), 1922-5.

209. Liu, S.; Hassink, M.; Selvaraj, R.; Yap, L. P.; Park, R.; Wang, H.; Chen, X.; Fox, J. M.; Li, Z.; Conti, P. S., Efficient 18F labeling of cysteine-containing peptides and proteins using tetrazine-trans-cyclooctene ligation. *Mol Imaging* **2013,** *12* (2), 121-8.

210. Rossin, R.; Verkerk, P. R.; van den Bosch, S. M.; Vulders, R. C.; Verel, I.; Lub, J.; Robillard, M. S., In vivo chemistry for pretargeted tumor imaging in live mice. *Angew Chem Int Ed Engl* **2010,** *49* (19), 3375-8.

211. (a) Zeglis, B. M.; Mohindra, P.; Weissmann, G. I.; Divilov, V.; Hilderbrand, S. A.; Weissleder, R.; Lewis, J. S., Modular strategy for the construction of radiometalated antibodies for positron emission tomography based on inverse electron demand Diels-Alder click chemistry. *Bioconjug Chem* **2011,** *22* (10), 2048-59; (b) Zeglis, B. M.; Sevak, K. K.; Reiner, T.; Mohindra, P.; Carlin, S. D.; Zanzonico, P.; Weissleder, R.; Lewis, J. S., A pretargeted PET imaging strategy based on bioorthogonal Diels-Alder click chemistry. *J Nucl Med* **2013,** *54* (8), 1389-96.

212. Hou, S.; Choi, J. S.; Garcia, M. A.; Xing, Y.; Chen, K. J.; Chen, Y. M.; Jiang, Z. K.; Ro, T.; Wu, L.; Stout, D. B.; Tomlinson, J. S.; Wang, H.; Chen, K.; Tseng, H. R.; Lin, W. Y., Pretargeted Positron Emission Tomography Imaging That Employs Supramolecular Nanoparticles with in Vivo Bioorthogonal Chemistry. *ACS Nano* **2016,** *10* (1), 1417-24.

213. Emmetiere, F.; Irwin, C.; Viola-Villegas, N. T.; Longo, V.; Cheal, S. M.; Zanzonico, P.; Pillarsetty, N.; Weber, W. A.; Lewis, J. S.; Reiner, T., (18)F-labeled-bioorthogonal liposomes for in vivo targeting. *Bioconjug Chem* **2013,** *24* (11), 1784-9.

214. Schoch, J.; Ameta, S.; Jaschke, A., Inverse electron-demand Diels-Alder reactions for the selective and efficient labeling of RNA. *Chem Commun (Camb)* **2011,** *47* (46), 12536-7.

215. Neef, A. B.; Luedtke, N. W., Dynamic metabolic labeling of DNA in vivo with arabinosyl nucleosides. *Proc Natl Acad Sci U S A* **2011,** *108* (51), 20404-9.

216. Hermanson, G. T., *Bioconjugate Techniques*. Academic Press: San Diego, 1996.

217. (a) Schoch, J.; Wiessler, M.; Jaschke, A., Post-synthetic modification of DNA by inverse-electron-demand Diels-Alder reaction. *J Am Chem Soc* **2010,** *132* (26), 8846-7; (b) Asare-Okai, P. N.; Agustin, E.; Fabris, D.; Royzen, M., Site-specific fluorescence labelling of RNA using bio-orthogonal reaction of trans-cyclooctene and tetrazine. *Chem Commun (Camb)* **2014,** *50* (58), 7844-7.

218. Pyka, A. M.; Domnick, C.; Braun, F.; Kath-Schorr, S., Diels-Alder cycloadditions on synthetic RNA in mammalian cells. *Bioconjug Chem* **2014,** *25* (8), 1438-43.

219. (a) Pitulle, C.; Kleineidam, R. G.; Sproat, B.; Krupp, G., Initiator oligonucleotides for the combination of chemical and enzymatic RNA synthesis. *Gene* **1992,** *112* (1), 101-5; (b) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C., Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res* **1987,** *15* (21), 8783-98.

220. Ren, X.; El-Sagheer, A. H.; Brown, T., Azide and trans-cyclooctene dUTPs: incorporation into DNA probes and fluorescent click-labelling. *Analyst* **2015,** *140* (8), 2671-8.

221. Domnick, C.; Eggert, F.; Kath-Schorr, S., Site-specific enzymatic introduction of a norbornene modified unnatural base into RNA and application in post-transcriptional labeling. *Chem Commun (Camb)* **2015,** *51* (39), 8253-6.

222. Eggert, F.; Kath-Schorr, S., A cyclopropene-modified nucleotide for site-specific RNA labeling using genetic alphabet expansion transcription. *Chem Commun (Camb)* **2016,** *52* (45), 7284-7.

223. Dhami, K.; Malyshev, D. A.; Ordoukhanian, P.; Kubelka, T.; Hocek, M.; Romesberg, F. E., Systematic exploration of a class of hydrophobic unnatural base pairs yields multiple new candidates for the expansion of the genetic alphabet. *Nucleic Acids Res* **2014,** *42* (16), 10235-44.

224. Wang, K.; Wang, D.; Ji, K.; Chen, W.; Zheng, Y.; Dai, C.; Wang, B., Post-synthesis DNA modifications using a trans-cyclooctene click handle. *Org Biomol Chem* **2015,** *13* (3), 909-15.

225. (a) Holstein, J. M.; Anhäuser, L.; Rentmeister, A., Modifying the 5′-Cap for Click Reactions of Eukaryotic mRNA and To Tune Translation Efficiency in Living Cells. *Angewandte Chemie International Edition* **2016,** *55* (36), 10899-10903; (b) Holstein, J. M.; Stummer, D.; Rentmeister, A., Enzymatic modification of 5′-capped RNA with a 4-vinylbenzyl group provides a platform for photoclick and inverse electron-demand Diels–Alder reaction. *Chemical Science* **2015,** *6* (2), 1362-1369.

226. Holstein, J. M.; Muttach, F.; Schiefelbein, S. H. H.; Rentmeister, A., Dual 5′ Cap Labeling Based on Regioselective RNA Methyltransferases and Bioorthogonal Reactions. *Chemistry – A European Journal* **2017,** *23* (25), 6165-6173.

227. (a) Bußkamp, H.; Batroff, E.; Niederwieser, A.; Abdel-Rahman, O. S.; Winter, R. F.; Wittmann, V.; Marx, A., Efficient labelling of enzymatically synthesized vinyl-modified DNA by an inverse-electron-demand Diels–Alder reaction. *Chemical Communications* **2014,** *50* (74), 10827-10829; (b) Rieder, U.; Luedtke, N. W., Alkene-tetrazine ligation for imaging cellular DNA. *Angew Chem Int Ed Engl* **2014,** *53* (35), 9168-72.

228. Seckute, J.; Yang, J.; Devaraj, N. K., Rapid oligonucleotide-templated fluorogenic tetrazine ligations. *Nucleic Acids Res* **2013,** *41* (15), e148.

229. Wu, H.; Cisneros, B. T.; Cole, C. M.; Devaraj, N. K., Bioorthogonal Tetrazine-Mediated Transfer Reactions Facilitate Reaction Turnover in Nucleic Acid-Templated Detection of MicroRNA. *Journal of the American Chemical Society* **2014,** *136* (52), 17942-17945.

230. Wu, H.; Alexander, S. C.; Jin, S.; Devaraj, N. K., A Bioorthogonal Near-Infrared Fluorogenic Probe for mRNA Detection. *Journal of the American Chemical Society* **2016,** *138* (36), 11429-11432.

231. (a) Agarwal, P.; Beahm, B. J.; Shieh, P.; Bertozzi, C. R., Systemic Fluorescence Imaging of Zebrafish Glycans with Bioorthogonal Chemistry. *Angew Chem Int Ed Engl* **2015,** *54* (39), 11504-10; (b) Laughlin, S. T.; Bertozzi, C. R., Imaging the glycome. *Proceedings of the National Academy of Sciences* **2009,** *106* (1), 12-17.

232. (a) Chang, P. V.; Bertozzi, C. R., Imaging beyond the proteome. *Chemical Communications* **2012,** *48* (71), 8864-8879; (b) Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R., In vivo imaging of membrane-associated glycans in developing zebrafish. *Science* **2008,** *320* (5876), 664-7; (c) Prescher, J. A.; Dube, D. H.; Bertozzi, C. R., Chemical remodelling of cell surfaces in living animals. *Nature* **2004,** *430* (7002), 873-877.

233. Cole, C. M.; Yang, J.; Seckute, J.; Devaraj, N. K., Fluorescent live-cell imaging of metabolically incorporated unnatural cyclopropene-mannosamine derivatives. *Chembiochem* **2013,** *14* (2), 205-8.

234. Spate, A. K.; Busskamp, H.; Niederwieser, A.; Schart, V. F.; Marx, A.; Wittmann, V., Rapid labeling of metabolically engineered cell-surface glycoconjugates with a carbamate-linked cyclopropene reporter. *Bioconjug Chem* **2014,** *25* (1), 147-54.

235. Spate, A. K.; Dold, J. E.; Batroff, E.; Schart, V. F.; Wieland, D. E.; Baudendistel, O. R.; Wittmann, V., Exploring the Potential of Norbornene-Modified Mannosamine Derivatives for Metabolic Glycoengineering. *Chembiochem* **2016,** *17* (14), 1374-83.

236. Niederwieser, A.; Spate, A. K.; Nguyen, L. D.; Jungst, C.; Reutter, W.; Wittmann, V., Two-color glycan labeling of live cells by a combination of Diels-Alder and click chemistry. *Angew Chem Int Ed Engl* **2013,** *52* (15), 4265-8.

237. Neves, A. A.; Stockmann, H.; Wainman, Y. A.; Kuo, J. C.; Fawcett, S.; Leeper, F. J.; Brindle, K. M., Imaging cell surface glycosylation in vivo using "double click" chemistry. *Bioconjug Chem* **2013,** *24* (6), 934-41.

238. Doll, F.; Buntz, A.; Späte, A.-K.; Schart, V. F.; Timper, A.; Schrimpf, W.; Hauck, C. R.; Zumbusch, A.; Wittmann, V., Visualization of Protein-Specific Glycosylation inside Living Cells. *Angewandte Chemie International Edition* **2016,** *55* (6), 2262-2266.

239. Schart, V. F.; Hassenrück, J.; Späte, A.-K.; Dold, J. E. G. A.; Fahrner, R.; Wittmann, V., Triple Orthogonal Labeling of Glycans by Applying Photoclick Chemistry. *ChemBioChem* **2019,** *20* (2), 166-171.

240. Machida, T.; Lang, K.; Xue, L.; Chin, J. W.; Winssinger, N., Site-Specific Glycoconjugation of Protein via Bioorthogonal Tetrazine Cycloaddition with a Genetically Encoded trans-Cyclooctene or Bicyclononyne. *Bioconjug Chem* **2015,** *26* (5), 802-6.

241. Erdmann, R. S.; Takakura, H.; Thompson, A. D.; Rivera-Molina, F.; Allgeyer, E. S.; Bewersdorf, J.; Toomre, D.; Schepartz, A., Super-resolution imaging of the Golgi in live cells with a bioorthogonal ceramide probe. *Angew Chem Int Ed Engl* **2014,** *53* (38), 10242-6.

242. Lang, K.; Chin, J. W., Bioorthogonal reactions for labeling proteins. *ACS Chem Biol* **2014,** *9* (1), 16-20.

243. Liang, Y.; Mackey, J. L.; Lopez, S. A.; Liu, F.; Houk, K. N., Control and Design of Mutual Orthogonality in Bioorthogonal Cycloadditions. *Journal of the American Chemical Society* **2012,** *134* (43), 17904-17907.

244. Willems, L. I.; Li, N.; Florea, B. I.; Ruben, M.; van der Marel, G. A.; Overkleeft, H. S., Triple Bioorthogonal Ligation Strategy for Simultaneous Labeling of Multiple Enzymatic Activities. *Angewandte Chemie International Edition* **2012,** *51* (18), 4431-4434.

245. (a) Sachdeva, A.; Wang, K.; Elliott, T.; Chin, J. W., Concerted, rapid, quantitative, and site-specific dual labeling of proteins. *J Am Chem Soc* **2014,** *136* (22), 7785-8; (b) Murrey, H. E.; Judkins, J. C.; Am Ende, C. W.; Ballard, T. E.; Fang, Y.; Riccardi, K.; Di, L.; Guilmette, E. R.; Schwartz, J. W.; Fox, J. M.; Johnson, D. S., Systematic Evaluation of Bioorthogonal Reactions in Live Cells with Clickable HaloTag Ligands: Implications for Intracellular Imaging. *J Am Chem Soc* **2015,** *137* (35), 11461-75.

246. Karver, M. R.; Weissleder, R.; Hilderbrand, S. A., Bioorthogonal Reaction Pairs Enable Simultaneous, Selective, Multi-Target Imaging. *Angewandte Chemie International Edition* **2012,** *51* (4), 920-922.

247. Wang, K.; Sachdeva, A.; Cox, D. J.; Wilf, N. M.; Lang, K.; Wallace, S.; Mehl, R. A.; Chin, J. W., Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nature Chemistry* **2014,** *6*, 393.

248. Narayanam, M. K.; Liang, Y.; Houk, K. N.; Murphy, J. M., Discovery of new mutually orthogonal bioorthogonal cycloaddition pairs through computational screening. *Chemical Science* **2016,** *7* (2), 1257-1261.

249. Wallace, S.; Chin, J. W., Strain-promoted sydnone bicyclo-[6.1.0]-nonyne cycloadditiondaggerElectronic supplementary information (ESI) available: Full experimental details, 1H/13C NMR spectral data, protein synthesis and purification. See DOI: 10.1039/c3sc53332h. *Chem Sci* **2014,** *5* (5), 1742-1744.

250. Rakhit, R.; Navarro, R.; Wandless, T. J., Chemical biology strategies for posttranslational control of protein function. *Chem Biol* **2014,** *21* (9), 1238-52.

251. (a) Li, J.; Yu, J.; Zhao, J.; Wang, J.; Zheng, S.; Lin, S.; Chen, L.; Yang, M.; Jia, S.; Zhang, X.; Chen, P. R., Palladium-triggered deprotection chemistry for protein activation in living cells. *Nat Chem* **2014,** *6* (4), 352-61; (b) Streu, C.; Meggers, E., Ruthenium-induced allylcarbamate cleavage in living cells. *Angew Chem Int Ed Engl* **2006,** *45* (34), 5645-8; (c) Yusop, R. M.; Unciti-Broceta, A.; Johansson, E. M.; Sanchez-Martin, R. M.; Bradley, M., Palladium-mediated intracellular chemistry. *Nat Chem* **2011,** *3* (3), 239-43.

252. Versteegen, R. M.; Rossin, R.; ten Hoeve, W.; Janssen, H. M.; Robillard, M. S., Click to release: instantaneous doxorubicin elimination upon tetrazine ligation. *Angew Chem Int Ed Engl* **2013,** *52* (52), 14112-6.

253. Rossin, R.; van Duijnhoven, S. M.; Ten Hoeve, W.; Janssen, H. M.; Kleijn, L. H.; Hoeben, F. J.; Versteegen, R. M.; Robillard, M. S., Triggered Drug Release from an Antibody-Drug Conjugate Using Fast "Click-to-Release" Chemistry in Mice. *Bioconjug Chem* **2016,** *27* (7), 1697-706.

254. Khan, I.; Agris, P. F.; Yigit, M. V.; Royzen, M., In situ activation of a doxorubicin prodrug using imaging-capable nanoparticles. *Chem Commun (Camb)* **2016,** *52* (36), 6174-7.

255. Zhang, G.; Li, J.; Xie, R.; Fan, X.; Liu, Y.; Zheng, S.; Ge, Y.; Chen, P. R., Bioorthogonal Chemical Activation of Kinases in Living Systems. *ACS Central Science* **2016,** *2* (5), 325-331.

256. (a) Fan, X.; Ge, Y.; Lin, F.; Yang, Y.; Zhang, G.; Ngai, W. S.; Lin, Z.; Zheng, S.; Wang, J.; Zhao, J.; Li, J.; Chen, P. R., Optimized Tetrazine Derivatives for Rapid Bioorthogonal Decaging in Living Cells. *Angew Chem Int Ed Engl* **2016,** *55* (45), 14046-14050; (b) Sarris, A. J. C.; Hansen, T.; de Geus, M. A. R.; Maurits, E.; Doelman, W.; Overkleeft, H. S.; Codée, J. D. C.; Filippov, D. V.; van Kasteren, S. I., Fast and pH-Independent Elimination of trans-Cyclooctene by Using Aminoethyl-Functionalized Tetrazines. *Chemistry – A European Journal* **2018,** *24* (68), 18075-18081; (c) Carlson, J. C. T.; Mikula, H.; Weissleder, R., Unraveling Tetrazine-Triggered Bioorthogonal Elimination Enables Chemical Tools for Ultrafast Release and Universal Cleavage. *Journal of the American Chemical Society* **2018,** *140* (10), 3603-3612.

257. Zhu, J.; Hiltz, J.; Lennox, R. B.; Schirrmacher, R., Chemical modification of single walled carbon nanotubes with tetrazine-tethered gold nanoparticles via a Diels-Alder reaction. *Chemical Communications* **2013,** *49* (87), 10275-10277.

258. Zhu, J.; Hiltz, J.; Mezour, M. A.; Bernard-Gauthier, V.; Lennox, R. B.; Schirrmacher, R., Facile Covalent Modification of a Highly Ordered Pyrolytic Graphite Surface via an Inverse Electron Demand Diels–Alder Reaction under Ambient Conditions. *Chemistry of Materials* **2014,** *26* (17), 5058-5062.

259. Hansell, C. F.; Lu, A.; Patterson, J. P.; O'Reilly, R. K., Exploiting the tetrazine-norbornene reaction for single polymer chain collapse. *Nanoscale* **2014,** *6* (8), 4102-4107.

260. Beckmann, H. S.; Niederwieser, A.; Wiessler, M.; Wittmann, V., Preparation of carbohydrate arrays by using Diels-Alder reactions with inverse electron demand. *Chemistry* **2012,** *18* (21), 6548-54.

261. Roling, O.; Mardyukov, A.; Lamping, S.; Vonhoren, B.; Rinnen, S.; Arlinghaus, H. F.; Studer, A.; Ravoo, B. J., Surface patterning with natural and synthetic polymers via an inverse electron demand Diels-Alder reaction employing microcontact chemistry. *Organic & Biomolecular Chemistry* **2014,** *12* (39), 7828-7835.

262. (a) Liu, S.; Zhang, H.; Remy, R. A.; Deng, F.; Mackay, M. E.; Fox, J. M.; Jia, X., Meter-Long Multiblock Copolymer Microfibers Via Interfacial Bioorthogonal Polymerization. *Advanced Materials* **2015,** *27* (17), 2783-2790; (b) Zhang, H.; Trout, W. S.; Liu, S.; Andrade, G. A.; Hudson, D. A.; Scinto, S. L.; Dicker, K. T.; Li, Y.; Lazouski, N.; Rosenthal, J.; Thorpe, C.; Jia, X.; Fox, J. M., Rapid Bioorthogonal Chemistry Turn-on through Enzymatic or Long Wavelength Photocatalytic Activation of Tetrazine Ligation. *J Am Chem Soc* **2016,** *138* (18), 5978-83.

263. Baumgartner, R.; Song, Z.; Zhang, Y.; Cheng, J., Functional polyesters derived from alternating copolymerization of norbornene anhydride and epoxides. *Polymer Chemistry* **2015,** *6* (19), 3586-3590.

264. Desai, R. M.; Koshy, S. T.; Hilderbrand, S. A.; Mooney, D. J.; Joshi, N. S., Versatile click alginate hydrogels crosslinked via tetrazine-norbornene chemistry. *Biomaterials* **2015,** *50*, 30-7.

265. Murnauer, A. Design einer photoinduzierbaren Diels-Alder Cycloaddition. Master's Thesis, Technische Universität München, 2018.

266. (a) Lim, R. K. V.; Lin, Q., Photoinducible Bioorthogonal Chemistry: A Spatiotemporally Controllable Tool to Visualize and Perturb Proteins in Live Cells. *Accounts of Chemical Research* **2011,** *44* (9), 828-839; (b) Ramil, C. P.; Lin, Q., Photoclick chemistry: a fluorogenic light-triggered in vivo ligation reaction. *Current Opinion in Chemical Biology* **2014,** *21*, 89-95.

267. Clovis, J. S.; Eckell, A.; Huisgen, R.; Sustmann, R., 1.3-Dipolare Cycloadditionen, XXV. Der Nachweis des freien Diphenylnitrilimins als Zwischenstufe bei Cycloadditionen. *Chemische Berichte* **1967,** *100* (1), 60-70.

268. Wang, Y.; Rivera Vera, C. I.; Lin, Q., Convenient Synthesis of Highly Functionalized Pyrazolines via Mild, Photoactivated 1,3-Dipolar Cycloaddition. *Organic Letters* **2007,** *9* (21), 4155-4158.

269. (a) Wang, Y.; Hu, W. J.; Song, W.; Lim, R. K. V.; Lin, Q., Discovery of Long-Wavelength Photoactivatable Diaryltetrazoles for Bioorthogonal 1,3-Dipolar Cycloaddition Reactions. *Organic Letters* **2008,** *10* (17), 3725-3728; (b) Yu, Z.; Ho, L. Y.; Wang, Z.; Lin, Q., Discovery of new photoactivatable diaryltetrazoles for photoclick chemistry via 'scaffold hopping'. *Bioorganic & medicinal chemistry letters* **2011,** *21* (17), 5033-6; (c) Li, F.; Zhang, H.; Sun, Y.; Pan, Y.; Zhou, J.; Wang, J., Expanding the Genetic Code for Photoclick Chemistry in E. coli, Mammalian Cells, and A. thaliana. *Angewandte Chemie International Edition* **2013,** *52* (37), 9700-9704; (d) Song, W.; Wang, Y.; Qu, J.; Lin, Q., Selective Functionalization of a Genetically Encoded Alkene-Containing Protein via "Photoclick Chemistry" in Bacterial Cells. *Journal of the American Chemical Society* **2008,** *130* (30), 9654-9655.

270. Lee, Y.-J.; Wu, B.; Raymond, J. E.; Zeng, Y.; Fang, X.; Wooley, K. L.; Liu, W. R., A Genetically Encoded Acrylamide Functionality. *ACS Chemical Biology* **2013,** *8* (8), 1664-1670.

271. Wang, Y.; Song, W.; Hu, W. J.; Lin, Q., Fast alkene functionalization in vivo by Photoclick chemistry: HOMO lifting of nitrile imine dipoles. *Angew Chem Int Ed Engl* **2009,** *48* (29), 5330-3.

272. Yu, Z.; Lin, Q., Design of Spiro[2.3]hex-1-ene, a Genetically Encodable Double-Strained Alkene for Superfast Photoclick Chemistry. *Journal of the American Chemical Society* **2014,** *136* (11), 4153-4156.

273. Yu, Z.; Pan, Y.; Wang, Z.; Wang, J.; Lin, Q., Genetically encoded cyclopropene directs rapid, photoclick-chemistry-mediated protein labeling in mammalian cells. *Angew Chem Int Ed Engl* **2012,** *51* (42), 10600-4.

274. Yu, Z.; Ohulchanskyy, T. Y.; An, P.; Prasad, P. N.; Lin, Q., Fluorogenic, Two-Photon-Triggered Photoclick Chemistry in Live Mammalian Cells. *Journal of the American Chemical Society* **2013,** *135* (45), 16766-16769.

275. (a) Tian, Y.; Jacinto, M. P.; Zeng, Y.; Yu, Z.; Qu, J.; Liu, W. R.; Lin, Q., Genetically Encoded 2-Aryl-5-carboxytetrazoles for Site-Selective Protein Photo-Cross-Linking. *Journal of the American Chemical Society* **2017,** *139* (17), 6078-6081; (b) Tian, Y.; Lin, Q., Genetic encoding of 2-aryl-5-carboxytetrazole-based protein photo-cross-linkers. *Chemical Communications* **2018,** *54* (35), 4449-4452.

276. (a) Nguyen, L. T.; De Proft, F.; Nguyen, M. T.; Geerlings, P., Theoretical study of cyclopropenones and cyclopropenethiones: decomposition via intermediates. *Journal of the Chemical Society, Perkin Transactions 2* **2001,** (6), 898-905; (b) Sung, K.; Fang, D.-C.; Glenn, D.; T. Tidwell, T., Substituent effects on decarbonylation: theoretical study of the interconversion of 1,2-bisketenes, cyclopropenones and alkynes. *Journal of the Chemical Society, Perkin Transactions 2* **1998,** (9), 2073-2080.

277. Poloukhtine, A.; Popik, V. V., Highly Efficient Photochemical Generation of a Triple Bond: Synthesis, Properties, and Photodecarbonylation of Cyclopropenones. *The Journal of Organic Chemistry* **2003,** *68* (20), 7833-7840.

278. Poloukhtine, A. A.; Mbua, N. E.; Wolfert, M. A.; Boons, G.-J.; Popik, V. V., Selective Labeling of Living Cells by a Photo-Triggered Click Reaction. *Journal of the American Chemical Society* **2009,** *131* (43), 15769-15776.

279. McNitt, C. D.; Popik, V. V., Photochemical generation of oxa-dibenzocyclooctyne (ODIBO) for metal-free click ligations. *Organic & Biomolecular Chemistry* **2012,** *10* (41), 8200-8202.

280. Nainar, S.; Kubota, M.; McNitt, C.; Tran, C.; Popik, V. V.; Spitale, R. C., Temporal Labeling of Nascent RNA Using Photoclick Chemistry in Live Cells. *Journal of the American Chemical Society* **2017,** *139* (24), 8090-8093.

281. Arumugam, S.; Popik, V. V., Sequential "Click" – "Photo-Click" Cross-Linker for Catalyst-Free Ligation of Azide-Tagged Substrates. *The Journal of Organic Chemistry* **2014,** *79* (6), 2702-2708.

282. Kii, I.; Shiraishi, A.; Hiramatsu, T.; Matsushita, T.; Uekusa, H.; Yoshida, S.; Yamamoto, M.; Kudo, A.; Hagiwara, M.; Hosoya, T., Strain-promoted double-click reaction for chemical modification of azido-biomolecules. *Organic & Biomolecular Chemistry* **2010,** *8* (18), 4051-4055.

283. Sutton, D. A.; Yu, S.-H.; Steet, R.; Popik, V. V., Cyclopropenone-caged Sondheimer diyne (dibenzo[a,e]cyclooctadiyne): a photoactivatable linchpin for efficient SPAAC crosslinking. *Chemical Communications* **2016,** *52* (3), 553-556.

284. Sutton, D. A.; Popik, V. V., Sequential Photochemistry of Dibenzo[a,e]dicyclopropa[c,g][8]annulene-1,6-dione: Selective Formation of Didehydrodibenzo[a,e][8]annulenes with Ultrafast SPAAC Reactivity. *The Journal of Organic Chemistry* **2016,** *81* (19), 8850-8857.

285. Friscourt, F.; Fahrni, C. J.; Boons, G.-J., A Fluorogenic Probe for the Catalyst-Free Detection of Azide-Tagged Molecules. *Journal of the American Chemical Society* **2012,** *134* (45), 18809-18815.

286. Martínek, M.; Filipová, L.; Galeta, J.; Ludvíková, L.; Klán, P., Photochemical Formation of Dibenzosilacyclohept-4-yne for Cu-Free Click Chemistry with Azides and 1,2,4,5-Tetrazines. *Organic Letters* **2016,** *18* (19), 4892-4895.

287. Wu, H.; Yang, J.; Šečkutė, J.; Devaraj, N. K., In Situ Synthesis of Alkenyl Tetrazines for Highly Fluorogenic Bioorthogonal Live-Cell Imaging Probes. *Angewandte Chemie International Edition* **2014,** *53* (23), 5805-5809.

288. Qu, Y.; Sauvage, F.-X.; Clavier, G.; Miomandre, F.; Audebert, P., Metal-Free Synthetic Approach to 3-Monosubstituted Unsymmetrical 1,2,4,5-Tetrazines Useful for Bioorthogonal Reactions. *Angewandte Chemie* **2018,** *130* (37), 12233-12237.

289. Luo, X.; Fu, G.; Wang, R. E.; Zhu, X.; Zambaldo, C.; Liu, R.; Liu, T.; Lyu, X.; Du, J.; Xuan, W.; Yao, A.; Reed, S. A.; Kang, M.; Zhang, Y.; Guo, H.; Huang, C.; Yang, P.-Y.; Wilson, I. A.; Schultz, P. G.; Wang, F., Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nature Chemical Biology* **2017,** *13*, 845.

290. (a) Lang, K.; Chin, J. W., Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chemical reviews* **2014,** *114* (9), 4764-806; (b) Liu, C. C.; Schultz, P. G., Adding New Chemistries to the Genetic Code. *Annu Rev Biochem* **2010,** *79*, 413-444.

291. (a) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G., Designing logical codon reassignment - Expanding the chemistry in biology. *Chemical science* **2015,** *6* (1), 50-69; (b) Wan, W.; Tharp, J. M.; Liu, W. R., Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochimica et biophysica acta* **2014,** *1844* (6), 1059-70.

292. Kavran, J. M.; Gundllapalli, S.; O'Donoghue, P.; Englert, M.; Soll, D.; Steitz, T. A., Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (27), 11268-73.

293. Kumar, P.; Laughlin, S. T., Modular activatable bioorthogonal reagents. *Methods in enzymology* **2019,** *622*, 153-182.

294. Poloukhtine, A. A.; Mbua, N. E.; Wolfert, M. A.; Boons, G. J.; Popik, V. V., Selective labeling of living cells by a photo-triggered click reaction. *Journal of the American Chemical Society* **2009,** *131* (43), 15769-76.

295. Friscourt, F.; Fahrni, C. J.; Boons, G. J., A fluorogenic probe for the catalyst-free detection of azide-tagged molecules. *Journal of the American Chemical Society* **2012,** *134* (45), 18809-15.

296. Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G. J., Visualizing metabolically labeled glycoconjugates of living cells by copper-free and fast huisgen cycloadditions. *Angewandte Chemie* **2008,** *47* (12), 2253-5.

297. Martinek, M.; Filipova, L.; Galeta, J.; Ludvikova, L.; Klan, P., Photochemical Formation of Dibenzosilacyclohept-4-yne for Cu-Free Click Chemistry with Azides and 1,2,4,5-Tetrazines. *Organic letters* **2016,** *18* (19), 4892-4895.

298. (a) McNitt, C. D.; Popik, V. V., Photochemical generation of oxa-dibenzocyclooctyne (ODIBO) for metal-free click ligations. *Organic & biomolecular chemistry* **2012,** *10* (41), 8200-2; (b) Bjerknes, M.; Cheng, H.; McNitt, C. D.; Popik, V. V., Facile Quenching and Spatial Patterning of Cylooctynes via Strain-Promoted Alkyne-Azide Cycloaddition of Inorganic Azides. *Bioconjugate chemistry* **2017,** *28* (5), 1560-1565; (c) Nainar, S.; Kubota, M.; McNitt, C.; Tran, C.; Popik, V. V.; Spitale, R. C., Temporal Labeling of Nascent RNA Using Photoclick Chemistry in Live Cells. *Journal of the American Chemical Society* **2017,** *139* (24), 8090-8093.

299. (a) Kumar, P.; Jiang, T.; Li, S.; Zainul, O.; Laughlin, S. T., Caged cyclopropenes for controlling bioorthogonal reactivity. *Organic & biomolecular chemistry* **2018,** *16* (22), 4081-4085; (b) Kumar, P.; Zainul, O.; Camarda, F. M.; Jiang, T.; Mannone, J. A.; Huang, W.; Laughlin, S. T., Caged Cyclopropenes with Improved Tetrazine Ligation Kinetics. *Organic letters* **2019,** *21* (10), 3721-3725.

300. (a) Gordon, C. G.; Mackey, J. L.; Jewett, J. C.; Sletten, E. M.; Houk, K. N.; Bertozzi, C. R., Reactivity of biarylazacyclooctynones in copper-free click chemistry. *Journal of the American Chemical Society* **2012,** *134* (22), 9199-208; (b) Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R., Rapid Cu-free click chemistry with readily synthesized biarylazacyclooctynones. *Journal of the American Chemical Society* **2010,** *132* (11), 3688-90; (c) Debets, M. F.; van Berkel, S. S.; Schoffelen, S.; Rutjes, F. P.; van Hest, J. C.; van Delft, F.

L., Aza-dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition. *Chemical communications* **2010,** *46* (1), 97-9.

301. (a) Arumugam, S.; Popik, V. V., Sequential "click" - "photo-click" cross-linker for catalyst-free ligation of azide-tagged substrates. *The Journal of organic chemistry* **2014,** *79* (6), 2702-8; (b) Sutton, D. A.; Yu, S. H.; Steet, R.; Popik, V. V., Cyclopropenone-caged Sondheimer diyne (dibenzo[a,e]cyclooctadiyne): a photoactivatable linchpin for efficient SPAAC crosslinking. *Chemical communications* **2016,** *52* (3), 553-6; (c) Sutton, D. A.; Popik, V. V., Sequential Photochemistry of Dibenzo[a,e]dicyclopropa[c,g][8]annulene-1,6-dione: Selective Formation of Didehydrodibenzo[a,e][8]annulenes with Ultrafast SPAAC Reactivity. *The Journal of organic chemistry* **2016,** *81* (19), 8850-8857.

302. Orski, S. V.; Poloukhtine, A. A.; Arumugam, S.; Mao, L.; Popik, V. V.; Locklin, J., High density orthogonal surface immobilization via photoactivated copper-free click chemistry. *Journal of the American Chemical Society* **2010,** *132* (32), 11024-6.

303. Greiss, S.; Chin, J. W., Expanding the genetic code of an animal. *Journal of the American Chemical Society* **2011,** *133* (36), 14196-9.

304. Liu, J.; Hemphill, J.; Samanta, S.; Tsang, M.; Deiters, A., Genetic Code Expansion in Zebrafish Embryos and Its Application to Optical Control of Cell Signaling. *Journal of the American Chemical Society* **2017,** *139* (27), 9100-9103.

305. Bianco, A.; Townsley, F. M.; Greiss, S.; Lang, K.; Chin, J. W., Expanding the genetic code of Drosophila melanogaster. *Nature chemical biology* **2012,** *8* (9), 748-50.

306. McNitt, C. D.; Cheng, H.; Ullrich, S.; Popik, V. V.; Bjerknes, M., Multiphoton Activation of Photo-Strain-Promoted Azide Alkyne Cycloaddition "Click" Reagents Enables in Situ Labeling with Submicrometer Resolution. *Journal of the American Chemical Society* **2017,** *139* (40), 14029-14032.

307. Collins, C. J.; Fisher, G. B.; Reem, A.; Goralski, C. T.; Singaram, B., Aminoborohydrides. 9. Selective reductions of aldehydes, ketones, esters, and epoxides in the presence of a nitrile using Lithium N,N-dialkylaminoborohydrides. *Tetrahedron Letters* **1997,** *38* (4), 529-532.

308. Wu, H.; Yang, J.; Šečkutė, J.; Devaraj, N. K., In Situ Synthesis of Alkenyl Tetrazines for Highly Fluorogenic Bioorthogonal Live-Cell Imaging Probes. *Angewandte Chemie* **2014,** *126* (23), 5915-5919.

309. Qu, Y.; Sauvage, F.-X.; Clavier, G.; Miomandre, F.; Audebert, P., Metal-Free Synthetic Approach to 3-Monosubstituted Unsymmetrical 1,2,4,5-Tetrazines Useful for Bioorthogonal Reactions. *Angewandte Chemie International Edition* **2018,** *57* (37), 12057-12061.

310. Kugele, A.; Silkenath, B.; Langer, J.; Wittmann, V.; Drescher, M., Protein Spin Labeling with a Photocaged Nitroxide Using Diels-Alder Chemistry. *ChemBioChem* **2019,** *0* (ja).

311. (a) Svoboda, P.; Stein, P.; Hayashi, H.; Schultz, R. M., Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* **2000,** *127* (19), 4147-56; (b) van der Krol, A. R.; Mur, L. A.; de Lange, P.; Mol, J. N.; Stuitje, A. R., Inhibition of flower pigmentation by antisense CHS genes: promoter and minimal sequence requirements

for the antisense effect. *Plant Mol Biol* **1990,** *14* (4), 457-66; (c) Ketting, R. F.; Plasterk, R. H., A genetic link between co-suppression and RNA interference in C. elegans. *Nature* **2000,** *404* (6775), 296-8; (d) Elbashir, S. M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T., Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **2001,** *411* (6836), 494-8; (e) Ratcliff, F.; Harrison, B. D.; Baulcombe, D. C., A similarity between viral defense and gene silencing in plants. *Science* **1997,** *276* (5318), 1558-60.

312. Lu, M.; Zhang, Q.; Deng, M.; Miao, J.; Guo, Y.; Gao, W.; Cui, Q., An Analysis of Human MicroRNA and Disease Associations. *PLOS ONE* **2008,** *3* (10), e3420.

313. (a) Carthew, R. W.; Sontheimer, E. J., Origins and Mechanisms of miRNAs and siRNAs. *Cell* **2009,** *136* (4), 642-655; (b) Wilson, R. C.; Doudna, J. A., Molecular Mechanisms of RNA Interference. *Annual Review of Biophysics* **2013,** *42* (1), 217-239.

314. Tants, J.-N.; Fesser, S.; Kern, T.; Stehle, R.; Geerlof, A.; Wunderlich, C.; Juen, M.; Hartlmüller, C.; Böttcher, R.; Kunzelmann, S.; Lange, O.; Kreutz, C.; Förstemann, K.; Sattler, M., Molecular basis for asymmetry sensing of siRNAs by the Drosophila Loqs-PD/Dcr-2 complex in RNA interference. *Nucleic Acids Research* **2017,** *45* (21), 12536-12550.

315. Klug, C. S.; Feix, J. B., Methods and Applications of Site-Directed Spin Labeling EPR Spectroscopy. In *Methods in Cell Biology*, Academic Press: 2008; Vol. 84, pp 617-658.

316. Klare, J. P., Site-directed spin labeling EPR spectroscopy in protein research. *Biol Chem* **2013,** *394* (10), 1281-1300.

317. Gillespie, J. R.; Shortle, D., Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. paramagnetic relaxation enhancement by nitroxide spin labels11Edited by P. E. Wright. *Journal of Molecular Biology* **1997,** *268* (1), 158-169.

318. Braun, T.; Drescher, M.; Summerer, D., Expanding the Genetic Code for Site-Directed Spin-Labeling. *International Journal of Molecular Sciences* **2019,** *20* (2).

319. Fielding, J. A.; Concilio, G. M.; Heaven, G.; Hollas, A. M., New Developments in Spin Labels for Pulsed Dipolar EPR. *Molecules* **2014,** *19* (10).

320. (a) Kugele, A.; Braun, T. S.; Widder, P.; Williams, L.; Schmidt, M. J.; Summerer, D.; Drescher, M., Site-directed spin labelling of proteins by Suzuki–Miyaura coupling via a genetically encoded aryliodide amino acid. *Chemical Communications* **2019,** *55* (13), 1923-1926; (b) Widder, P.; Berner, F.; Summerer, D.; Drescher, M., Double Nitroxide Labeling by Copper-Catalyzed Azide–Alkyne Cycloadditions with Noncanonical Amino Acids for Electron Paramagnetic Resonance Spectroscopy. *ACS Chemical Biology* **2019,** *14* (5), 839-844.

321. Molander, G. A.; Katona, B. W.; Machrouhi, F., Development of the suzuki-miyaura cross-coupling reaction: use of air-stable potassium alkynyltrifluoroborates in aryl alkynylations. *J Org Chem* **2002,** *67* (24), 8416-23.

322. Devaraj, N. K.; Karver, M. R.; Hilderbrand, S.; Weissleder, R. Functionalized 1,2,4,5-tetrazine compounds for use in bioorthogonal coupling reactions. WO Patent 2014/065860 Al. May 1st, 2014.

323. Chen, G.; Shigenari, T.; Jain, P.; Zhang, Z.; Jin, Z.; He, J.; Li, S.; Mapelli, C.; Miller, M. M.; Poss, M. A.; Scola, P. M.; Yeung, K. S.; Yu, J. Q., Ligand-enabled beta-C-H arylation of alpha-amino acids using a simple and practical auxiliary. *J Am Chem Soc* **2015,** *137* (9), 3338-51.

324. Tran, L. D.; Daugulis, O., Nonnatural amino acid synthesis by using carbon-hydrogen bond functionalization methodology. *Angew Chem Int Ed Engl* **2012,** *51* (21), 5188-91.

325. Wieczorek, A.; Buckup, T.; Wombacher, R., Rigid tetrazine fluorophore conjugates with fluorogenic properties in the inverse electron demand Diels-Alder reaction. *Org Biomol Chem* **2014,** *12* (24), 4177-85.

326. Shabashov, D.; Daugulis, O., Auxiliary-Assisted Palladium-Catalyzed Arylation and Alkylation of sp2 and sp3 Carbon−Hydrogen Bonds. *Journal of the American Chemical Society* **2010,** *132* (11), 3965-3972.

327. Negishi, E.-i., Magical Power of Transition Metals: Past, Present, and Future (Nobel Lecture). *Angewandte Chemie International Edition* **2011,** *50* (30), 6738-6764.

328. Sase, S.; Jaric, M.; Metzger, A.; Malakhov, V.; Knochel, P., One-Pot Negishi Cross-Coupling Reactions of In Situ Generated Zinc Reagents with Aryl Chlorides, Bromides, and Triflates. *The Journal of Organic Chemistry* **2008,** *73* (18), 7380-7382.

329. (a) Jackson, R. F. W.; Moore, R. J.; Dexter, C. S.; Elliott, J.; Mowbray, C. E., Concise Synthesis of Enantiomerically Pure Phenylalanine, Homophenylalanine, and Bishomophenylalanine Derivatives Using Organozinc Chemistry: NMR Studies of Amino Acid-Derived Organozinc Reagents. *The Journal of Organic Chemistry* **1998,** *63* (22), 7875-7884; (b) Thierry, J.; Yue, C.; Potier, P., 2-Phenyl isopropyl and t-butyl trichloroacetimidates: Useful reagents for ester preparation of N-protected amino acids under neutral conditions. *Tetrahedron Letters* **1998,** *39* (12), 1557-1560.

330. Pinner, A.; Klein, F., Umwandlung der Nitrile in Imide. *Berichte der deutschen chemischen Gesellschaft* **1877,** *10* (2), 1889-1897.

331. Yadav, Veejendra K.; Babu, K. G., A Remarkably Efficient Markovnikov Hydrochlorination of Olefins and Transformation of Nitriles into Imidates by Use of AcCl and an Alcohol. *European Journal of Organic Chemistry* **2005,** *2005* (2), 452-456.

332. Horner, K. A.; Valette, N. M.; Webb, M. E., Strain-Promoted Reaction of 1,2,4-Triazines with Bicyclononynes. *Chemistry – A European Journal* **2015,** *21* (41), 14376-14381.

333. (a) Doyle, M. P.; Siegfried, B.; Dellaria, J. F., Alkyl nitrite-metal halide deamination reactions. 2. Substitutive deamination of arylamines by alkyl nitrites and copper(II) halides. A direct and remarkably efficient conversion of arylamines to aryl halides. *The Journal of Organic Chemistry* **1977,** *42* (14), 2426-2431; (b) Just, M.; Thalmann, F. Verfahren zur Herstellung von 6-substituierten 3-Alkylthio-1,2,4,5-tetrazinen. DD Patent 000000290011 A5. May 16th,1991; (c) Wainwright, P.; Maddaford, A.; Simms, M.; Forrest, N.; Glen, R.; Hart, J.; Zhang, X.; Pryde, D. C.; Stephenson, P. T.; Middleton, D. S.; Guyot, T.; Sutton, S. C., Synthesis of Novel 2-Cyano-7-deaza-8-azapurine- and 2-Cyano-8-azapurine-Derived Nucleosides. *Synlett* **2011,** *2011* (13), 1900-1904.

334. Melzig, L.; Metzger, A.; Knochel, P., Pd- and Ni-Catalyzed Cross-Coupling Reactions of Functionalized Organozinc Reagents with Unsaturated Thioethers. *Chemistry – A European Journal* **2011,** *17* (10), 2948-2956.

335. Negishi, E.-i.; Hu, Q.; Huang, Z.; Qian, M.; Wang, G., Palladium-Catalyzed Alkenylation by the Negishi Coupling. *Aldrichimica Acta* **2005,** *38* (3), 81-87.

336. Usuki, T.; Yanuma, H.; Hayashi, T.; Yamada, H.; Suzuki, N.; Masuyama, Y., Improved Negishi Cross-Coupling Reactions of an Organozinc Reagent Derived from l-Aspartic Acid with Monohalopyridines. *Journal of Heterocyclic Chemistry* **2014,** *51* (1), 269-273.

337. (a) Farina, V.; Baker, S. R.; Benigni, D. A.; Sapino, C., Palladium-catalyzed coupling between cephalosporin derivatives and unsaturated stannanes: A new ligand for palladium chemistry. *Tetrahedron Letters* **1988,** *29* (45), 5739-5742; (b) Farina, V.; Krishnan, B., Large rate accelerations in the stille reaction with tri-2-furylphosphine and triphenylarsine as palladium ligands: mechanistic and synthetic implications. *Journal of the American Chemical Society* **1991,** *113* (25), 9585-9595.

338. (a) Kavran, J. M.; Gundllapalli, S.; Donoghue, P.; Englert, M.; Söll, D.; Steitz, T. A., Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proceedings of the National Academy of Sciences* **2007,** *104* (27), 11268; (b) Ko, J.-h.; Wang, Y.-S.; Nakamura, A.; Guo, L.-T.; Söll, D.; Umehara, T., Pyrrolysyl-tRNA synthetase variants reveal ancestral aminoacylation function. *FEBS letters* **2013,** *587* (19), 3243-3248; (c) Takimoto, J. K.; Dellas, N.; Noel, J. P.; Wang, L., Stereochemical Basis for Engineered Pyrrolysyl-tRNA Synthetase and the Efficient in Vivo Incorporation of Structurally Divergent Non-native Amino Acids. *ACS Chemical Biology* **2011,** *6* (7), 733-743.

339. (a) Tharp, J. M.; Wang, Y.-S.; Lee, Y.-J.; Yang, Y.; Liu, W. R., Genetic Incorporation of Seven ortho-Substituted Phenylalanine Derivatives. *ACS Chemical Biology* **2014,** *9* (4), 884-890; (b) Tuley, A.; Wang, Y.-S.; Fang, X.; Kurra, Y.; Rezenom, Y. H.; Liu, W. R., The genetic incorporation of thirteen novel non-canonical amino acids. *Chemical Communications* **2014,** *50* (20), 2673-2675; (c) Wang, Y.-S.; Fang, X.; Chen, H.-Y.; Wu, B.; Wang, Z. U.; Hilty, C.; Liu, W. R., Genetic Incorporation of Twelve meta-Substituted Phenylalanine Derivatives Using a Single Pyrrolysyl-tRNA Synthetase Mutant. *ACS Chemical Biology* **2013,** *8* (2), 405-415; (d) Wang, Y.-S.; Fang, X.; Wallace, A. L.; Wu, B.; Liu, W. R., A Rationally Designed Pyrrolysyl-tRNA Synthetase Mutant with a Broad Substrate Spectrum. *Journal of the American Chemical Society* **2012,** *134* (6), 2950-2953; (e) Xiang, Z.; Lacey, V. K.; Ren, H.; Xu, J.; Burban, D. J.; Jennings, P. A.; Wang, L., Proximity-Enabled Protein Crosslinking through Genetically Encoding Haloalkane Unnatural Amino Acids. *Angewandte Chemie International Edition* **2014,** *53* (8), 2190-2193.

340. Lacey, V. K.; Louie, G. V.; Noel, J. P.; Wang, L., Expanding the Library and Substrate Diversity of the Pyrrolysyl-tRNA Synthetase to Incorporate Unnatural Amino Acids Containing Conjugated Rings. *ChemBioChem* **2013,** *14* (16), 2100-2105.

341. Brunner, A. D. Genetic code expansion by PylRS engineering; organic synthesis and incorporation of unnatural amino acids. Technical University Munich, 2017.

342. Zacharias, D. A.; Violin, J. D.; Newton, A. C.; Tsien, R. Y., Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **2002,** *296* (5569), 913-6.

343. Studier, F. W., Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **2005,** *41* (1), 207-34.

344. Gaussmann, S. Analysis of dsRBD domain orientation in Loqs-PD using a novel spin labeling method for PRE experiments. Eberhard Karls Universität Tübingen, 2017.

345. Kniss, A.; Schuetz, D.; Kazemi, S.; Pluska, L.; Spindler, P. E.; Rogov, V. V.; Husnjak, K.; Dikic, I.; Güntert, P.; Sommer, T.; Prisner, T. F.; Dötsch, V., Chain Assembly and Disassembly Processes Differently Affect the Conformational Space of Ubiquitin Chains. *Structure* **2018,** *26* (2), 249-258.e4.

346. Barfoot, C. W.; Harvey, J. E.; Kenworthy, M. N.; Kilburn, J. P.; Ahmed, M.; Taylor, R. J. K., Highly functionalised organolithium and organoboron reagents for the preparation of enantiomerically pure α-amino acids. *Tetrahedron* **2005,** *61* (13), 3403-3417.

347. Wang, P.; Na, Z.; Fu, J.; Tan, C. Y. J.; Zhang, H.; Yao, S. Q.; Sun, H., Microarray immobilization of biomolecules using a fast trans-cyclooctene (TCO)–tetrazine reaction. *Chemical Communications* **2014,** *50* (80), 11818-11821.

348. (a) Gautier, A.; Nguyen, D. P.; Lusic, H.; An, W.; Deiters, A.; Chin, J. W., Genetically Encoded Photocontrol of Protein Localization in Mammalian Cells. *Journal of the American Chemical Society* **2010,** *132* (12), 4086-4088; (b) Nguyen, D. P.; Mahesh, M.; Elsässer, S. J.; Hancock, S. M.; Uttamapinant, C.; Chin, J. W., Genetic Encoding of Photocaged Cysteine Allows Photoactivation of TEV Protease in Live Mammalian Cells. *Journal of the American Chemical Society* **2014,** *136* (6), 2240-2243; (c) Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G., A Genetically Encoded Photocaged Amino Acid. *Journal of the American Chemical Society* **2004,** *126* (44), 14306-14307.

349. (a) Liu, L.; Liu, Y.; Zhang, G.; Ge, Y.; Fan, X.; Lin, F.; Wang, J.; Zheng, H.; Xie, X.; Zeng, X.; Chen, P. R., Genetically Encoded Chemical Decaging in Living Bacteria. *Biochemistry* **2018,** *57* (4), 446-450; (b) Böcker, J. K.; Dörner, W.; Mootz, H. D., Light-control of the ultra-fast Gp41-1 split intein with preserved stability of a genetically encoded photo-caged amino acid in bacterial cells. *Chemical Communications* **2019,** *55* (9), 1287-1290.

350. (a) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W., Genetically encoding Nε-acetyllysine in recombinant proteins. *Nature Chemical Biology* **2008,** *4*, 232; (b) Park, H.-S.; Hohn, M. J.; Umehara, T.; Guo, L.-T.; Osborne, E. M.; Benner, J.; Noren, C. J.; Rinehart, J.; Söll, D., Expanding the Genetic Code of Escherichia coli with Phosphoserine. *Science* **2011,** *333* (6046), 1151.

351. Waldmann, H.; Sebastian, D., Enzymic Protecting Group Techniques. *Chemical Reviews* **1994,** *94* (4), 911-937.

352. (a) Wang, Q. C.; Fei, J.; Cui, D. F.; Zhu, S. G.; Xu, L. G., Application of an immobilized penicillin acylase to the deprotection of N-phenylacetyl insulin. *Biopolymers* **1986,** *25 Suppl*, S109-14; (b) Žáková, L.; Zyka, D.; Ježek, J.; Hančlová, I.; Šanda, M.; Brzozowski, A. M.; Jiráček, J., The use of Fmoc-Lys(Pac)-OH and penicillin G acylase in the preparation of novel semisynthetic insulin analogs. *Journal of Peptide Science* **2007,** *13* (5), 334-341.

353. Burroughs, A. M.; Balaji, S.; Iyer, L. M.; Aravind, L., Small but versatile: the extraordinary functional and structural diversity of the β-grasp fold. *Biology Direct* **2007,** *2* (1), 18.

354. Hoyer, K. M.; Mahlert, C.; Marahiel, M. A., The Iterative Gramicidin S Thioesterase Catalyzes Peptide Ligation and Cyclization. *Chemistry & Biology* **2007,** *14* (1), 13-22.

355. (a) Gulick, A. M., Conformational Dynamics in the Acyl-CoA Synthetases, Adenylation Domains of Non-ribosomal Peptide Synthetases, and Firefly Luciferase. *ACS Chemical Biology* **2009,** *4* (10), 811-827; (b) Yonus, H.; Neumann, P.; Zimmermann, S.; May, J. J.; Marahiel, M. A.; Stubbs, M. T., Crystal structure of DltA. Implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains. *The Journal of biological chemistry* **2008,** *283* (47), 32484-91.

356. Alfermann, J.; Sun, X.; Mayerthaler, F.; Morrell, T. E.; Dehling, E.; Volkmann, G.; Komatsuzaki, T.; Yang, H.; Mootz, H. D., FRET monitoring of a nonribosomal peptide synthetase. *Nature Chemical Biology* **2017,** *13*, 1009.

357. (a) Dehling, E.; Volkmann, G.; Matern, J. C. J.; Dörner, W.; Alfermann, J.; Diecker, J.; Mootz, H. D., Mapping of the Communication-Mediating Interface in Nonribosomal Peptide Synthetases Using a Genetically Encoded Photocrosslinker Supports an Upside-Down Helix-Hand Motif. *Journal of Molecular Biology* **2016,** *428* (21), 4345-4360; (b) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A., Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *The Journal of biological chemistry* **1998,** *273* (35), 22773-81.

358. (a) Choudhary, C.; Weinert, B. T.; Nishida, Y.; Verdin, E.; Mann, M., The growing landscape of lysine acetylation links metabolism and cell signalling. *Nature Reviews Molecular Cell Biology* **2014,** *15*, 536; (b) Seidel, J.; Klockenbusch, C.; Schwarzer, D., Investigating Deformylase and Deacylase Activity of Mammalian and Bacterial Sirtuins. *ChemBioChem* **2016,** *17* (5), 398-402.

359. Srirangan, K.; Orr, V.; Akawi, L.; Westbrook, A.; Moo-Young, M.; Chou, C. P., Biotechnological advances on Penicillin G acylase: Pharmaceutical implications, unique expression mechanism and production strategies. *Biotechnology Advances* **2013,** *31* (8), 1319-1332.

IV. APPENDIX

IV. Appendix

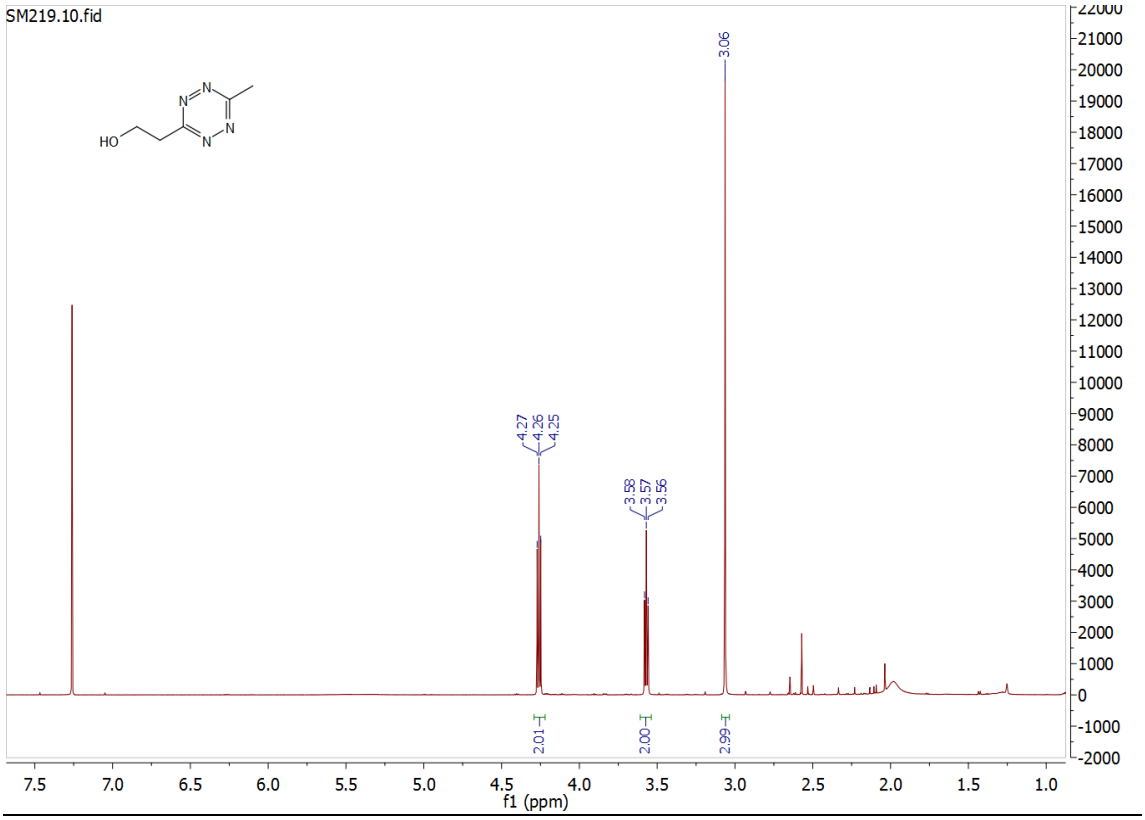
IV.1 NMR spectra

IV.1.1 Chapter 1

IV.1.1.1 TetK

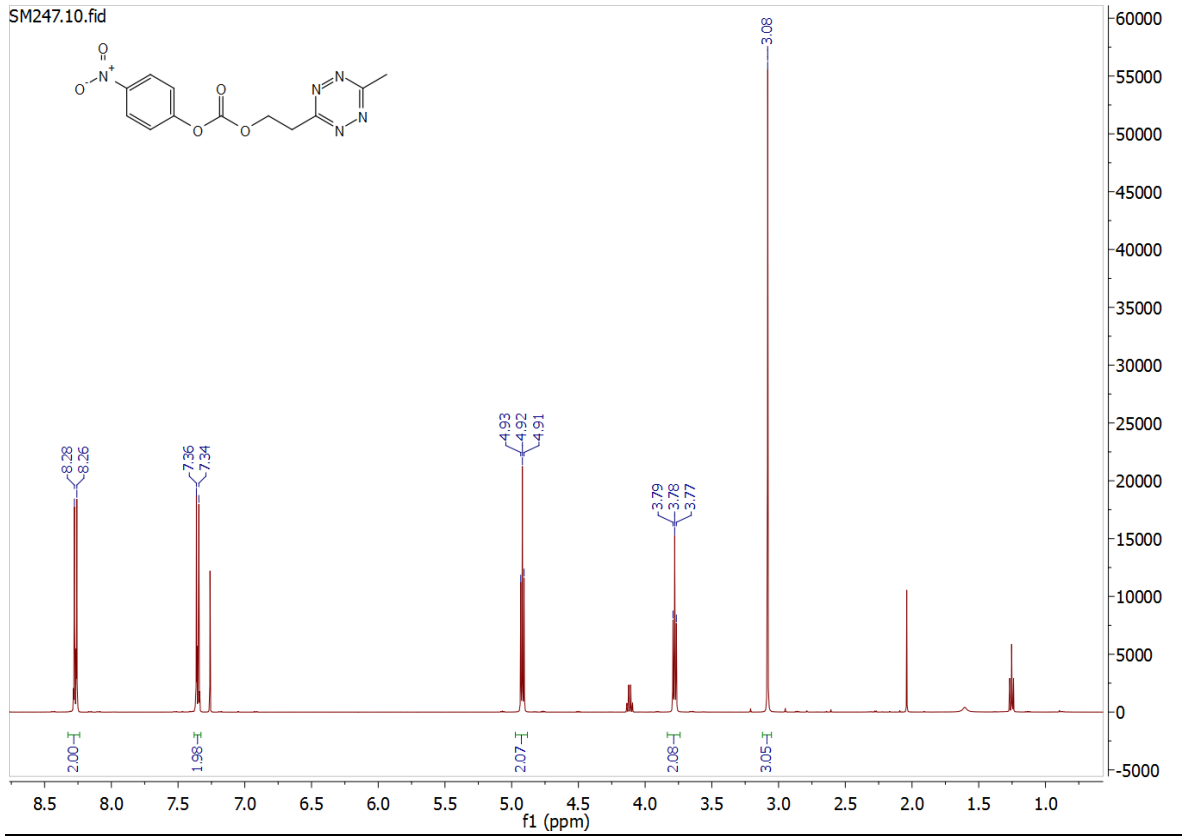
3-(2-Hydroxyethyl)-6-methyl-1,2,4,5-tetrazine

^{1}H:



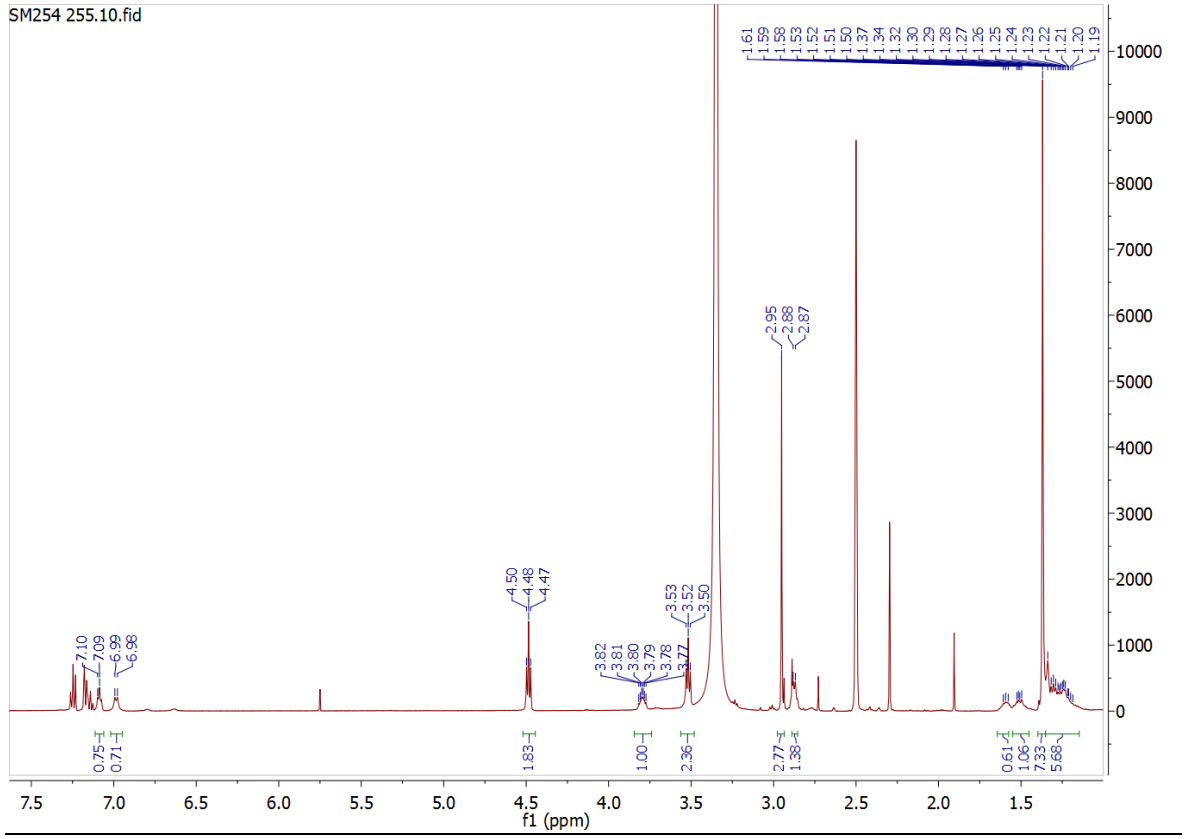
2-(6-methyl-1,2,4,5-tetrazin-3-yl)ethyl-(4-nitrophenyl)-carbonate

^{1}H:



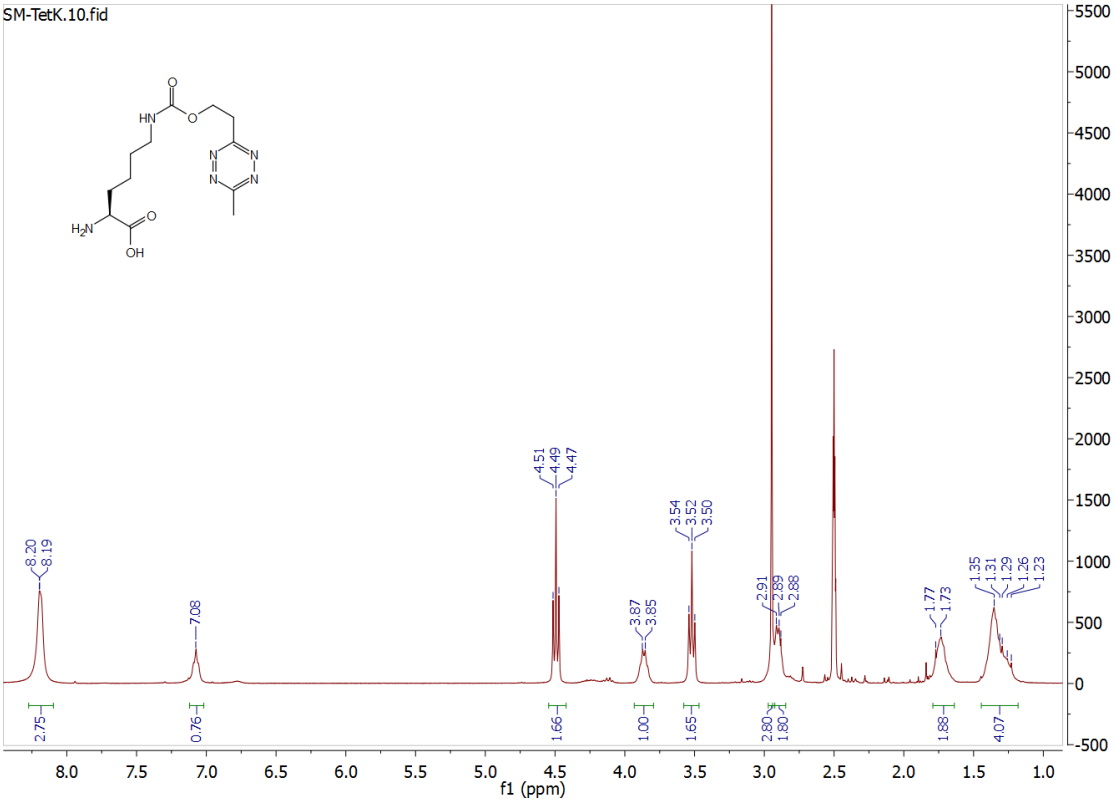
BocTetK

^{1}H:

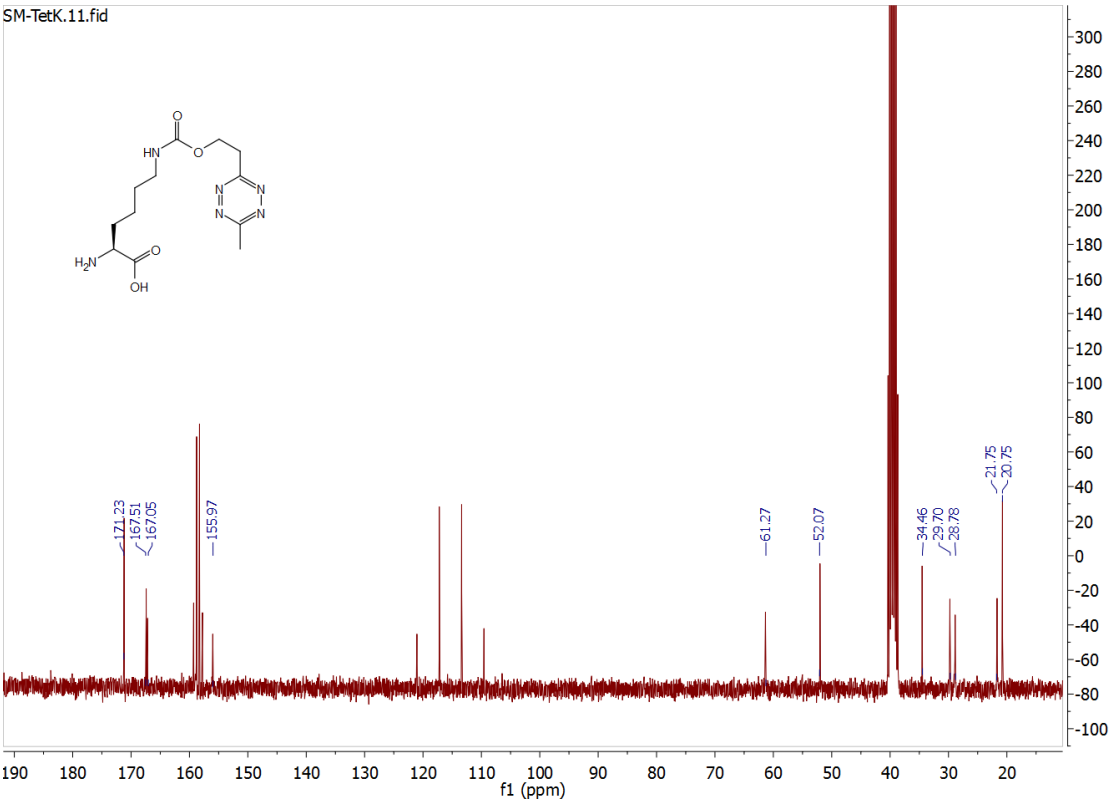


TetK

^{1}H:


SM-TetK.10.fid
8.20
8.19
7.08
4.51
4.49
4.47
3.87
3.85
3.54
3.52
3.50
2.91
2.89
2.88
1.77
1.73
1.35
1.31
1.29
1.26
1.23
2.75
0.76
1.66
1.00
1.65
2.80
1.80
1.88
4.07
f1 (ppm)

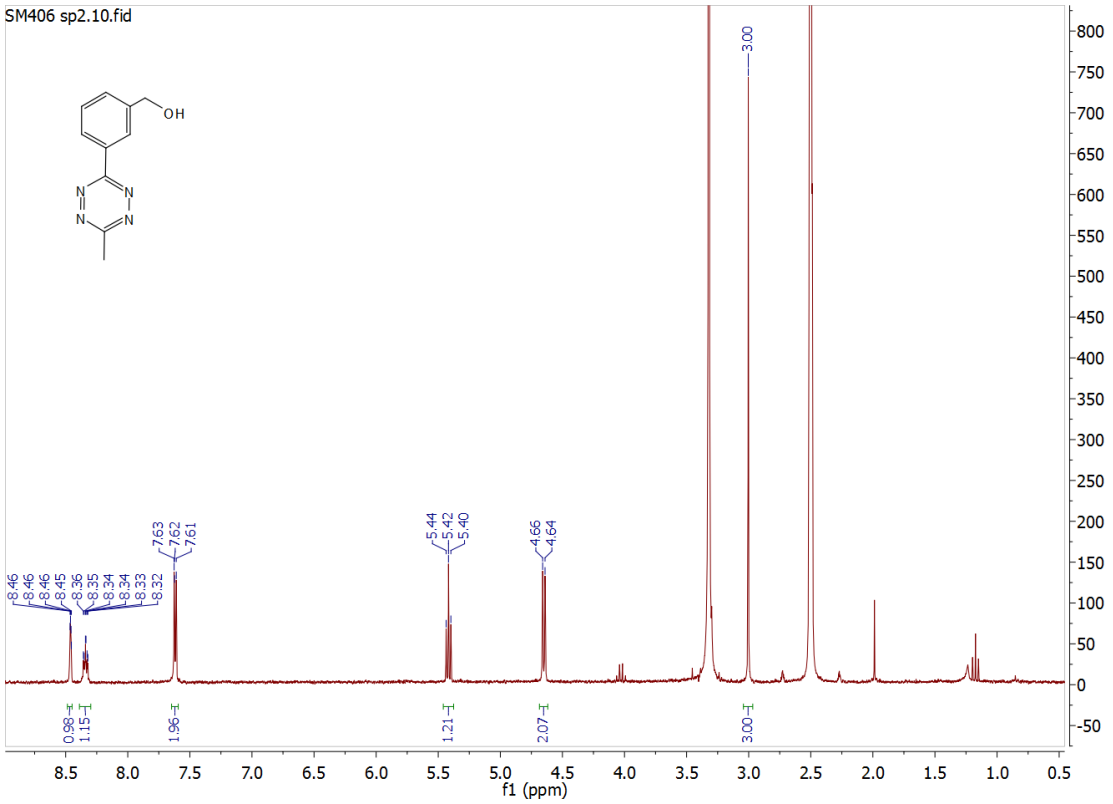
^{13}C:


SM-TetK.11.fid
171.23
167.51
167.05
155.97
61.27
52.07
34.46
29.70
28.78
21.75
20.75
f1 (ppm)

IV.1.1.2 mTetK

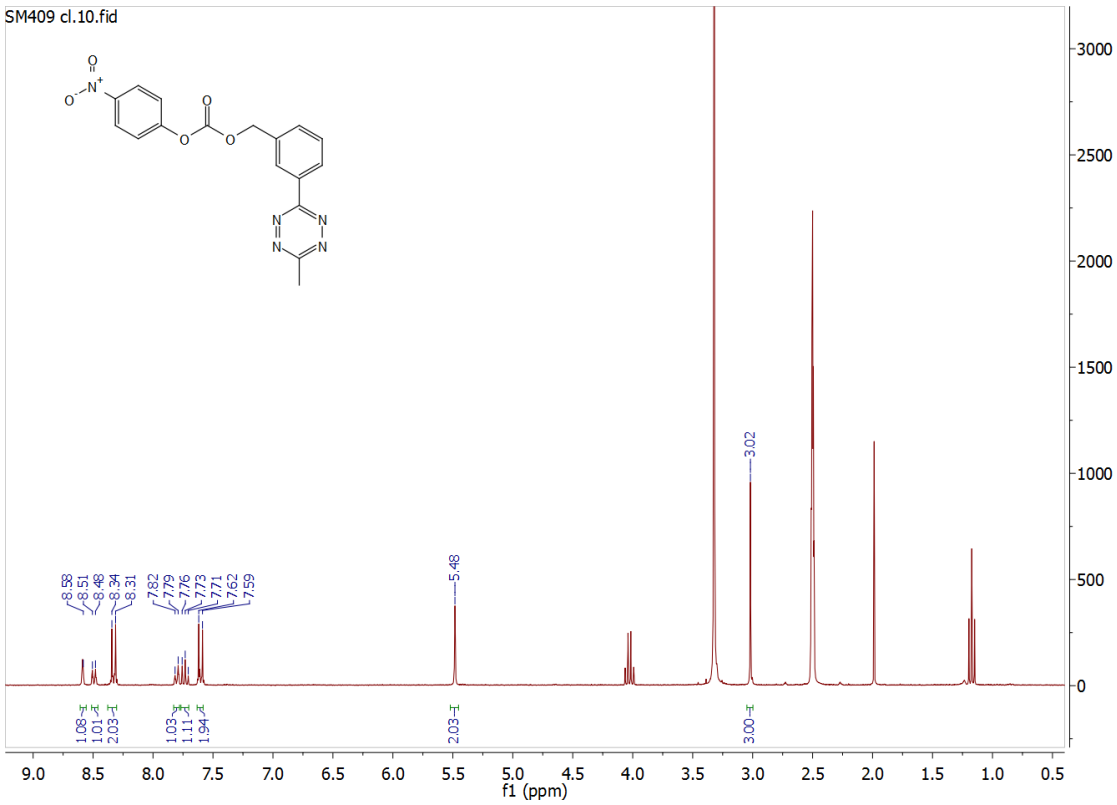
3-(3-(Hydroxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine

^{1}H:



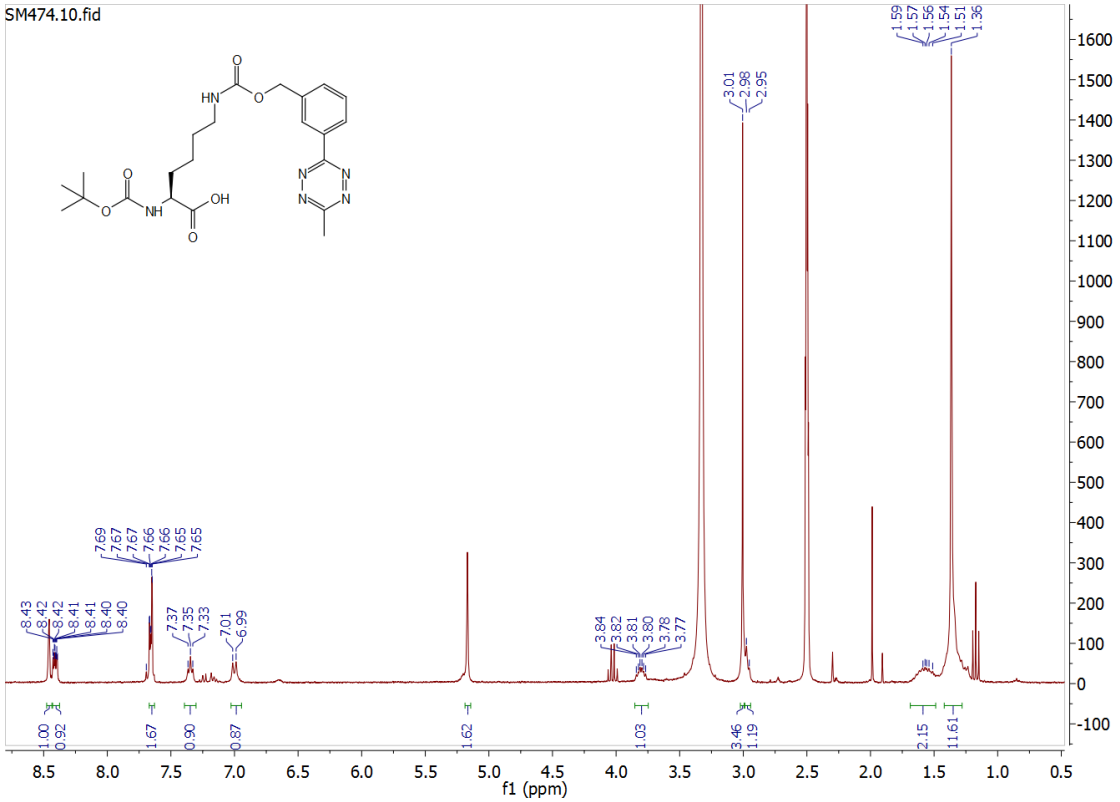
3-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate

^{1}H:



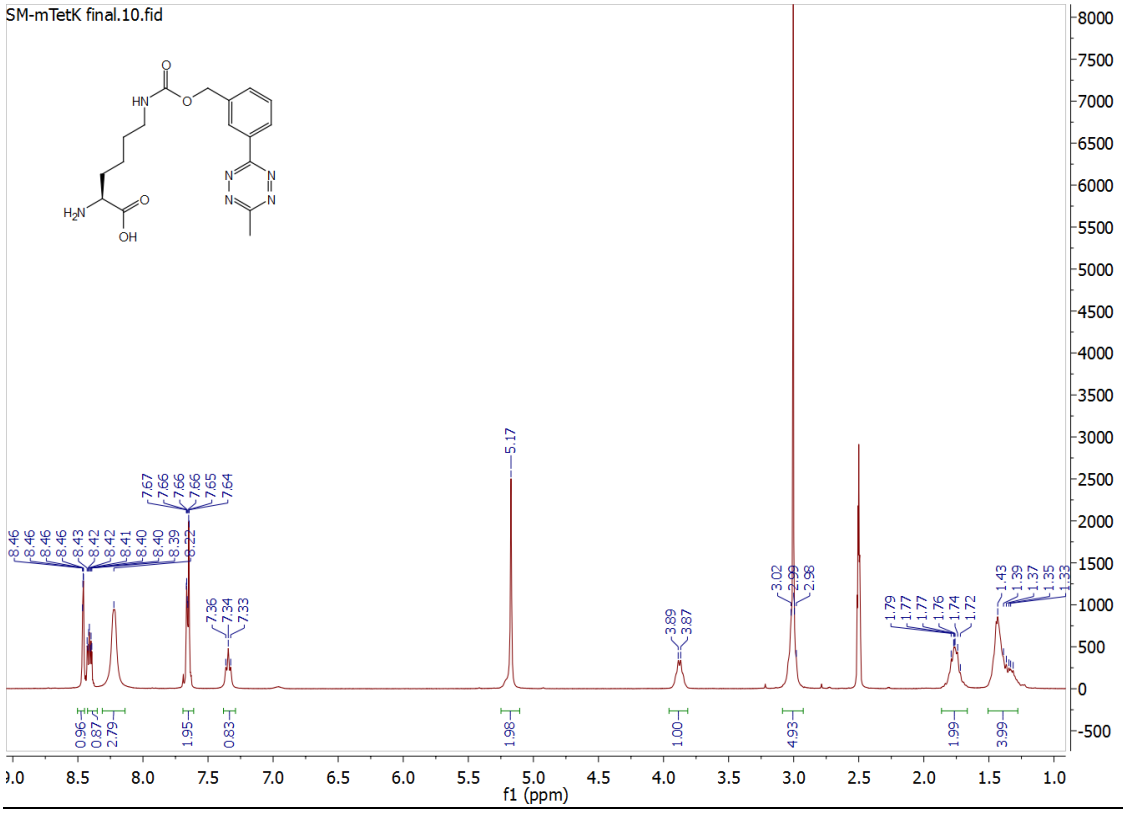
BocmTetK

^{1}H:

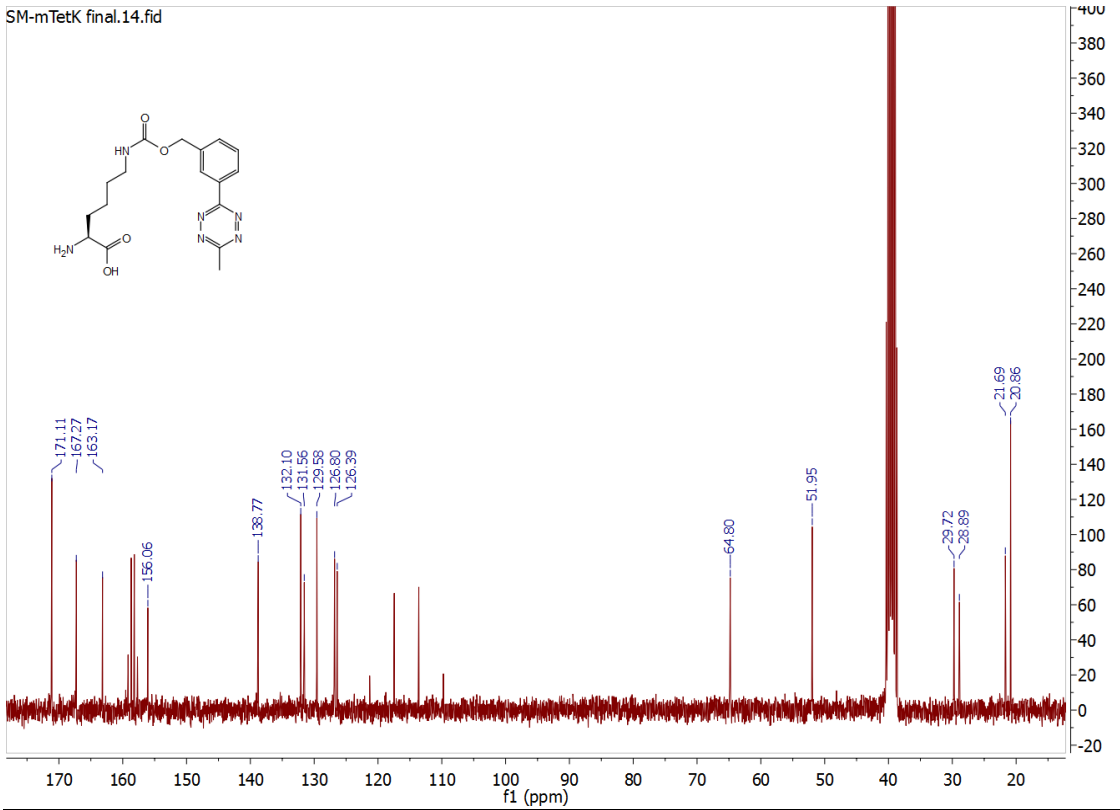


mTetK

^{1}H:

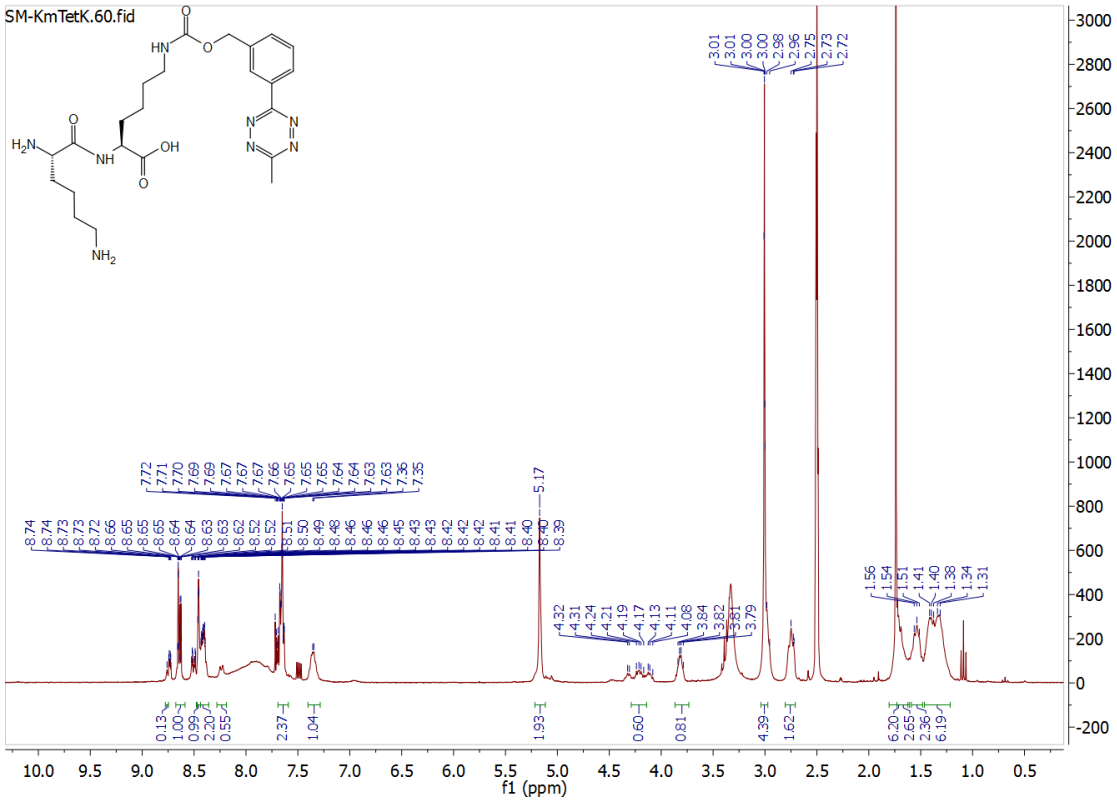


^{13}C:

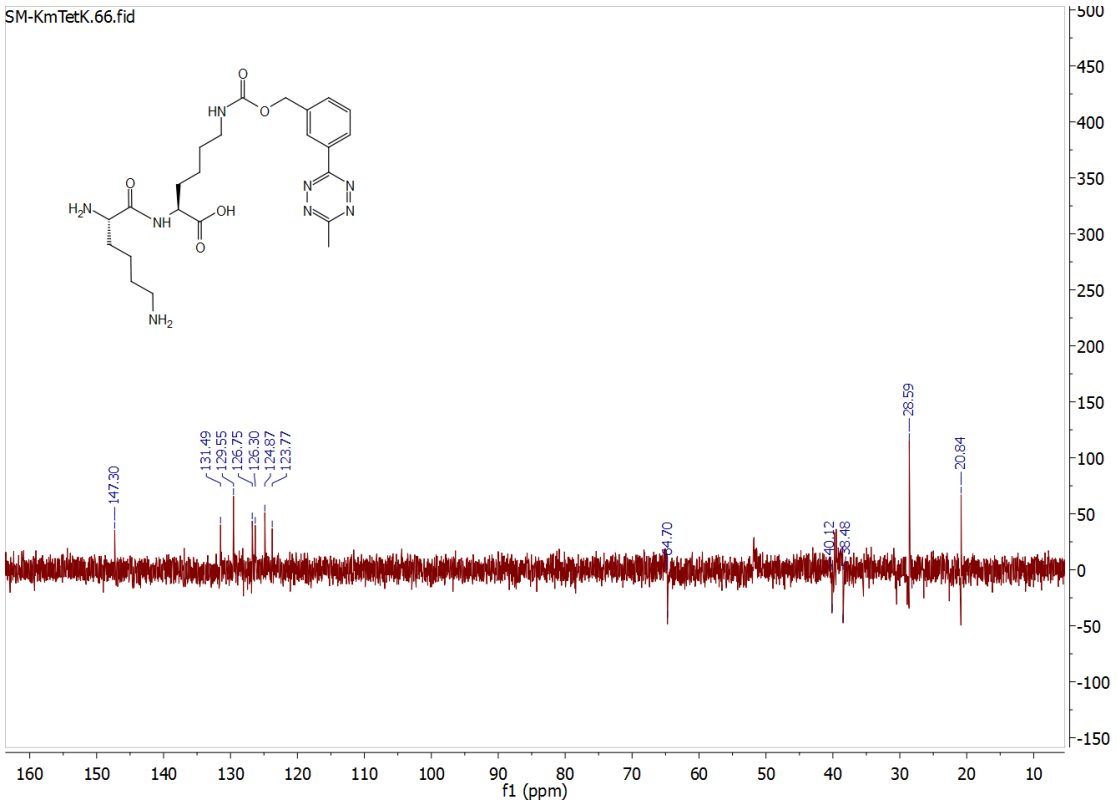


IV.1.1.3 K-mTetK

^{1}H:



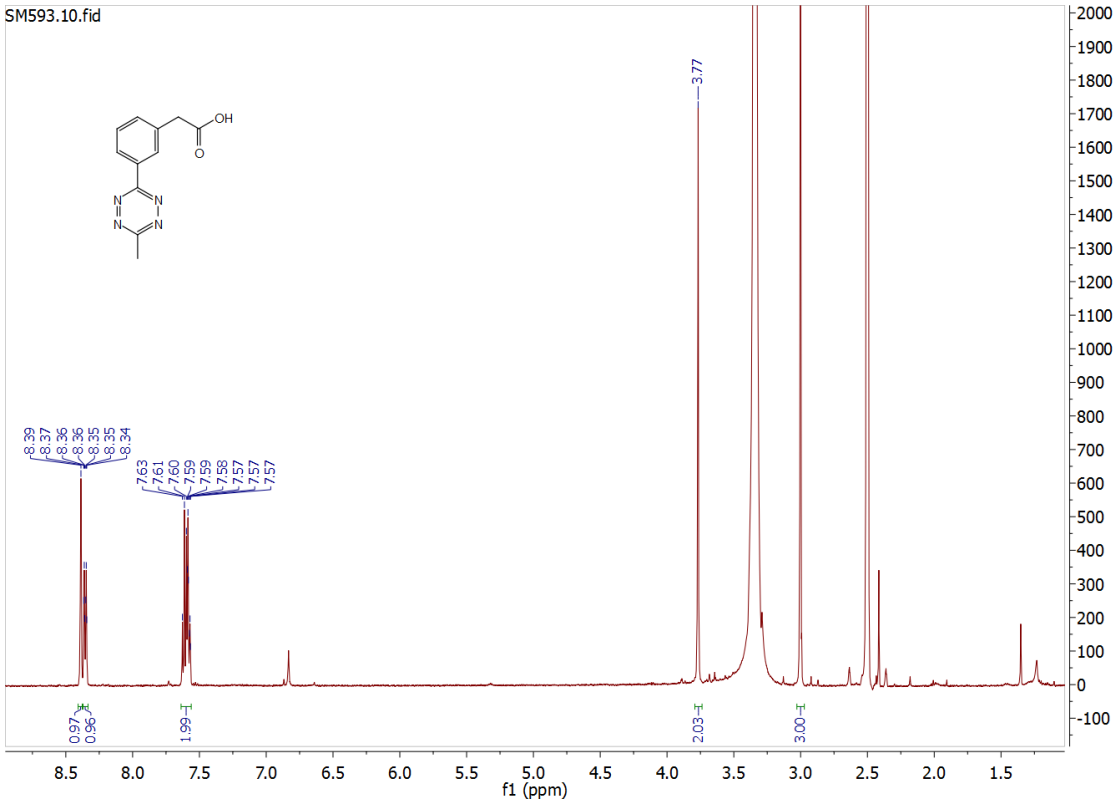
^{13}C - DEPT-135:



IV.1.1.4 mTet2K

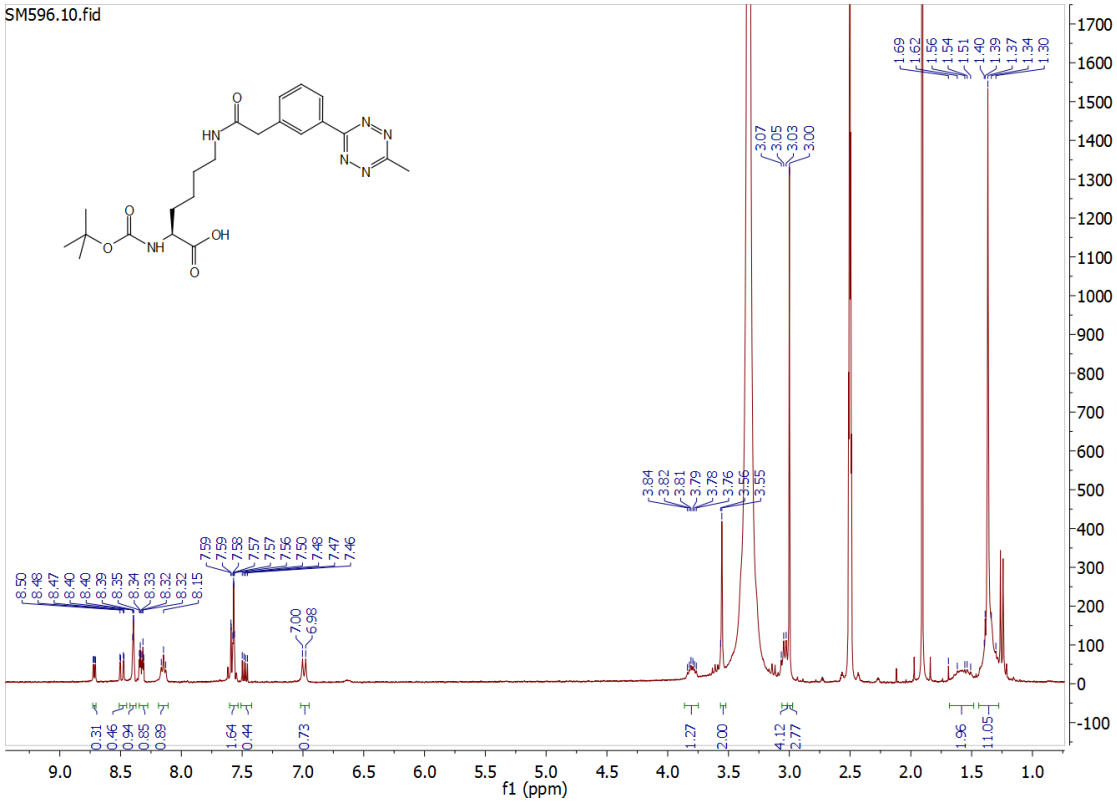
3-(3-(Carboxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine

^{1}H:



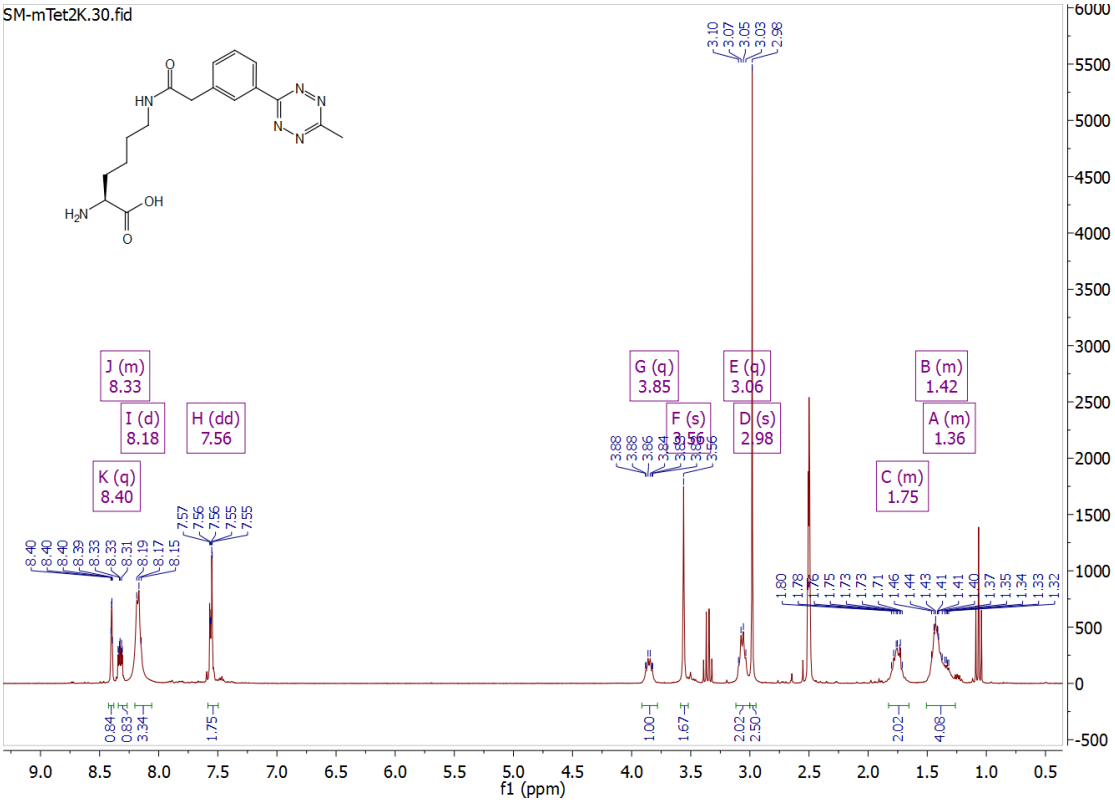
BocmTet2K

^{1}H:

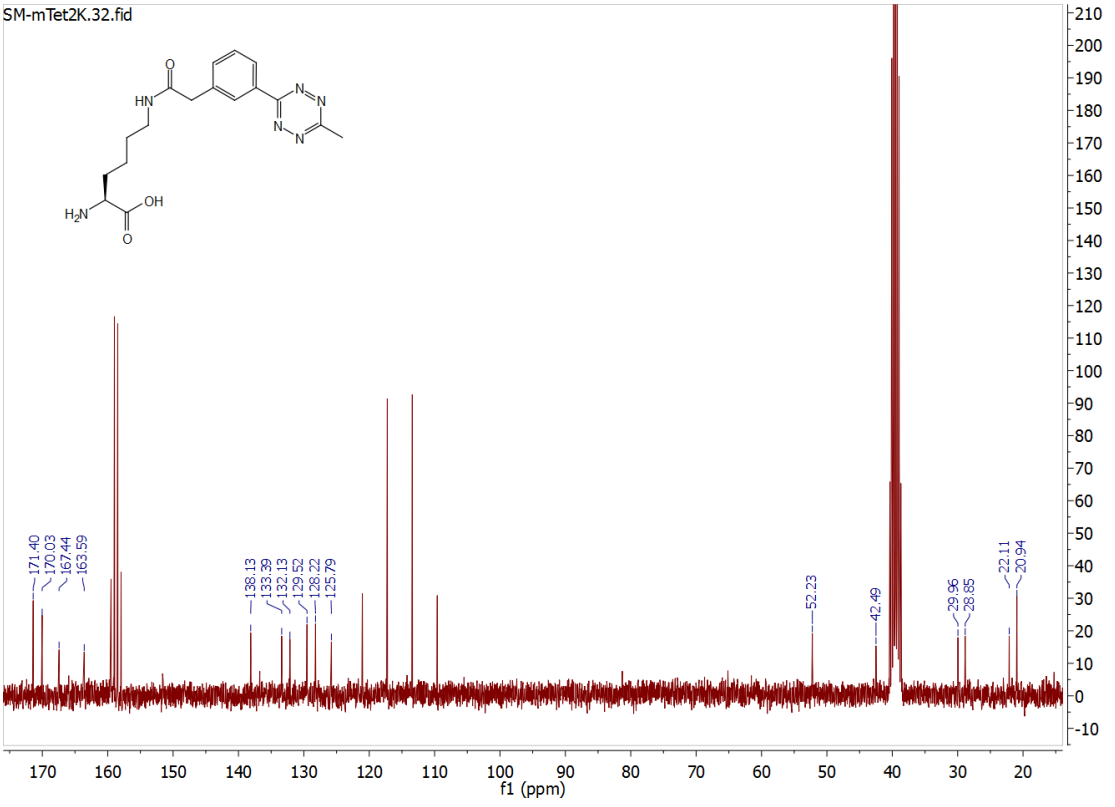


mTet2K

^{1}H:



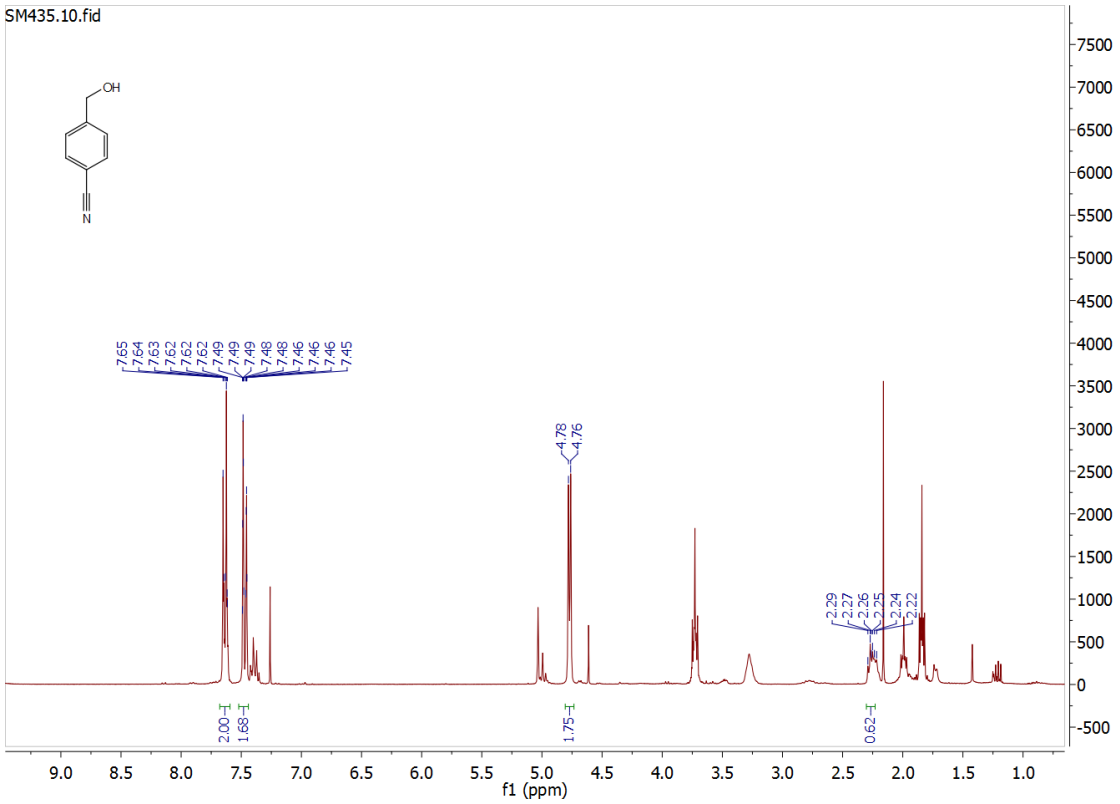
^{13}C:



IV.1.1.5 pTetK

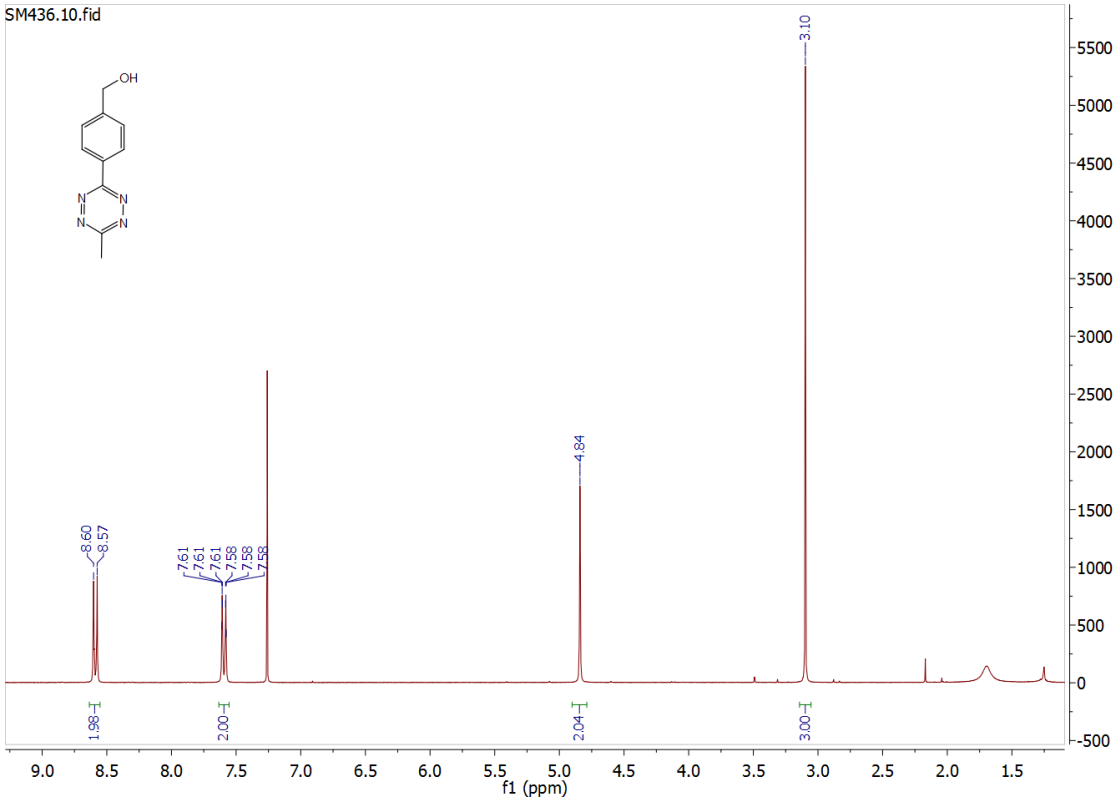
4-(Hydroxymethyl)benzonitrile

^{1}H:



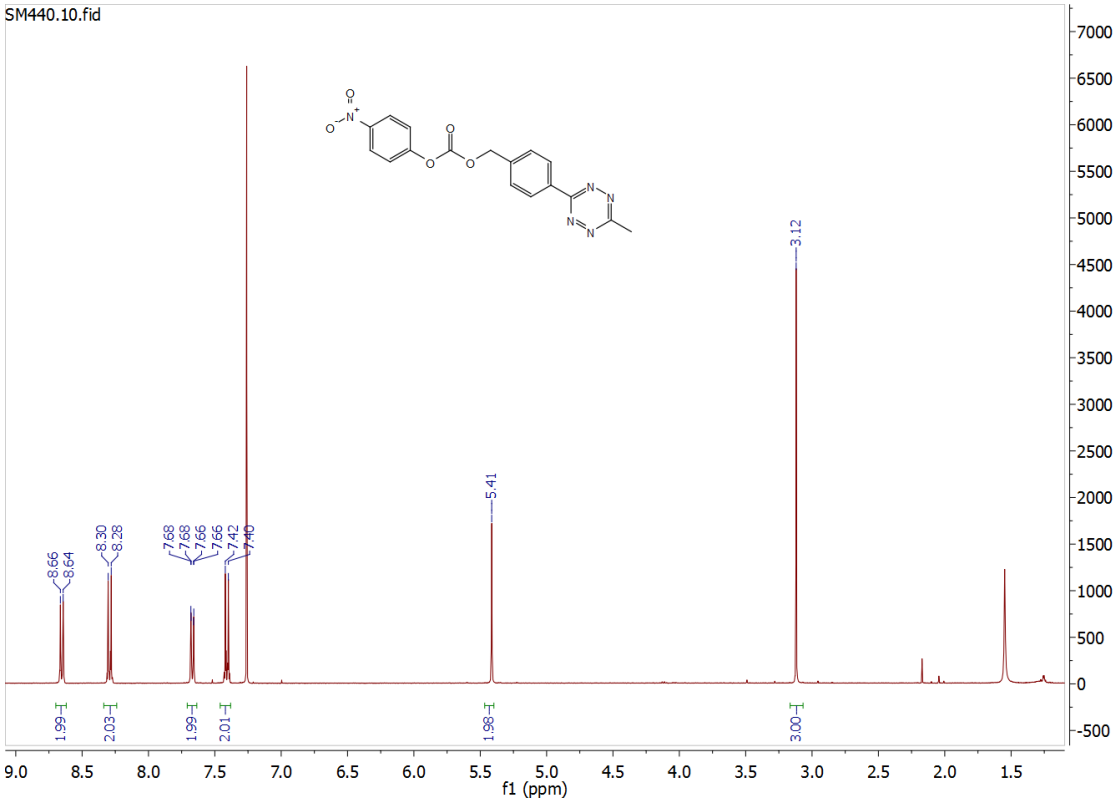
3-(4-(Hydroxymethyl)phenyl)-6-methyl-1,2,4,5-tetrazine

^{1}H:



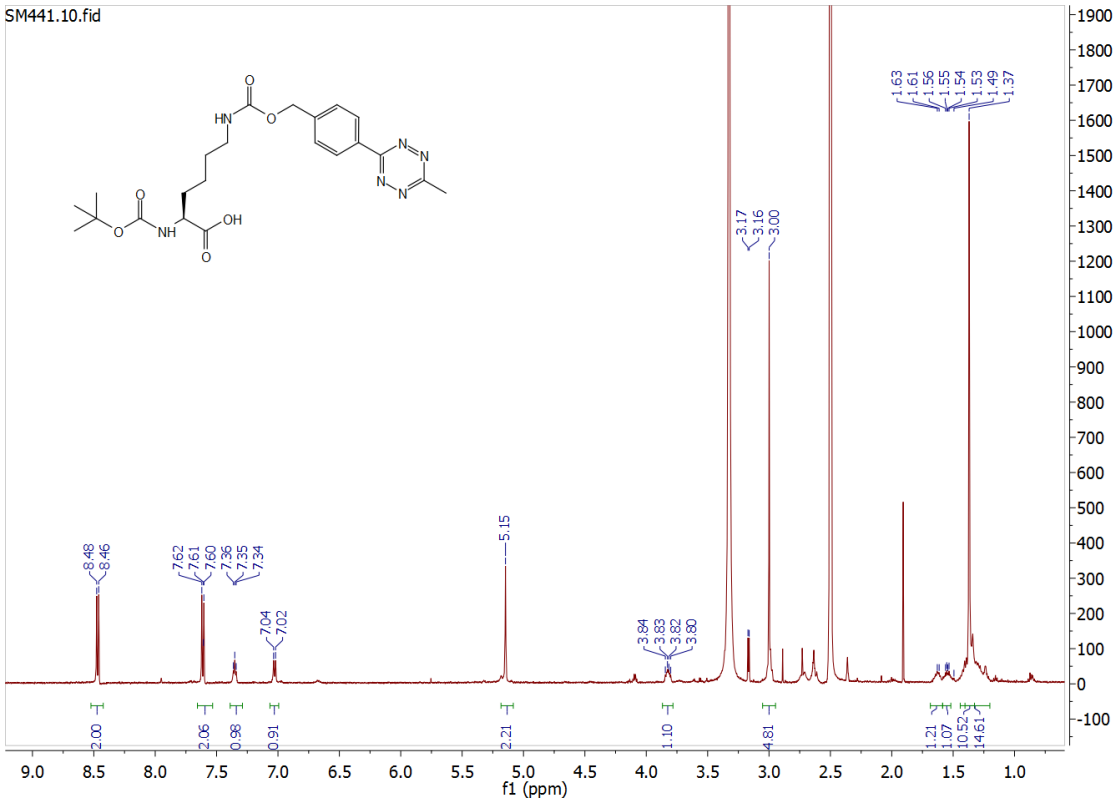
4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate

^{1}H:



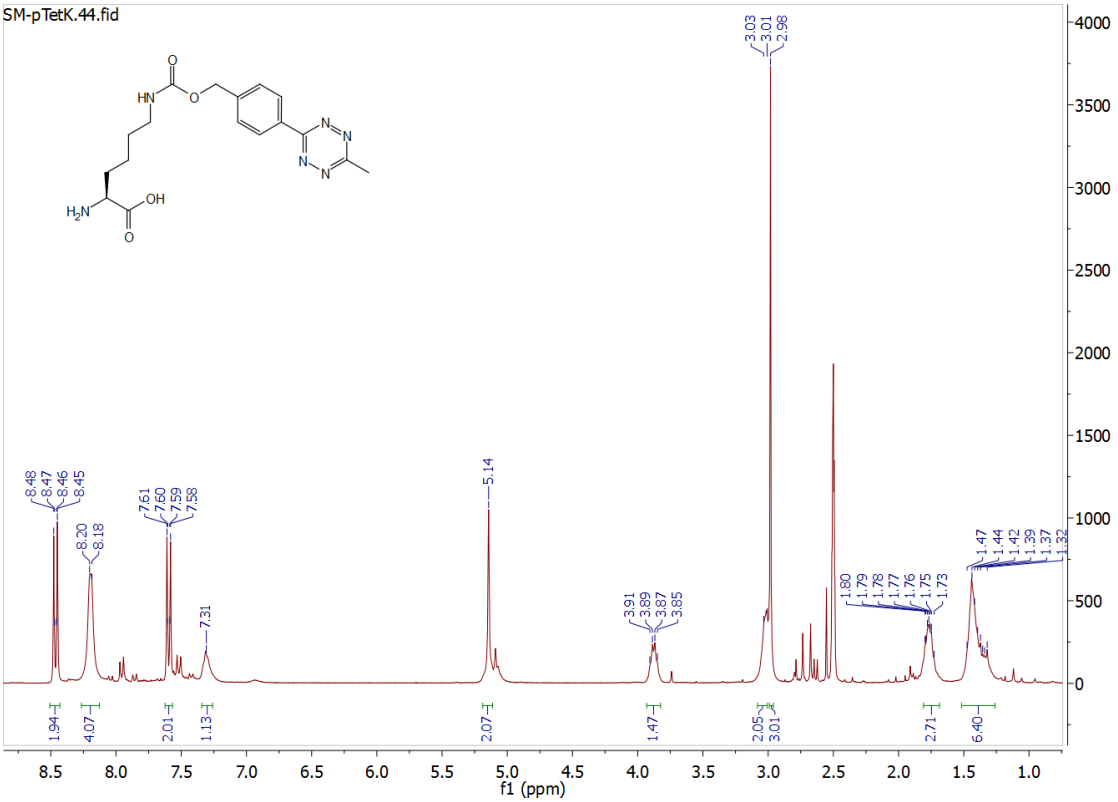
BocpTetK

^{1}H:

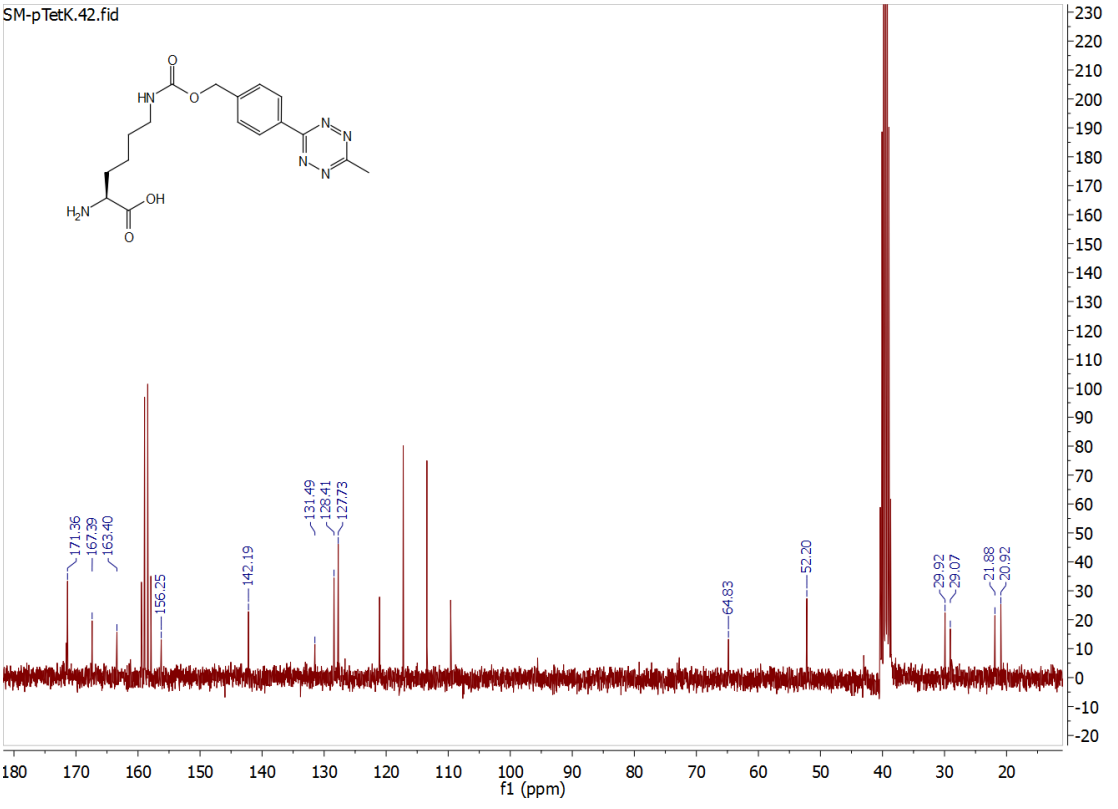


pTetK

^{1}H:

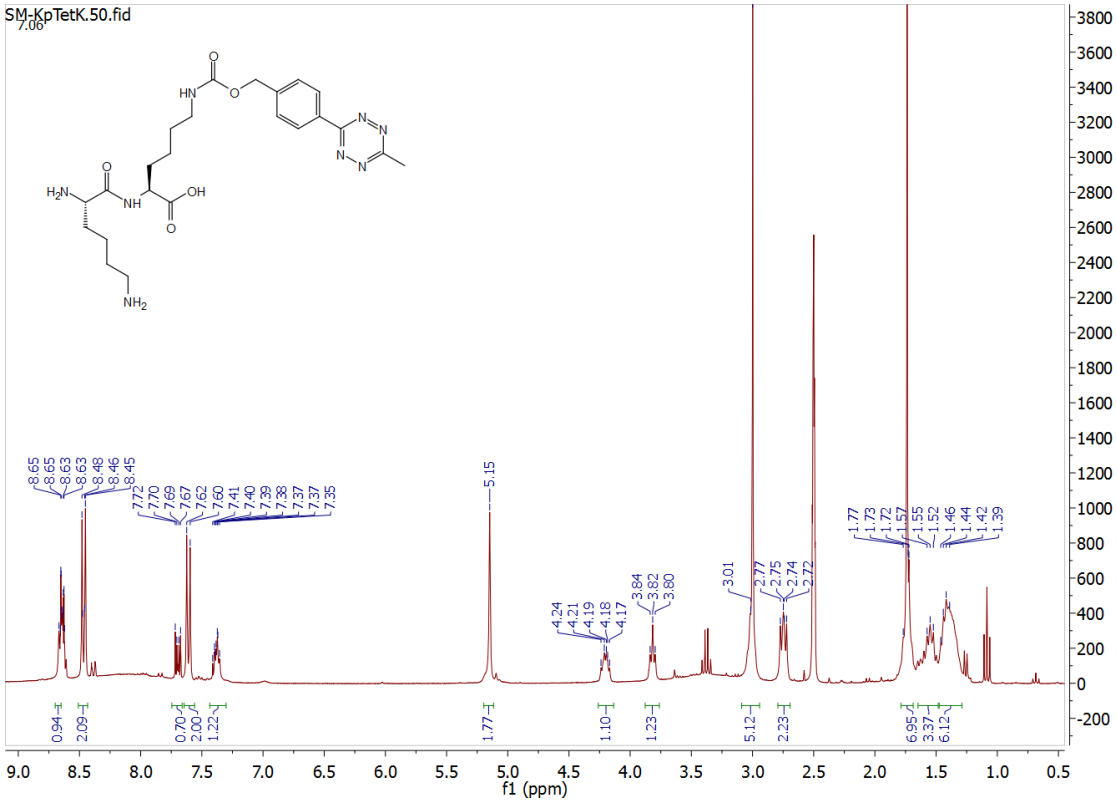

SM-pTetK.44.fid
8.48
8.47
8.46
8.45
8.20
8.18
7.61
7.60
7.59
7.58
7.31
5.14
3.91
3.89
3.87
3.85
3.03
3.01
2.98
1.80
1.79
1.78
1.77
1.76
1.75
1.73
1.47
1.44
1.42
1.39
1.37
1.32
1.94
4.07
2.01
1.13
2.07
1.47
2.05
3.01
2.71
6.40
f1 (ppm)

^{13}C:

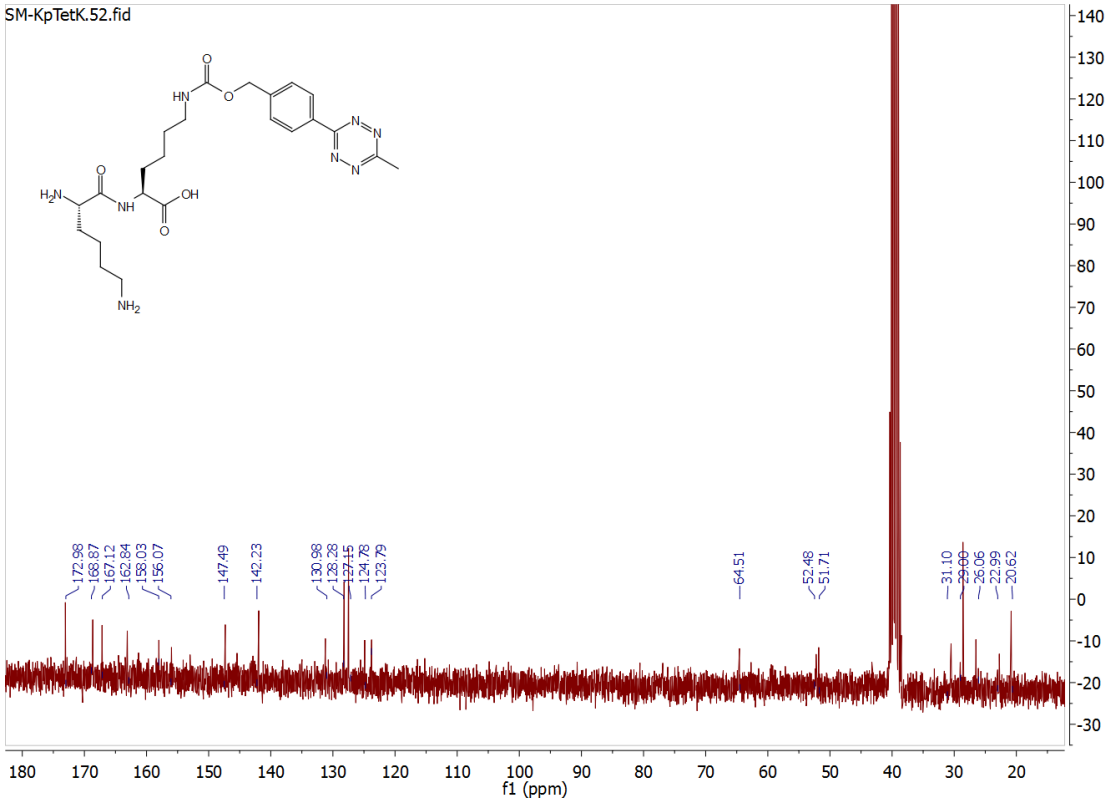

SM-pTetK.42.fid
171.36
167.39
163.40
156.25
142.19
131.49
128.41
127.73
64.83
52.20
29.92
29.07
21.88
20.92
f1 (ppm)

IV.1.1.6 K-pTetK

^{1}H:



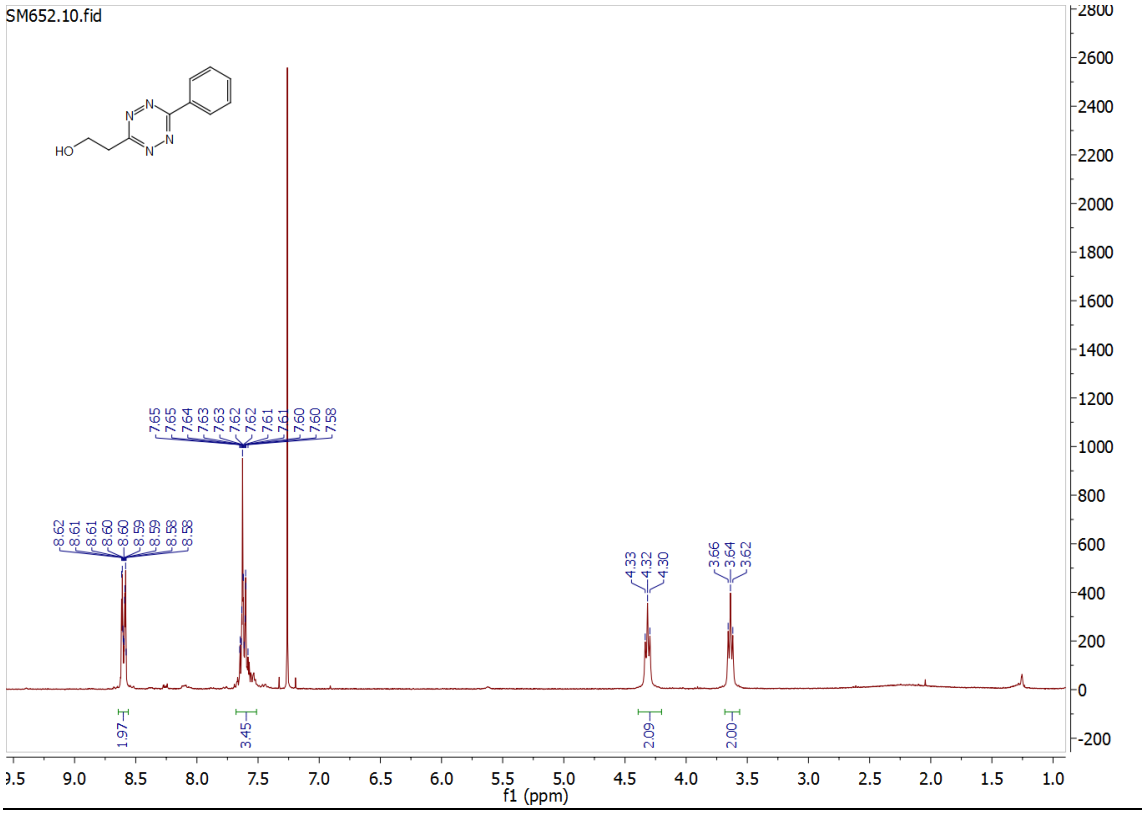
^{13}C:



IV.1.1.7 PhTetK

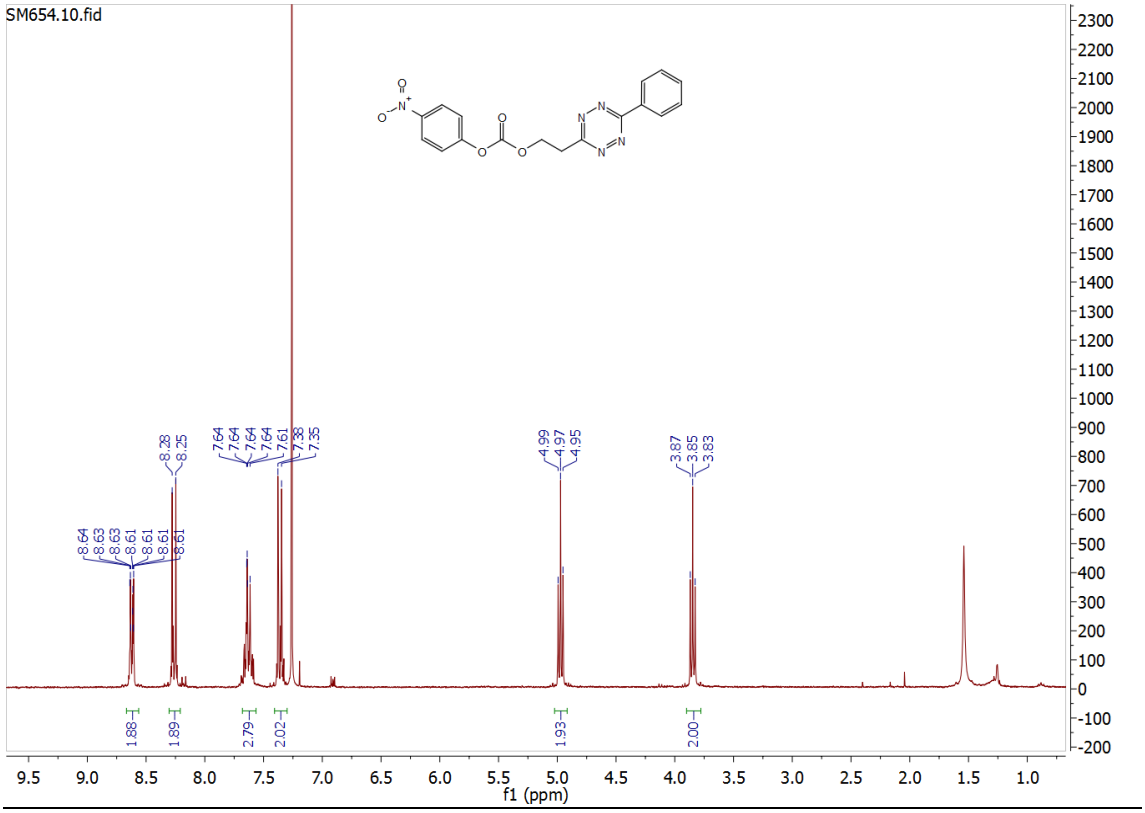
3-(2-Hydroxyethyl)-6-phenyl-1,2,4,5-tetrazine

^{1}H:



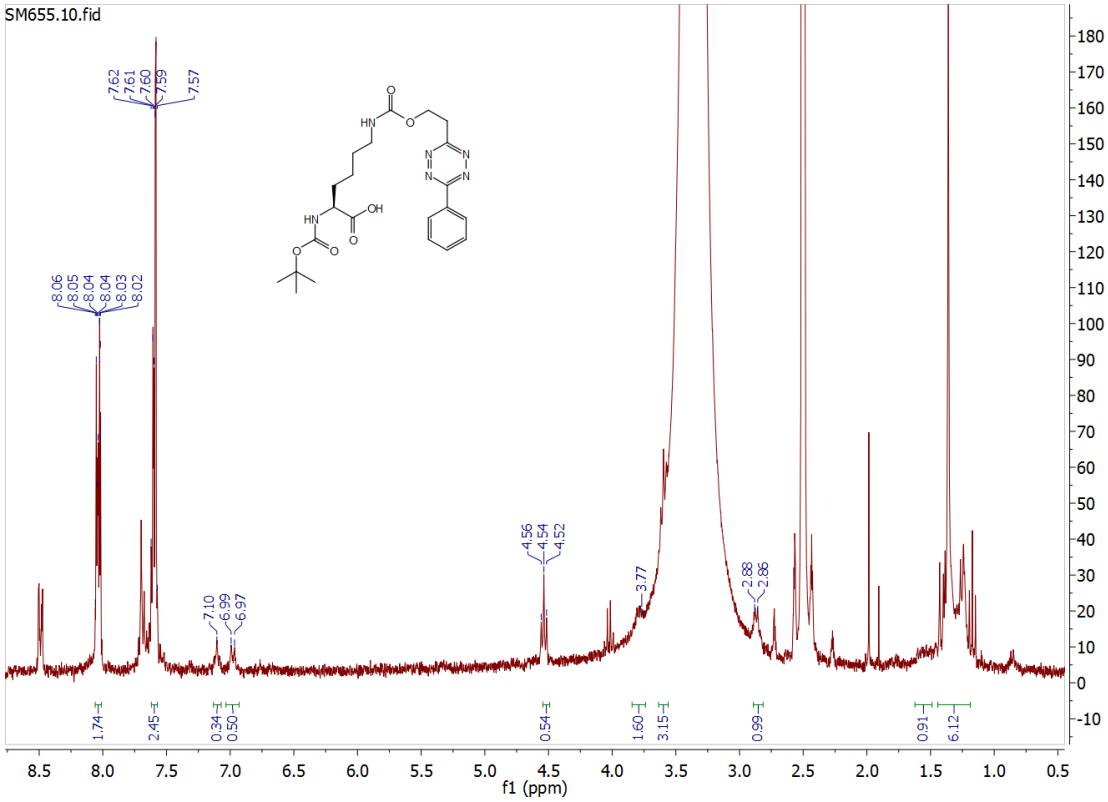
2-(6-phenyl-1,2,4,5-tetrazin-3-yl)ethyl-(4-nitrophenyl)-carbonate

^{1}H:



BocPhTetK

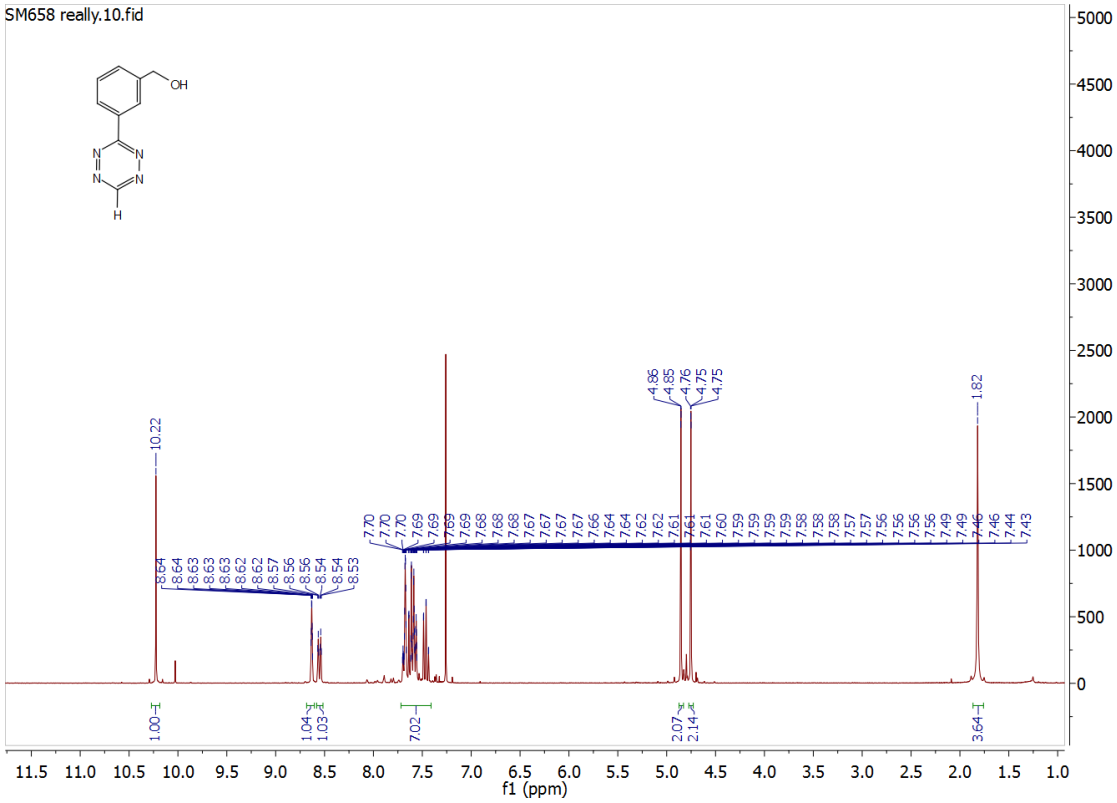
^{1}H:



IV.1.1.8 HmTetK

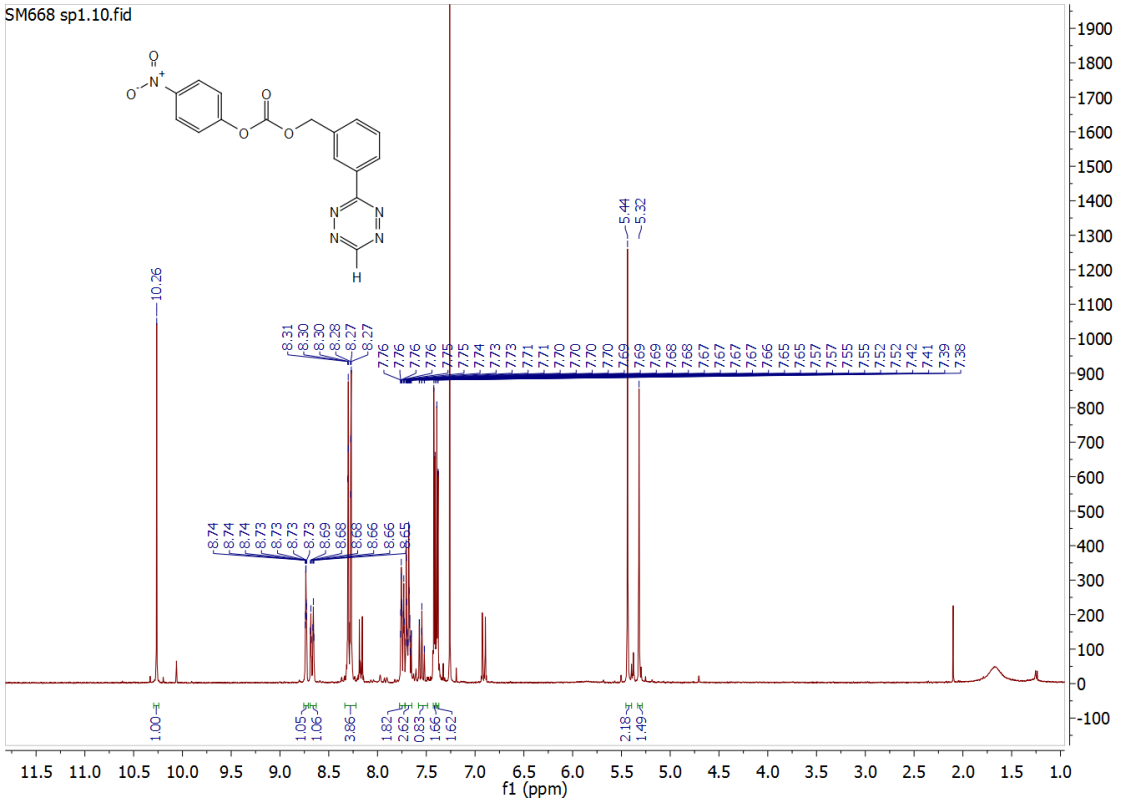
3-(3-(Hydroxymethyl)phenyl)-1,2,4,5-tetrazine

^{1}H:



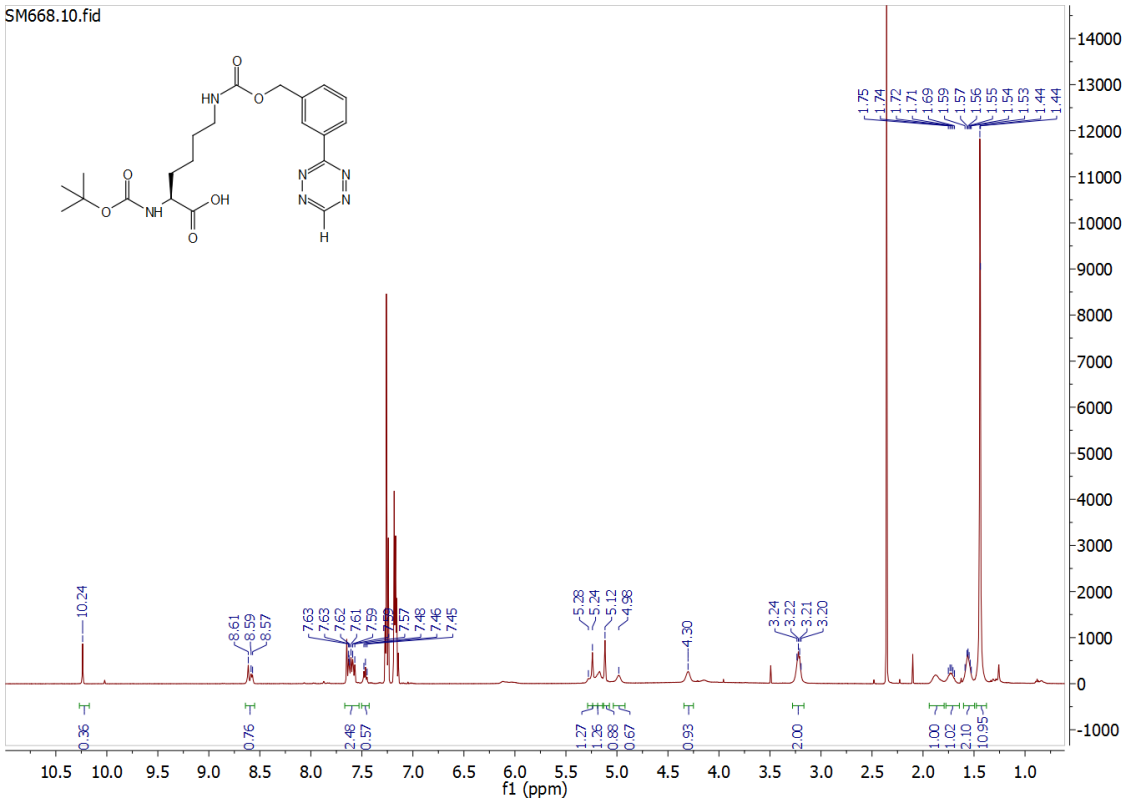
3-(1,2,4,5-tetrazin-3-yl)benzyl-(4-nitrophenyl)-carbonate

^{1}H:



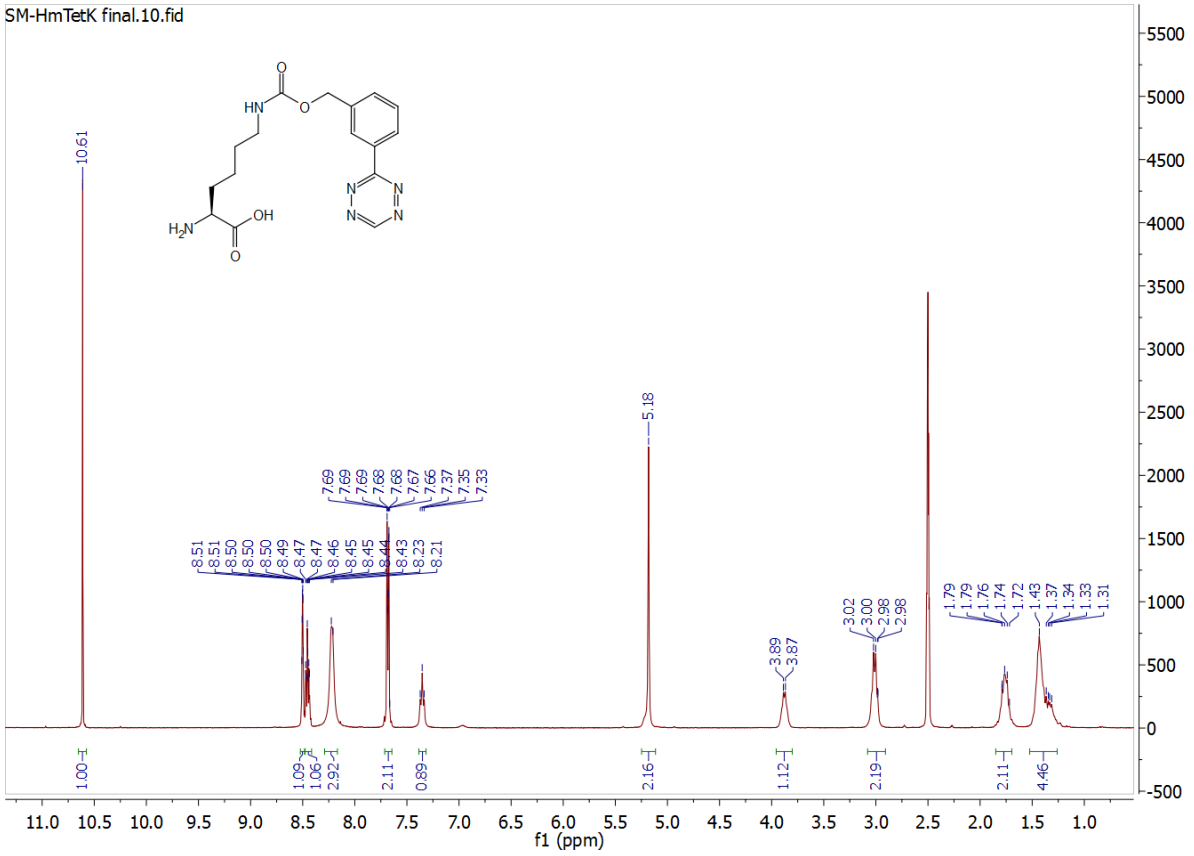
BocHmTetK

^{1}H:

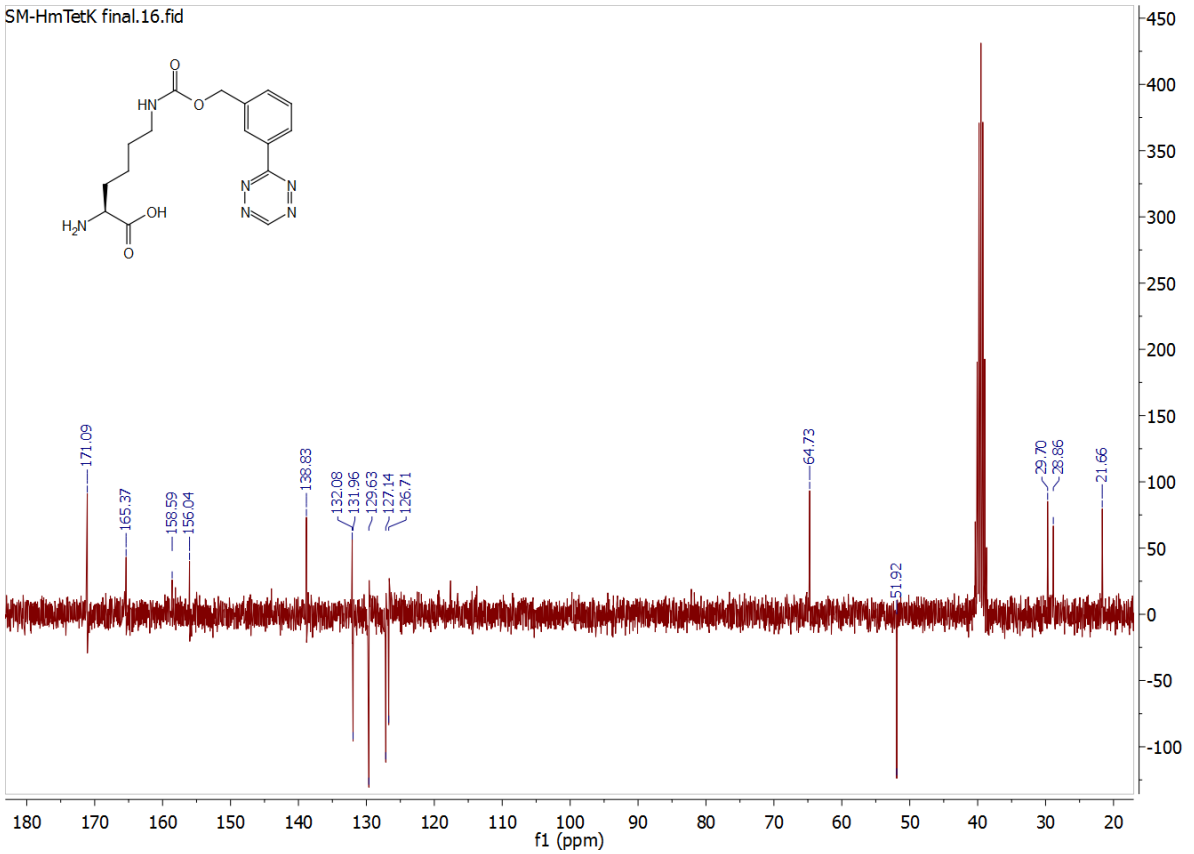


HmTetK

^{1}H:



^{13}C:



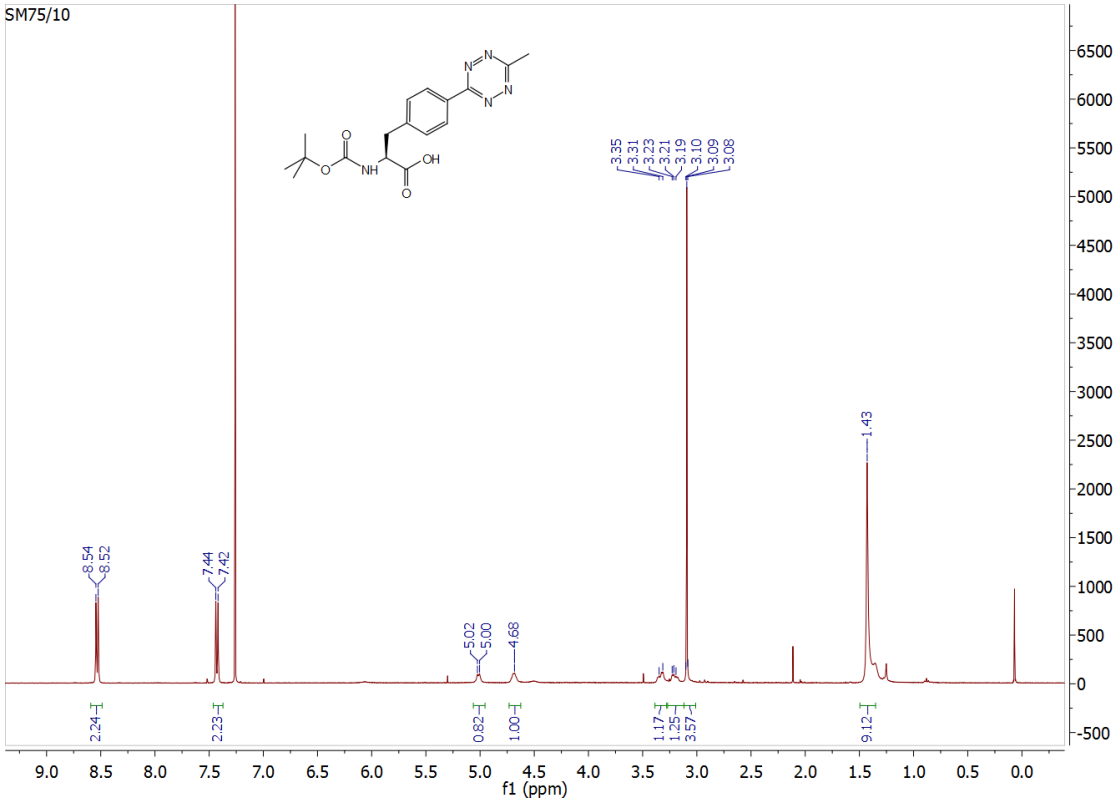
IV.1.2 Chapter 2

IV.1.2.1 TetF

IV.1.2.1.1 Nitrile condensation

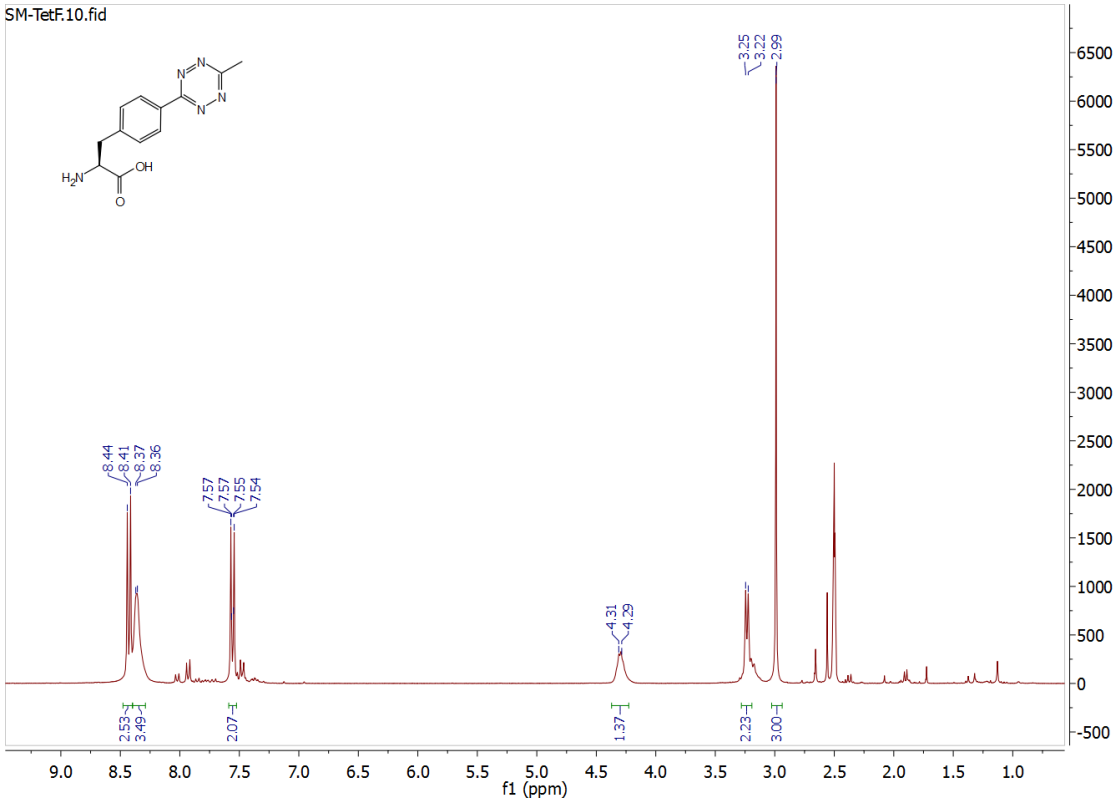
BocTetF

^{1}H:

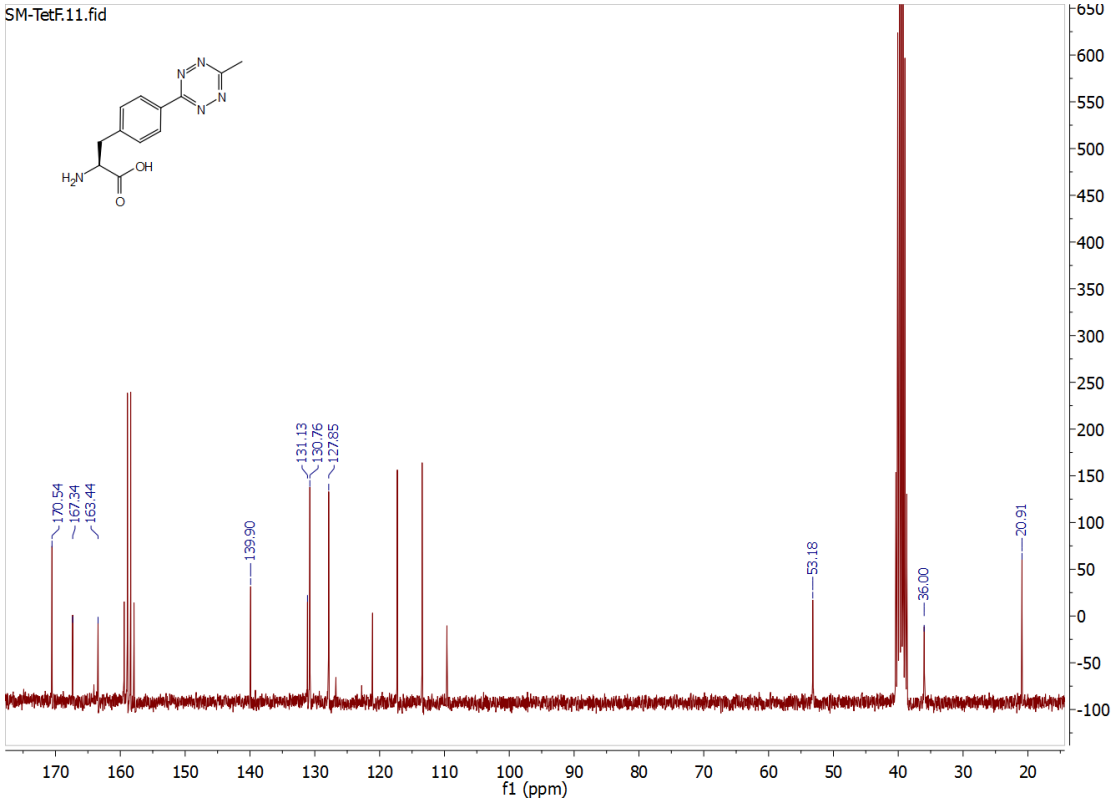


TetF

^{1}H:


SM-TetF.10.fid
8.44
8.41
8.37
8.36
7.57
7.57
7.55
7.54
4.31
4.29
3.25
3.22
2.99
2.53
3.49
2.07
1.37
2.23
3.00
f1 (ppm)

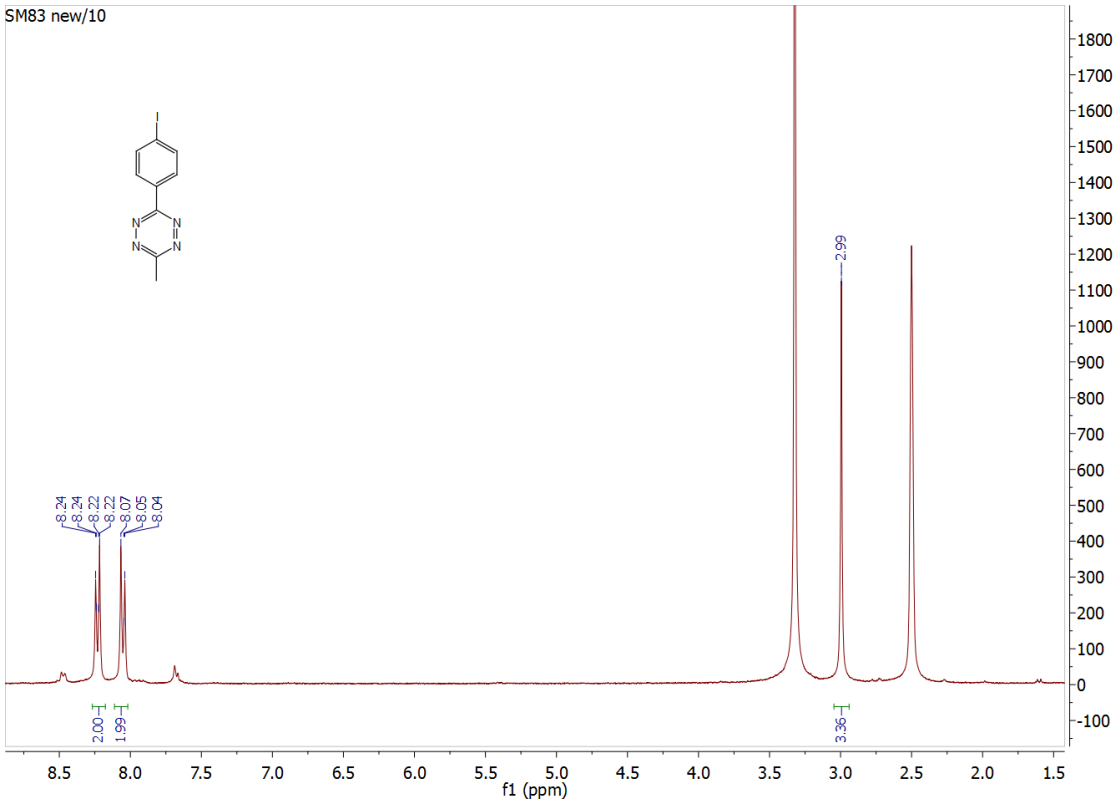
^{13}C:


SM-TetF.11.fid
170.54
167.34
163.44
139.90
131.13
130.76
127.85
53.18
36.00
20.91
f1 (ppm)

IV.1.2.1.2 ß-CH arylation

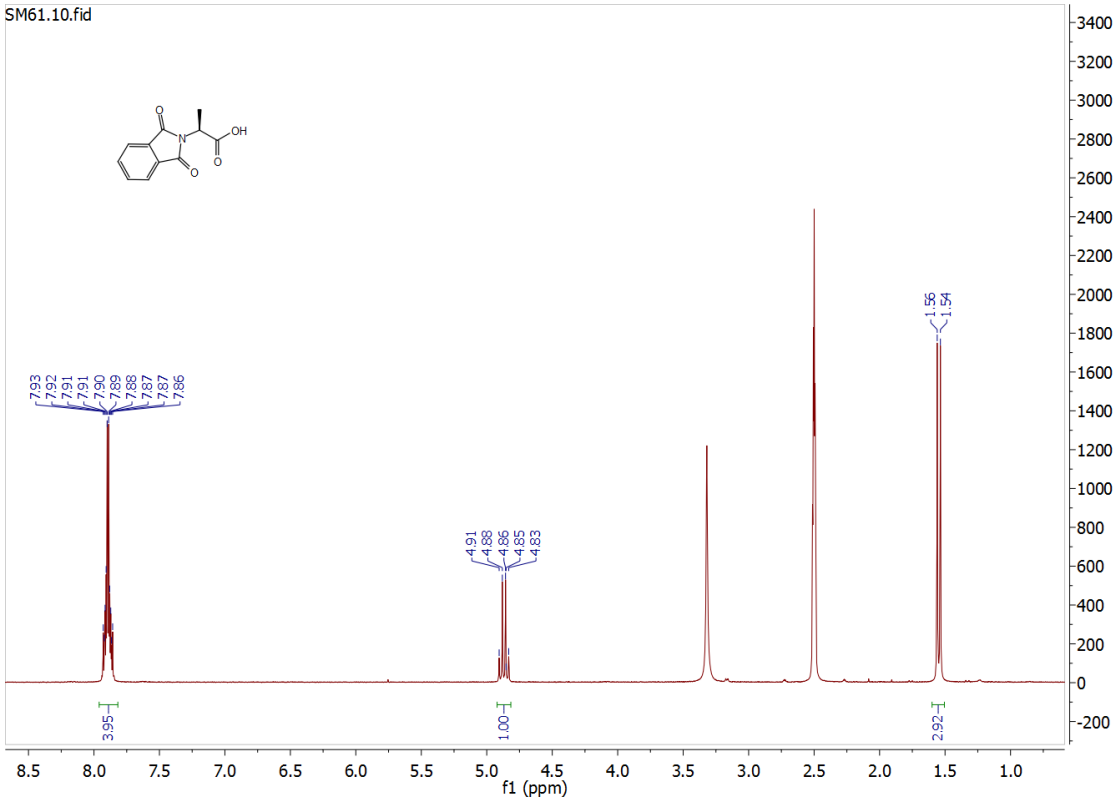
3-(4-iodophenyl)-6-methyl-1,2,4,5-tetrazine

^{1}H:



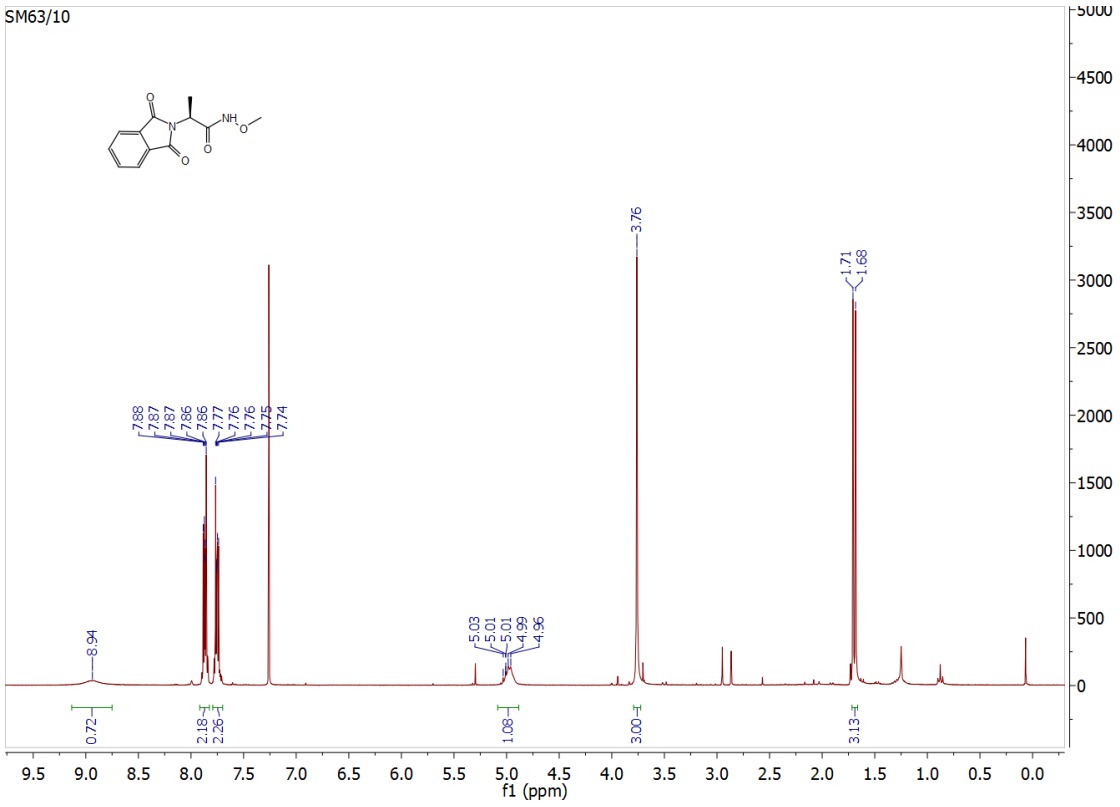
N^2,*N*^{2}-Phthaloyl-L-alanine

^{1}H:



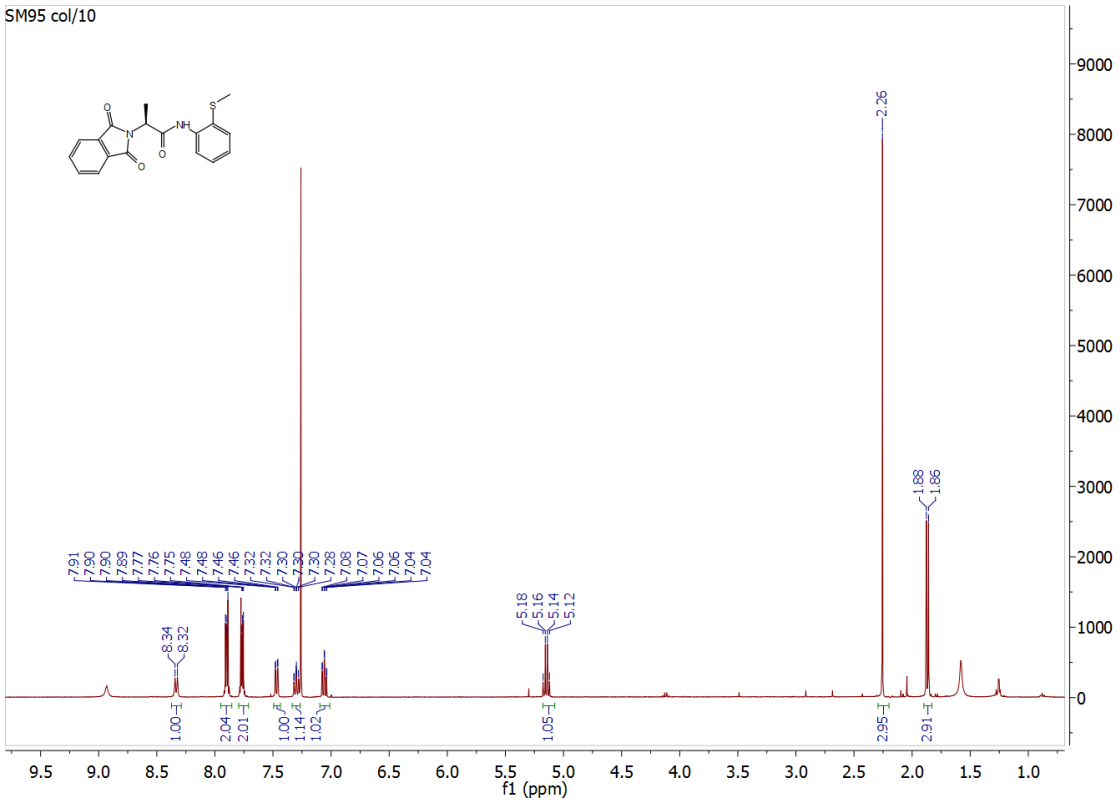
N^1-Methoxy-N^2,N^2-phthaloyl-L-alaninamide

^{1}H:



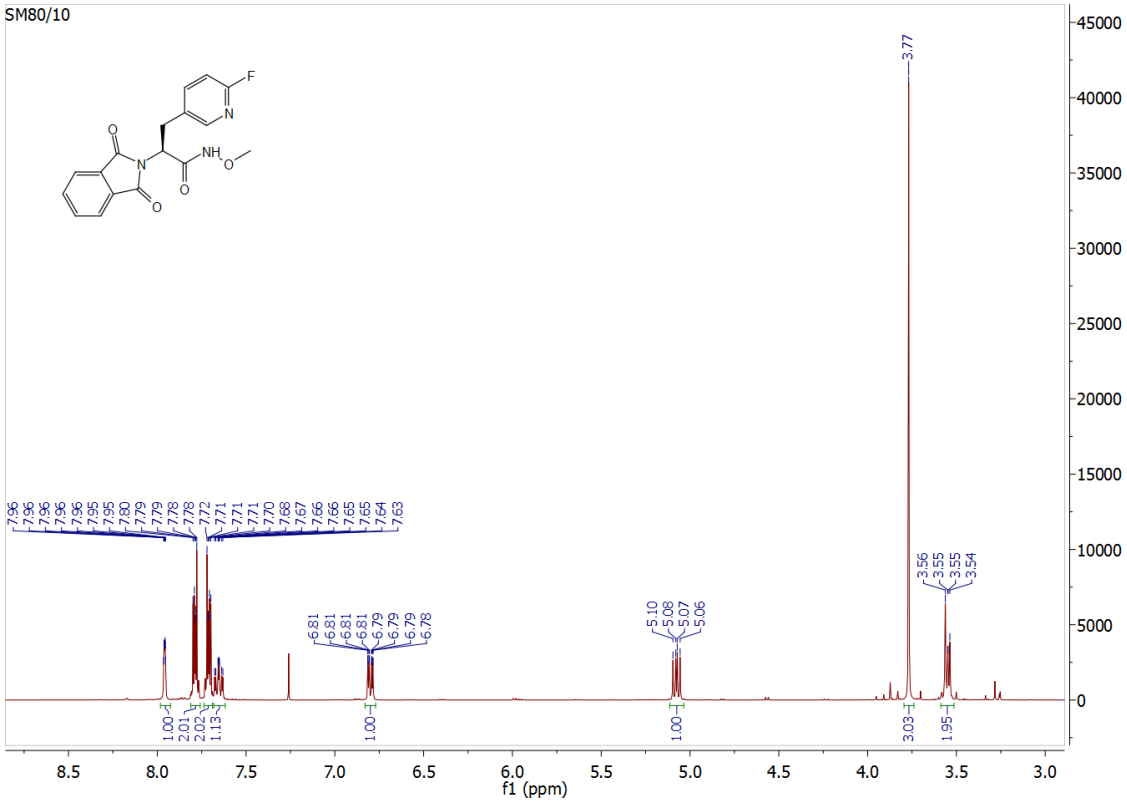
N^1-(2-(methylthio)phenyl)-N^2,N^2-phthaloyl-L-alaninamide

^{1}H:



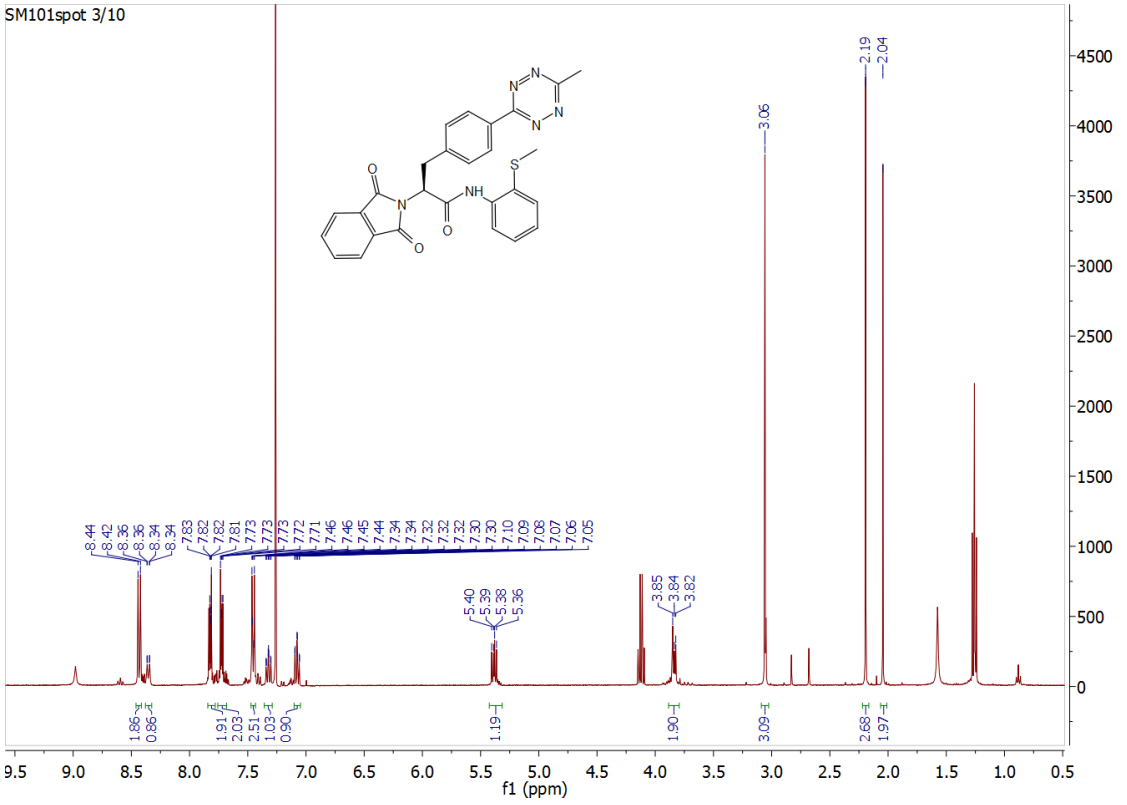
Methyl N^2,N^2-phthaloyl-ß-(6-fluoropyridin-3-yl)-L-alaninate

^{1}H:



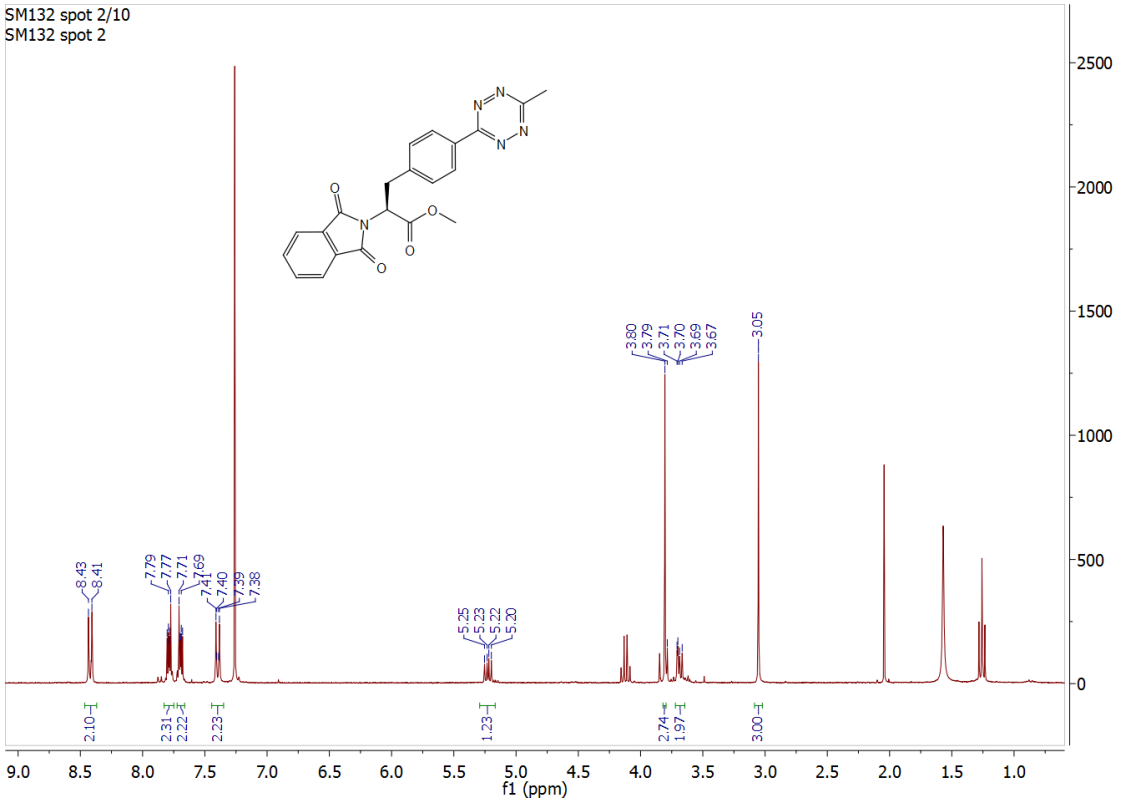
N^1-(2-(methylthio)phenyl)-N^2,N^2-phthaloyl-ß-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alaninamide

^{1}H:



Methyl N^2,N^2-phthaloyl-ß-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)phenyl)-L-alaninate

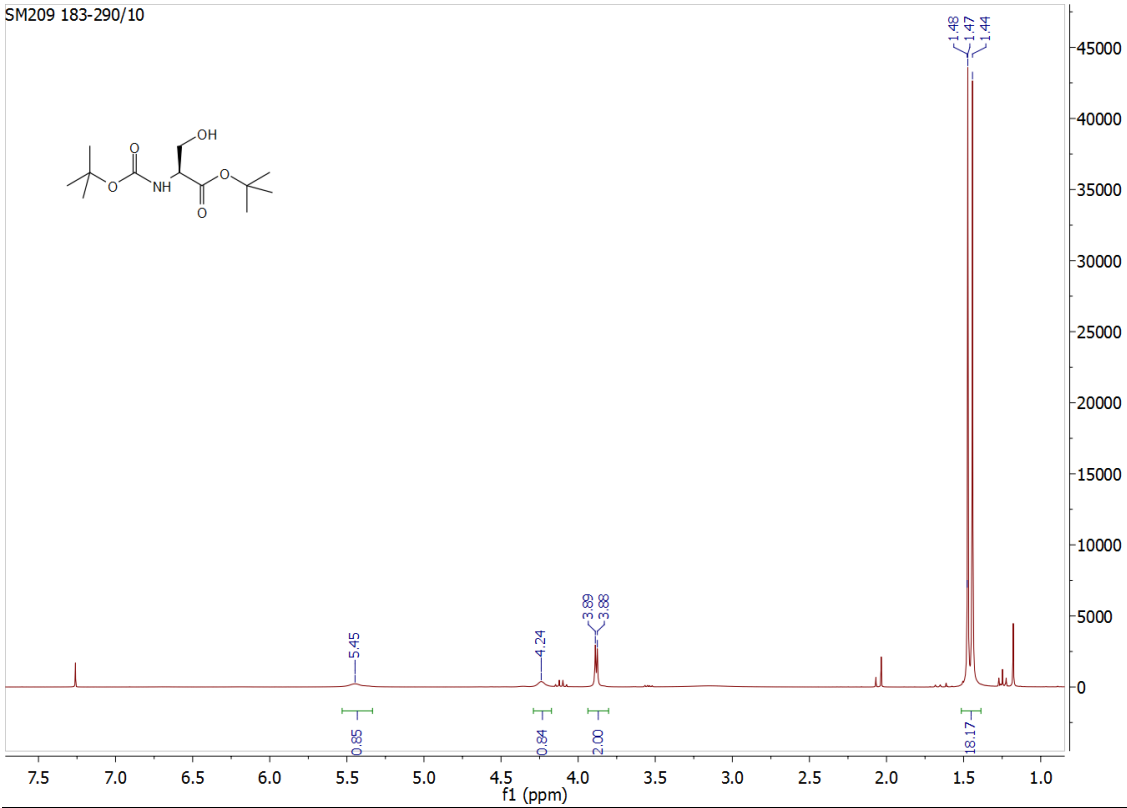
^{1}H:



IV.1.2.1.3 Negishi cross-coupling

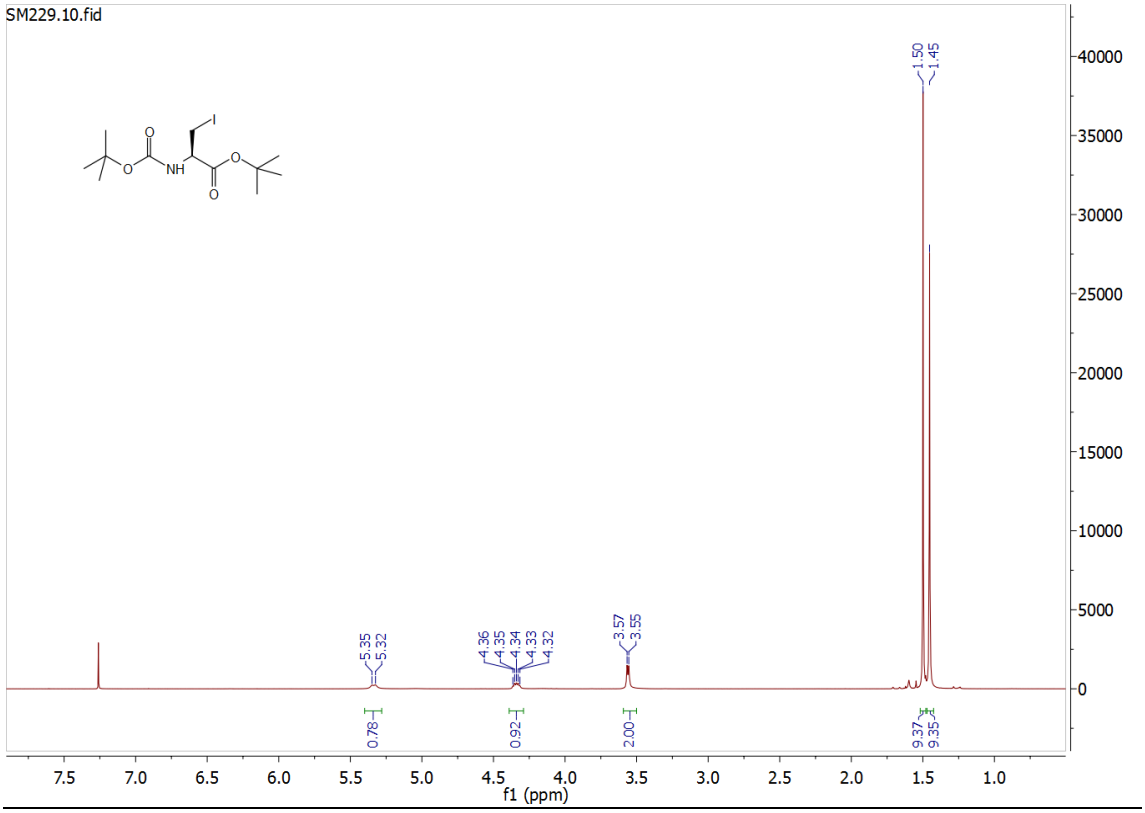
tert-butyl (*tert*-butoxycarbonyl)-L-serinate

^{1}H:



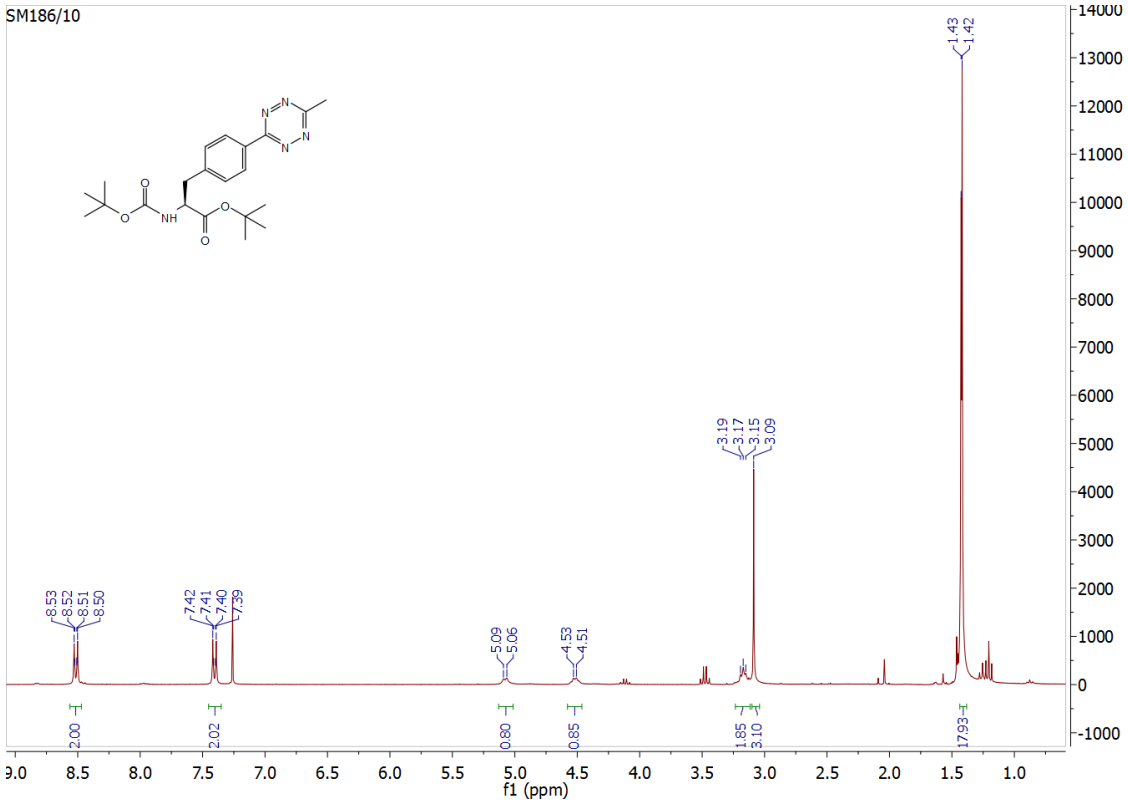
tert-butyl (*tert*-butoxycarbonyl)-*ß*-iodo-L-alaninate

^{1}H:



BocTetFOtBu

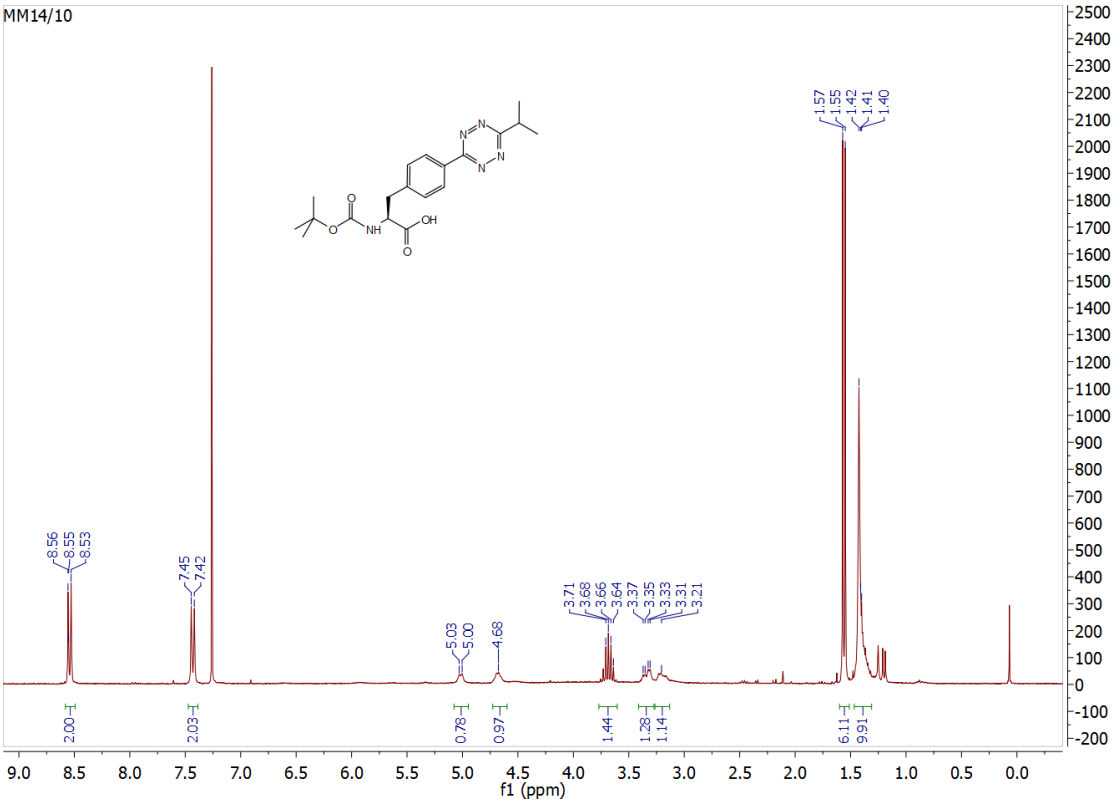
^{1}H:



IV.1.2.2 IpTetF

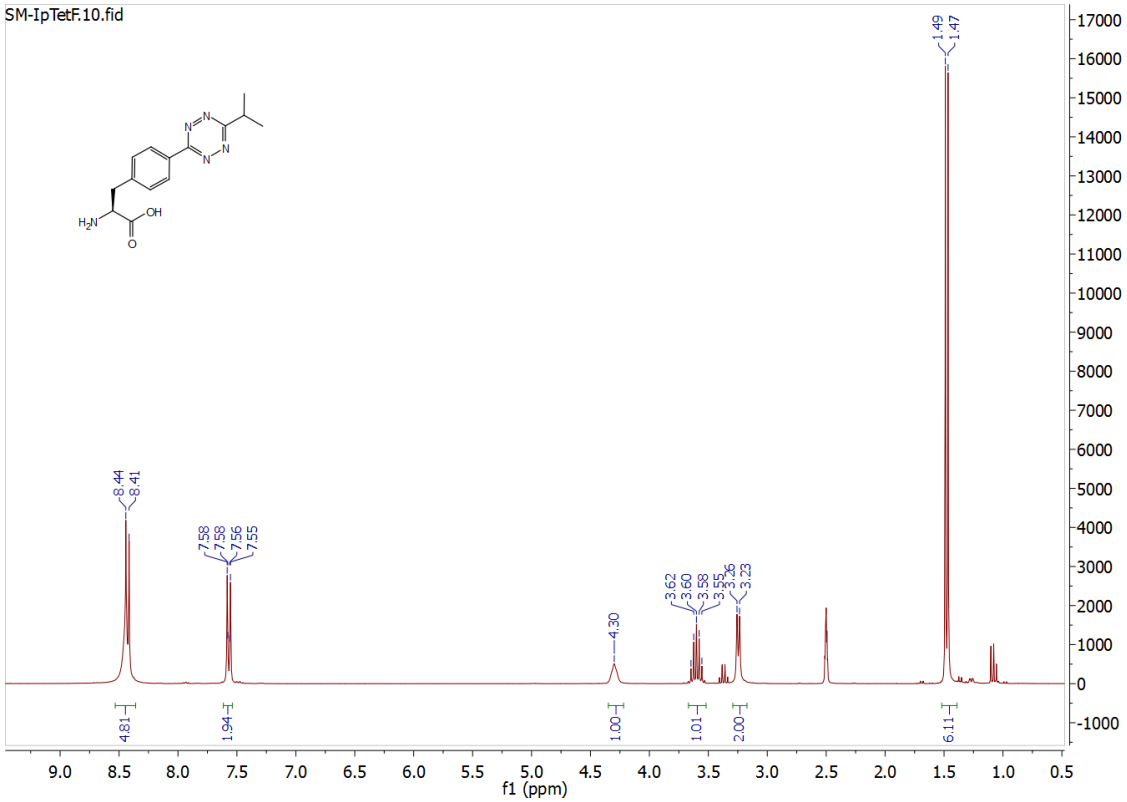
BocIpTetF

^{1}H:

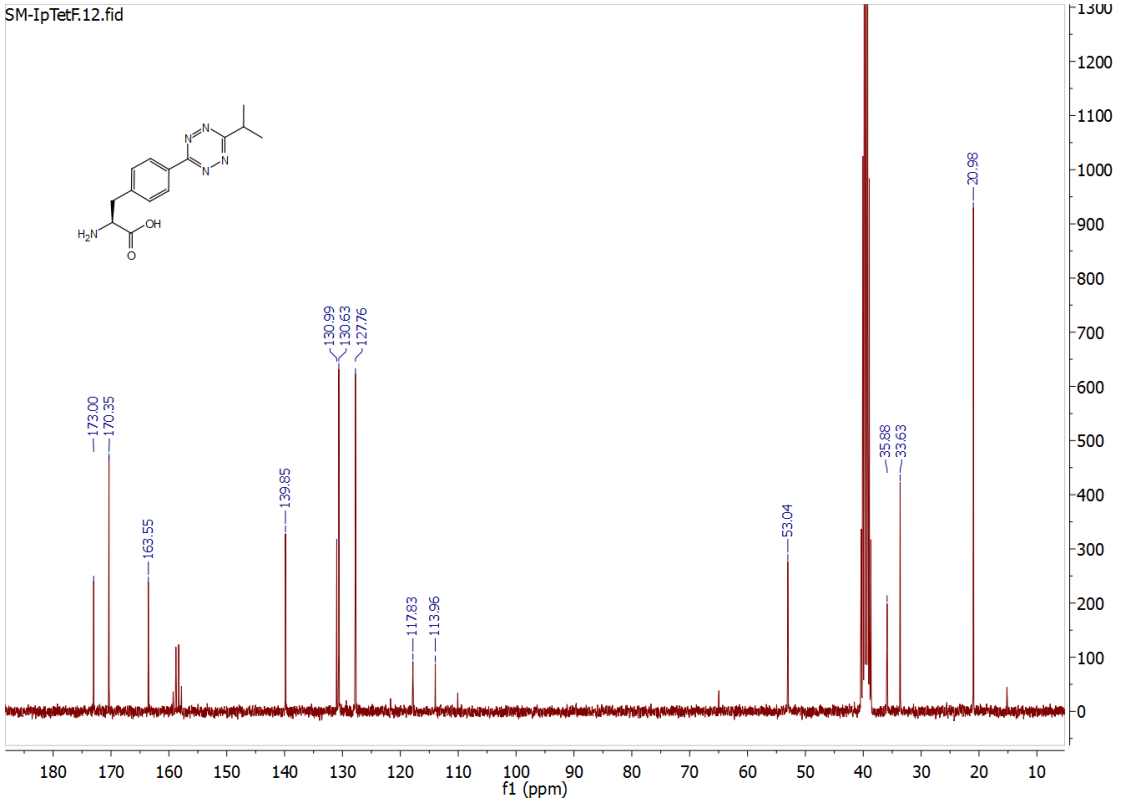


IpTetF

^{1}H:



^{13}C:

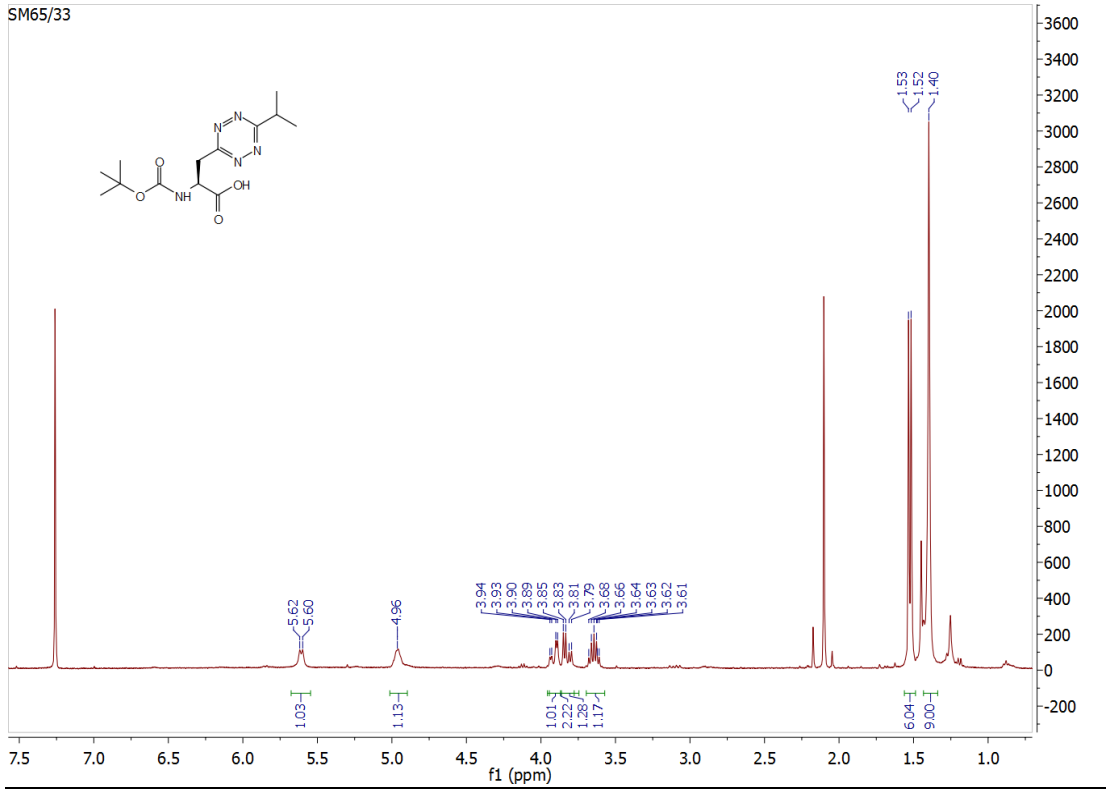


IV.1.2.3 TetA

IV.1.2.3.1 Nitrile condensation

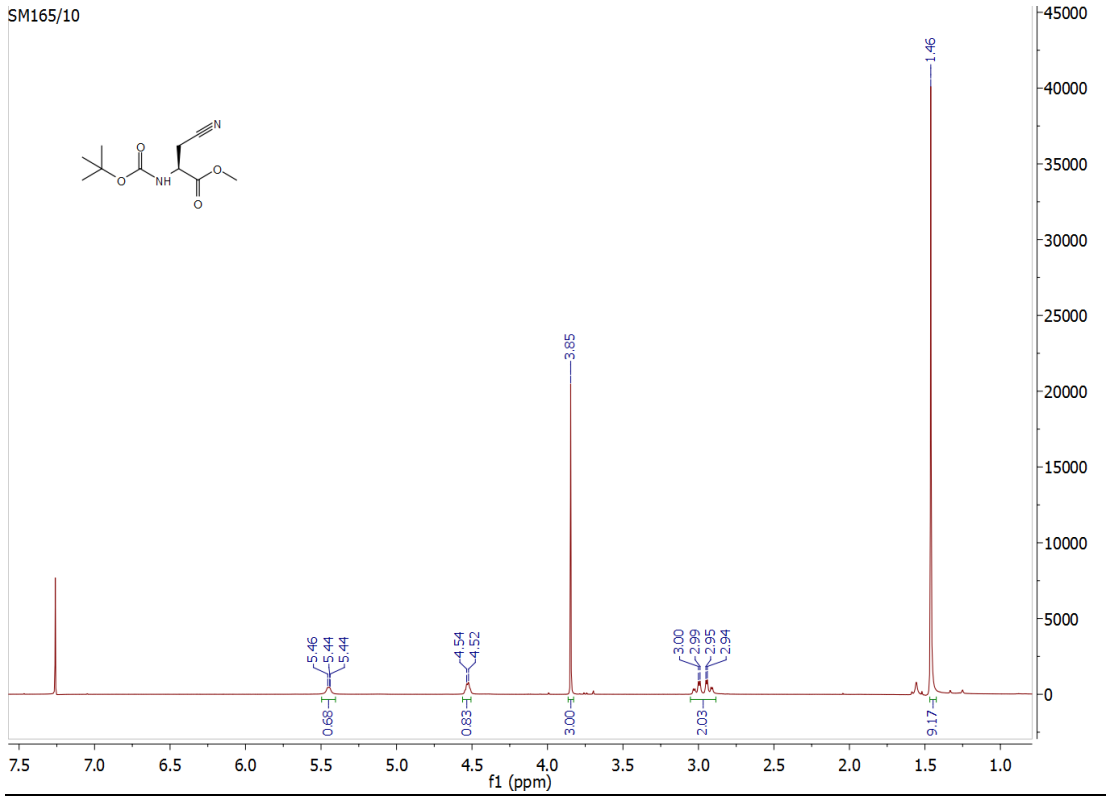
(*tert*-butoxycarbonyl)-*ß*-(6-isopropyl-1,2,4,5-tetrazin-3-yl)-L-alanine

^{1}H:



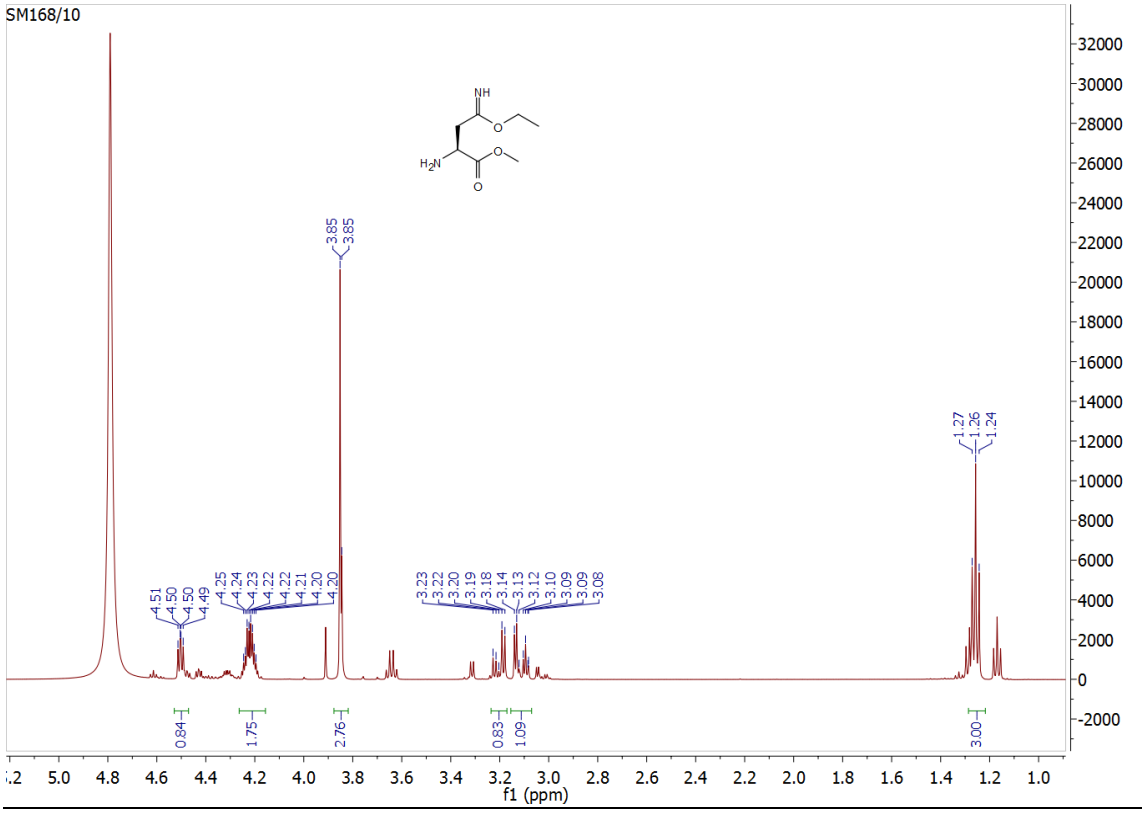
Methyl (*tert*-butoxycarbonyl)-*ß*-(4-cyano)-L-alaninate

^{1}H:



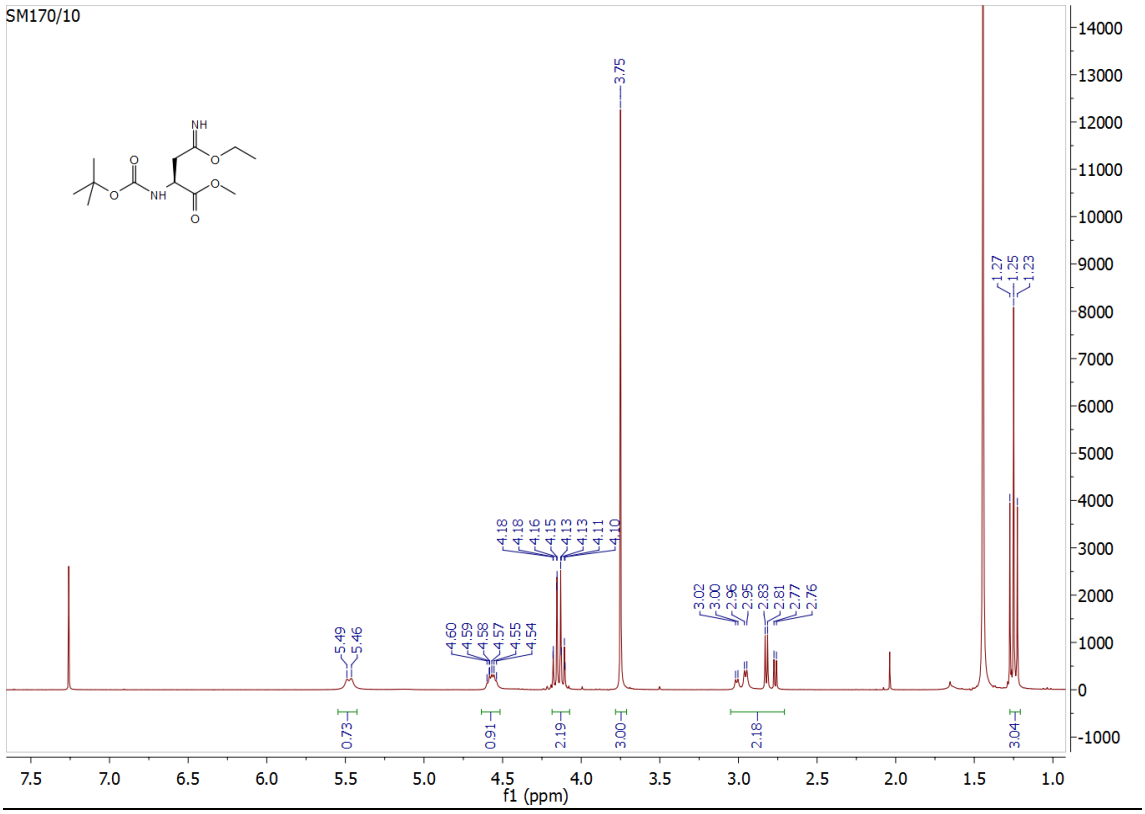
Methyl *ß*-(ethoxy(imino)methyl)-L-alaninate hydrochloride

^{1}H:



Methyl (*tert*-butoxycarbonyl)-*ß*-(ethoxy(imino)methyl)-L-alaninate

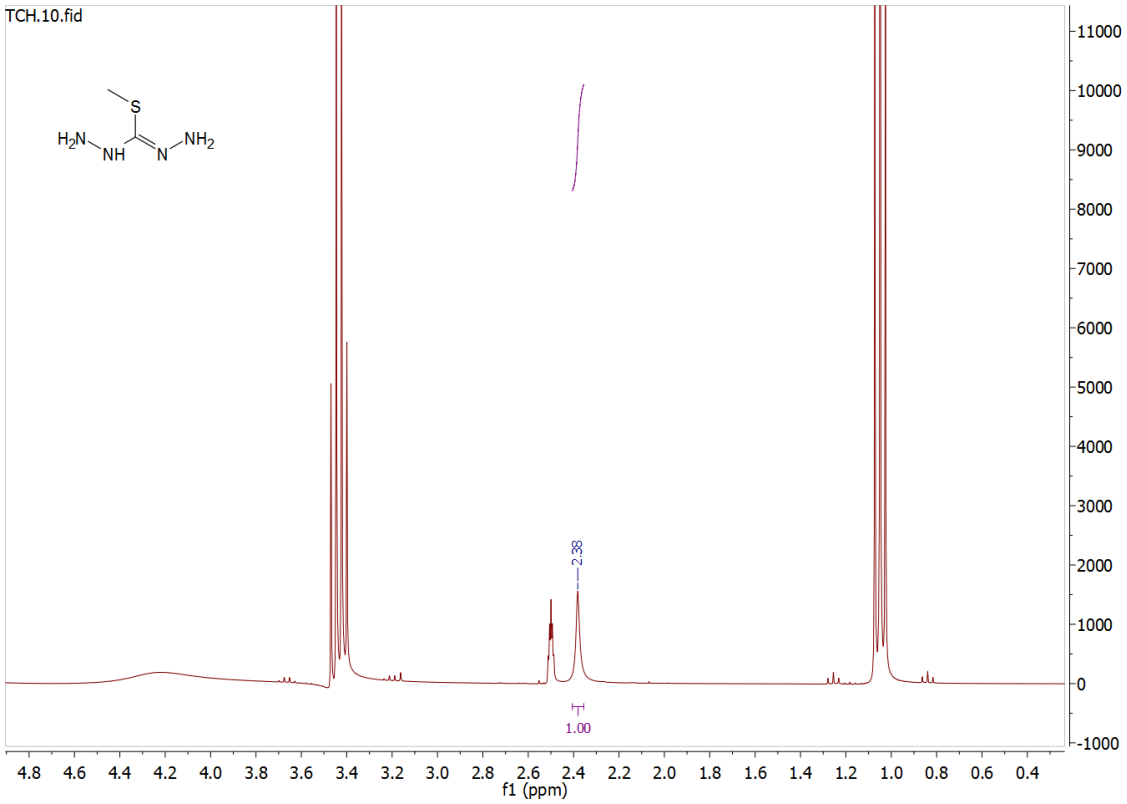
^{1}H:



IV.1.2.3.1 Negishi cross-coupling

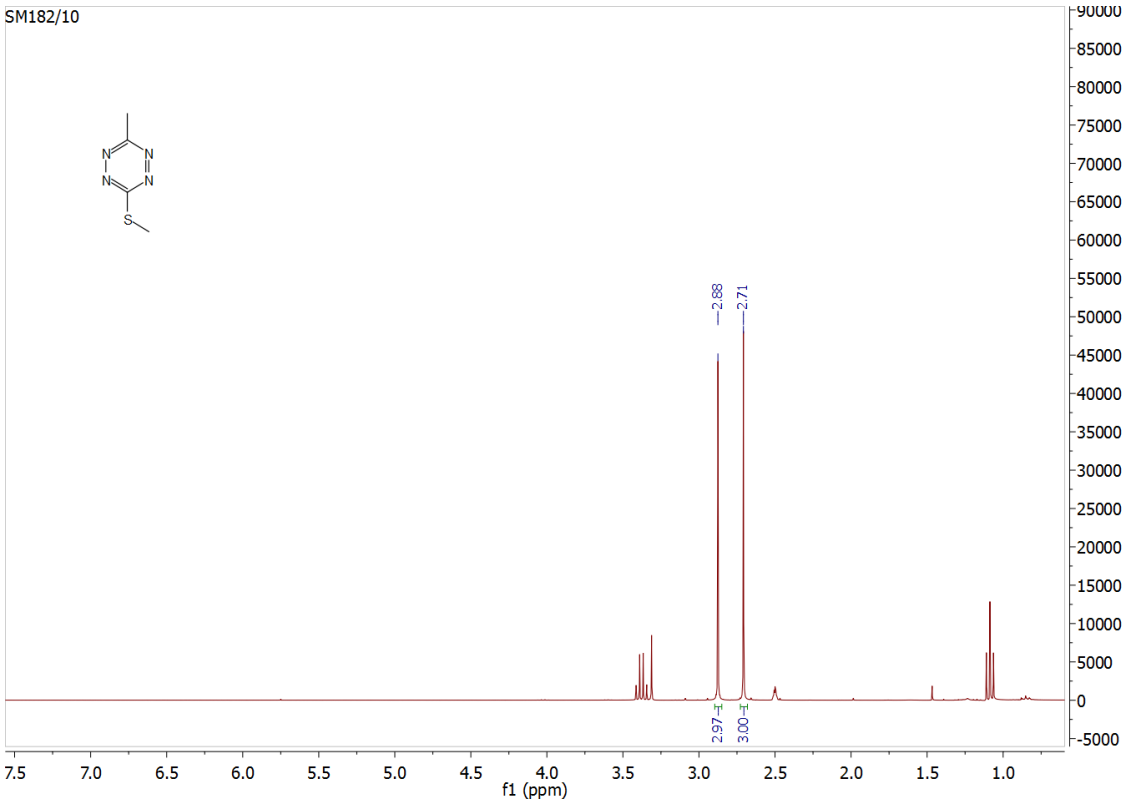
S-Methylisothiocarbohydrazide hydroiodide

^{1}H:



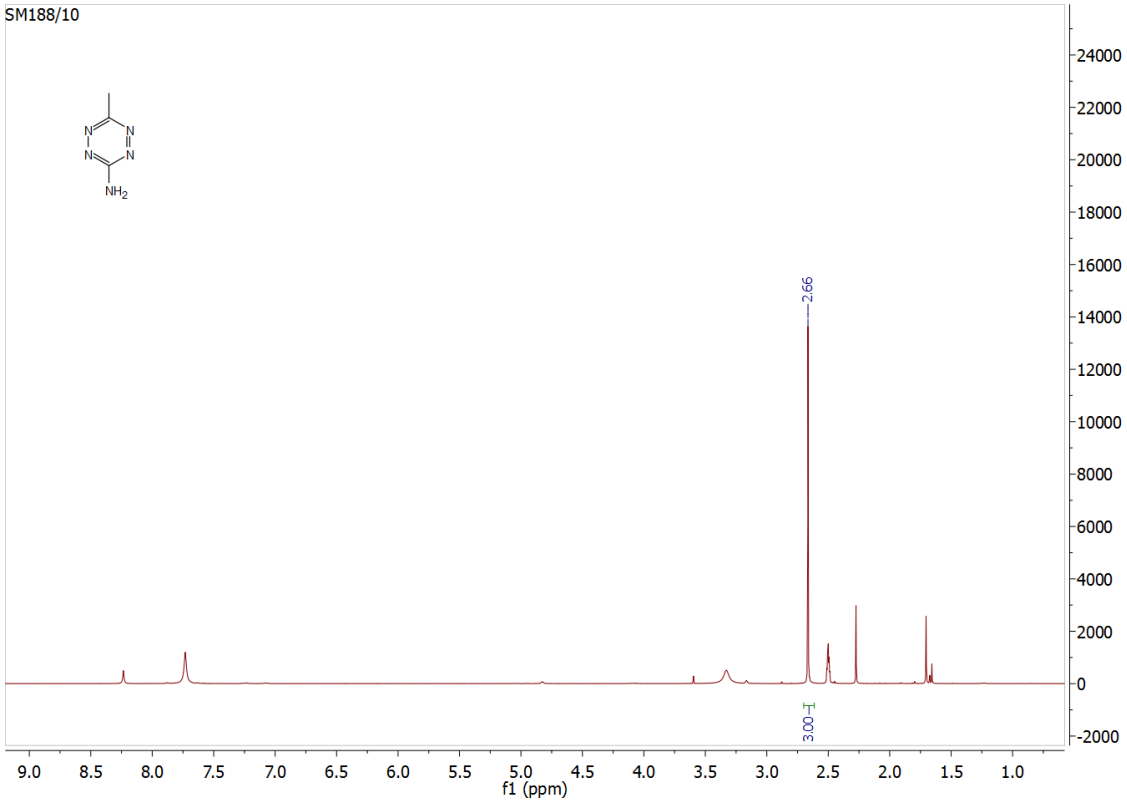
3-*S*-Methylthio-6-methyl-1,2,4,5-tetrazine

^{1}H:

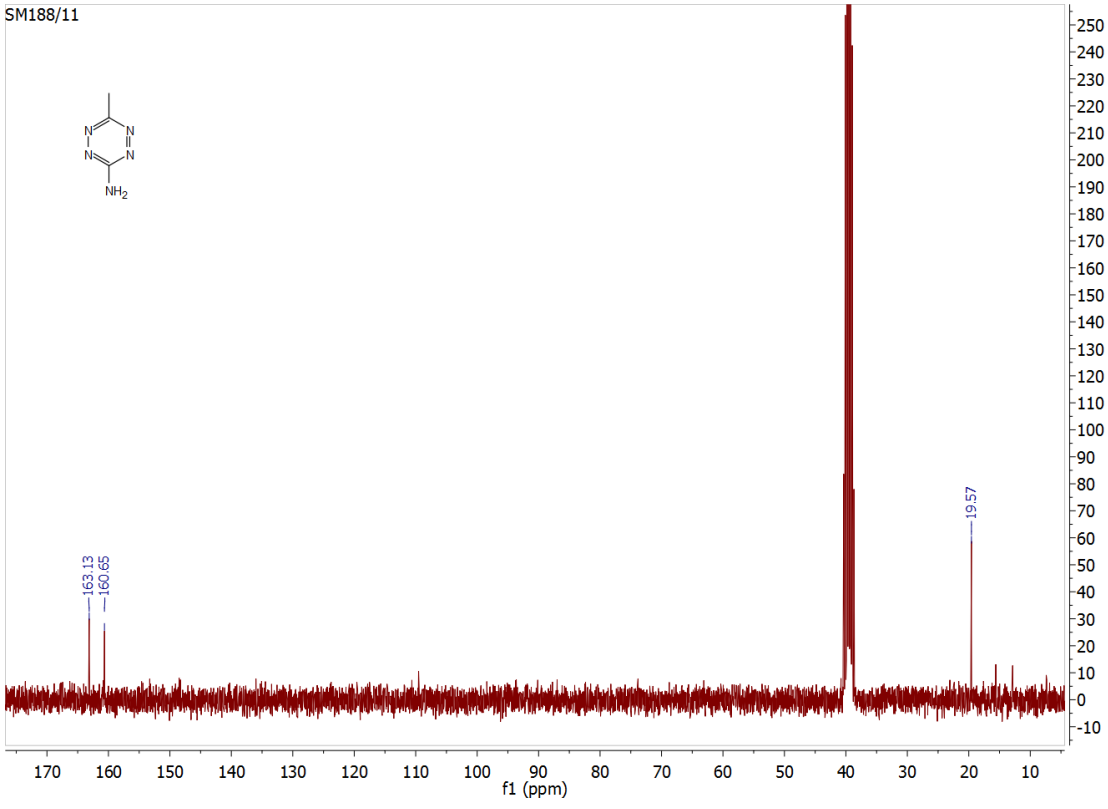


3-Amino-6-methyl-1,2,4,5-tetrazine

^{1}H:

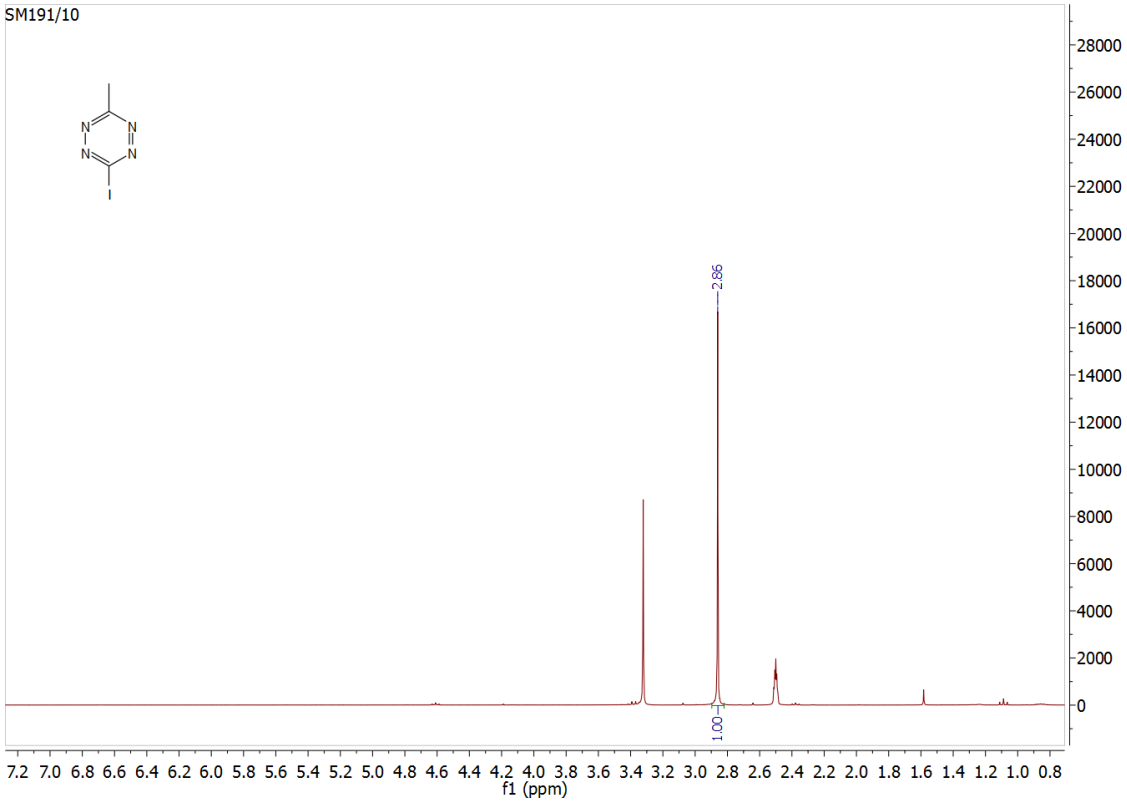


^{13}C:

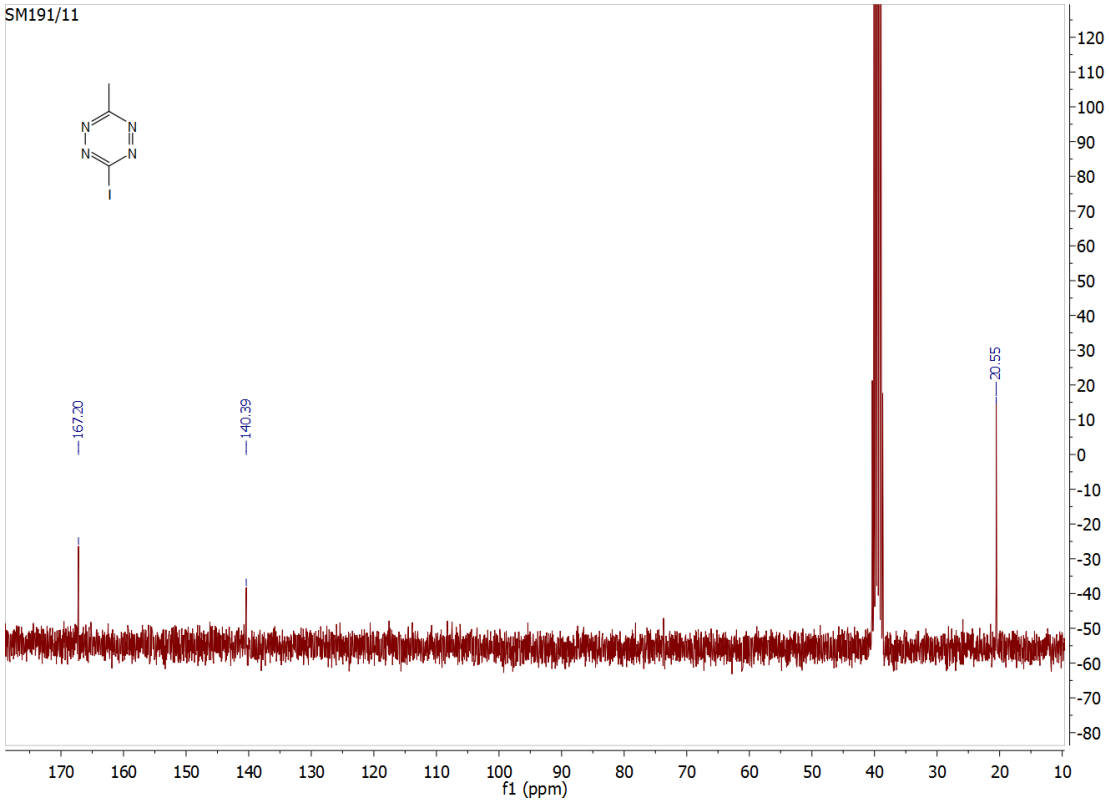


3-Iodo-6-methyl-1,2,4,5-tetrazine

^{1}H:

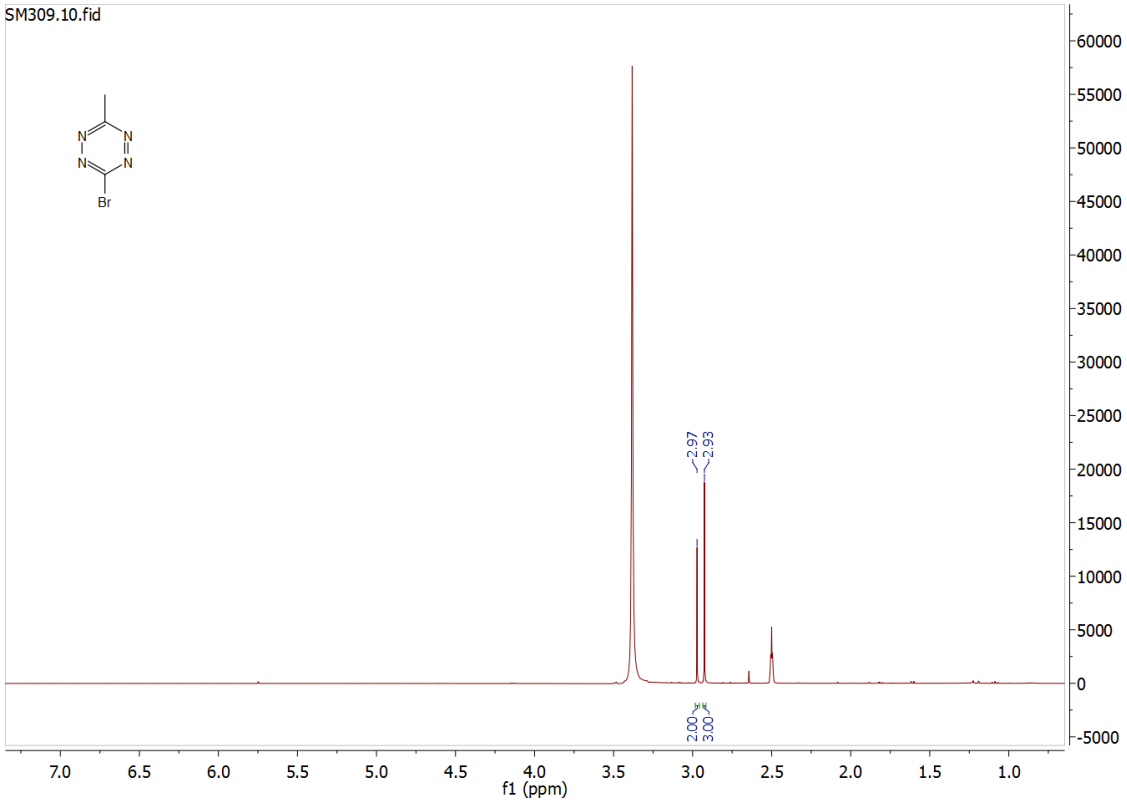


^{13}C:

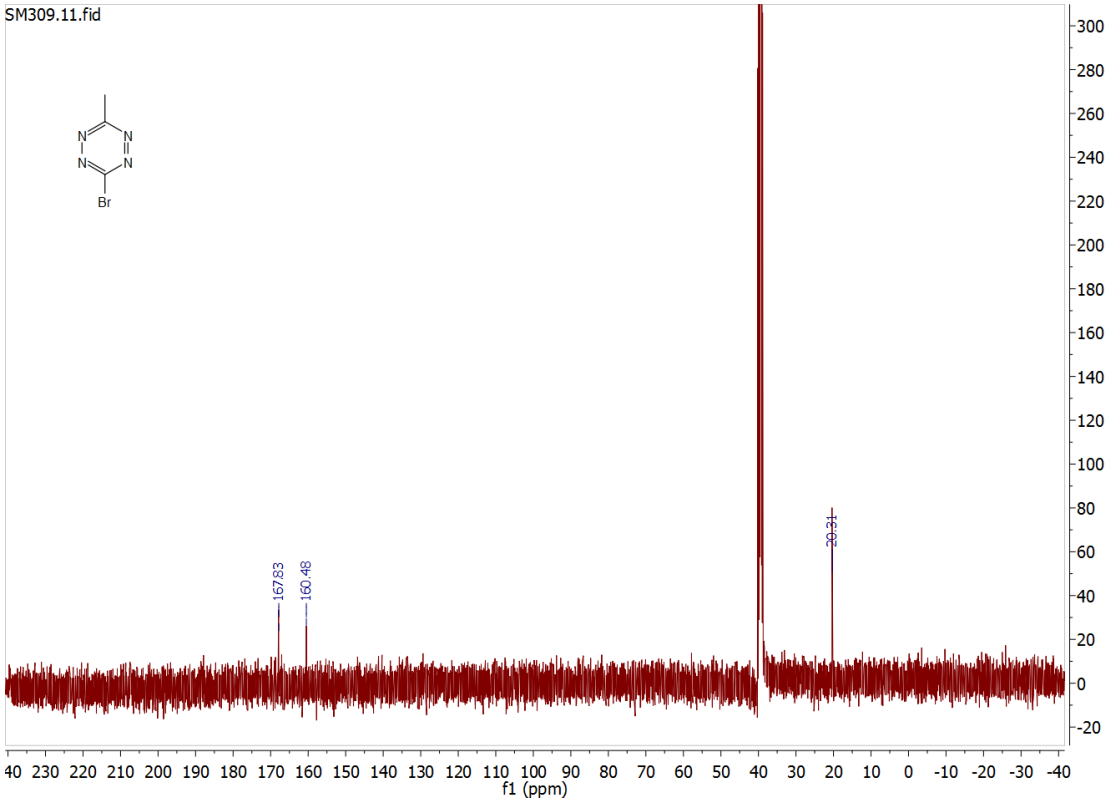


3-Bromo-6-methyl-1,2,4,5-tetrazine

^{1}H:

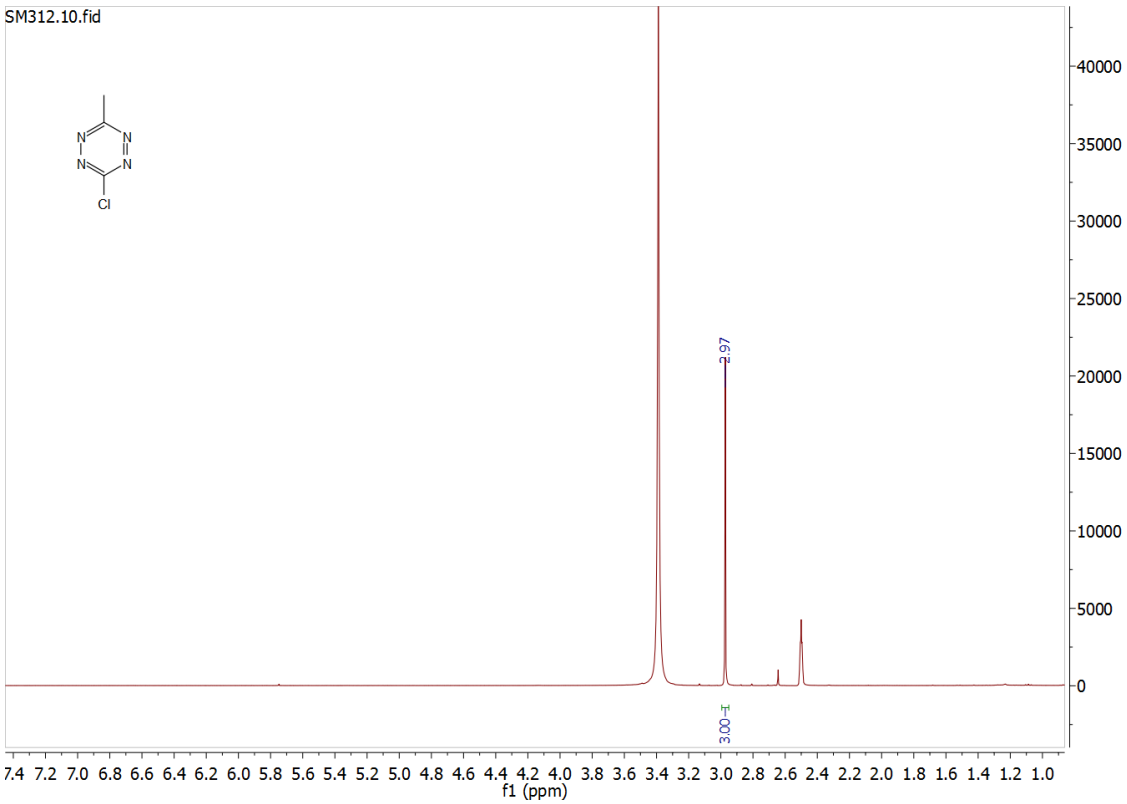


^{13}C:

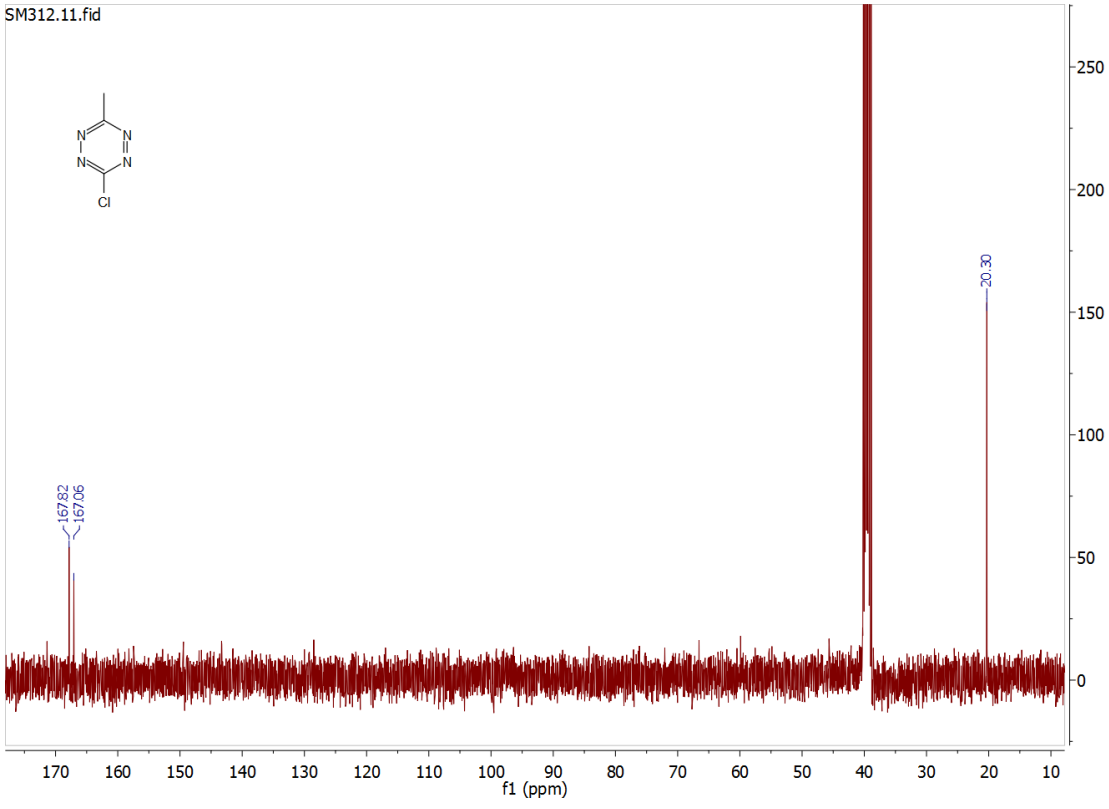


3-Chloro-6-methyl-1,2,4,5-tetrazine

^{1}H:

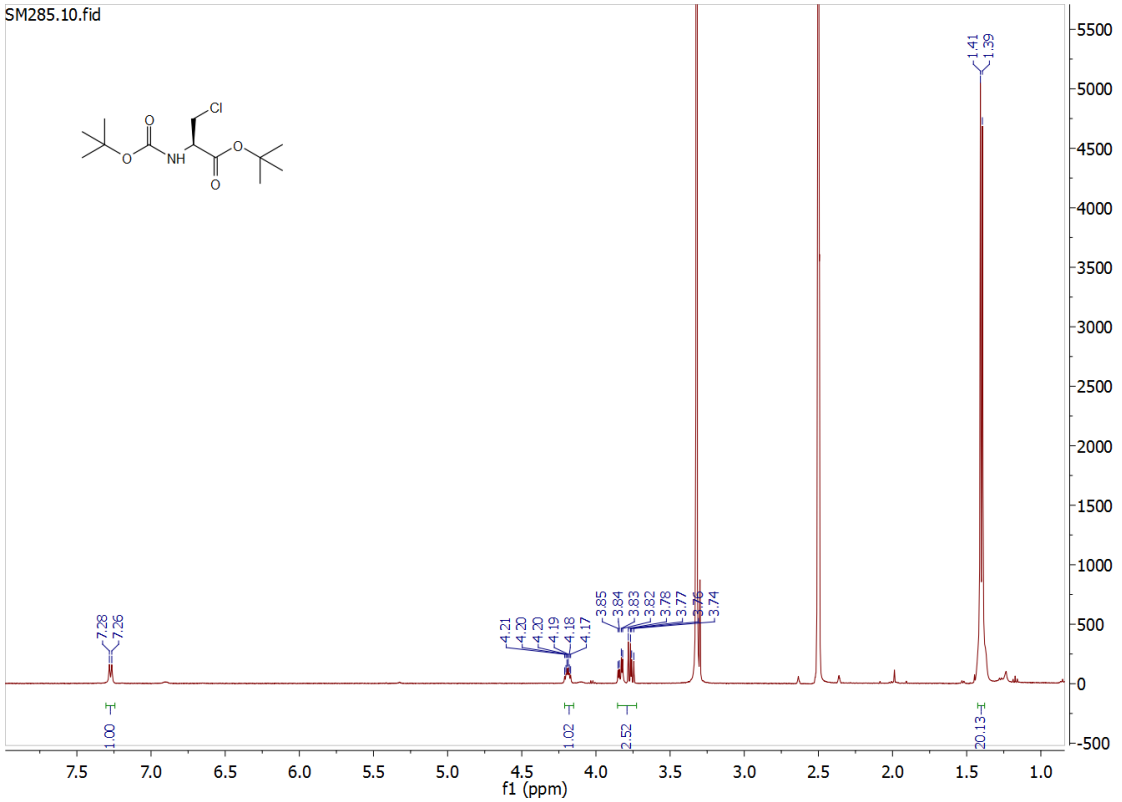


^{13}C:



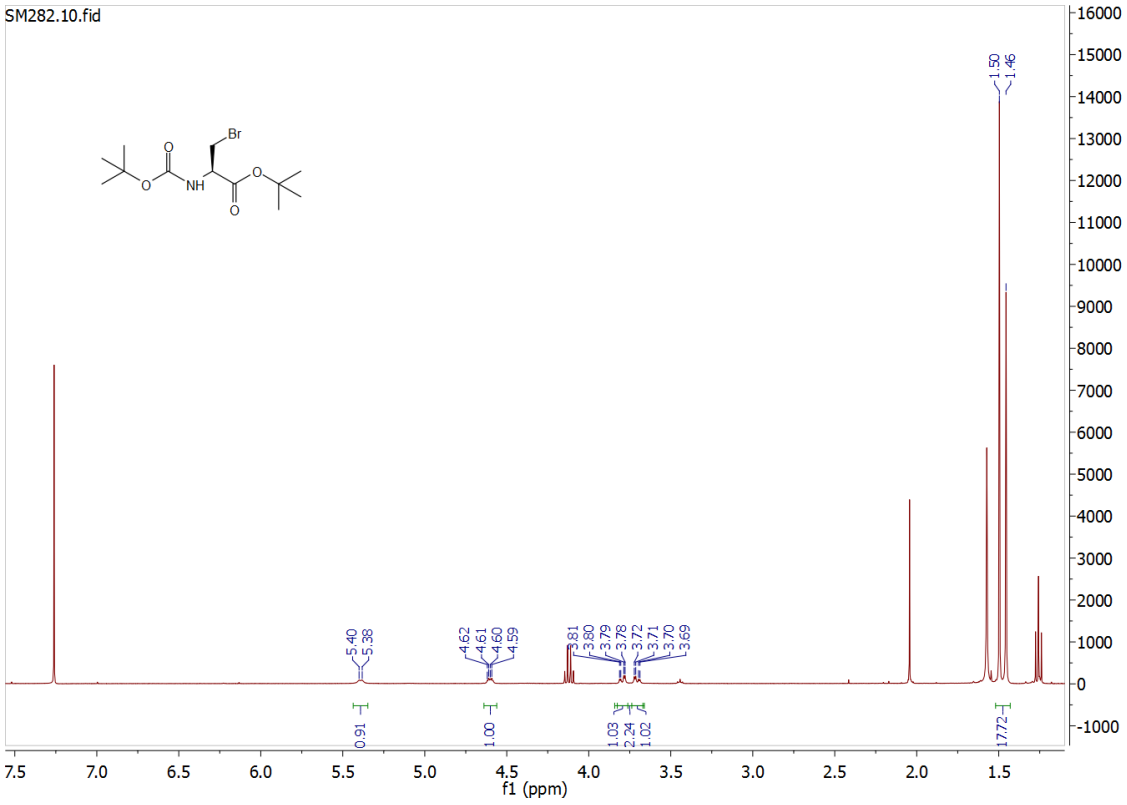
tert-butyl (*tert*-butoxycarbonyl)-*ß*-chloro-L-alaninate

^{1}H:



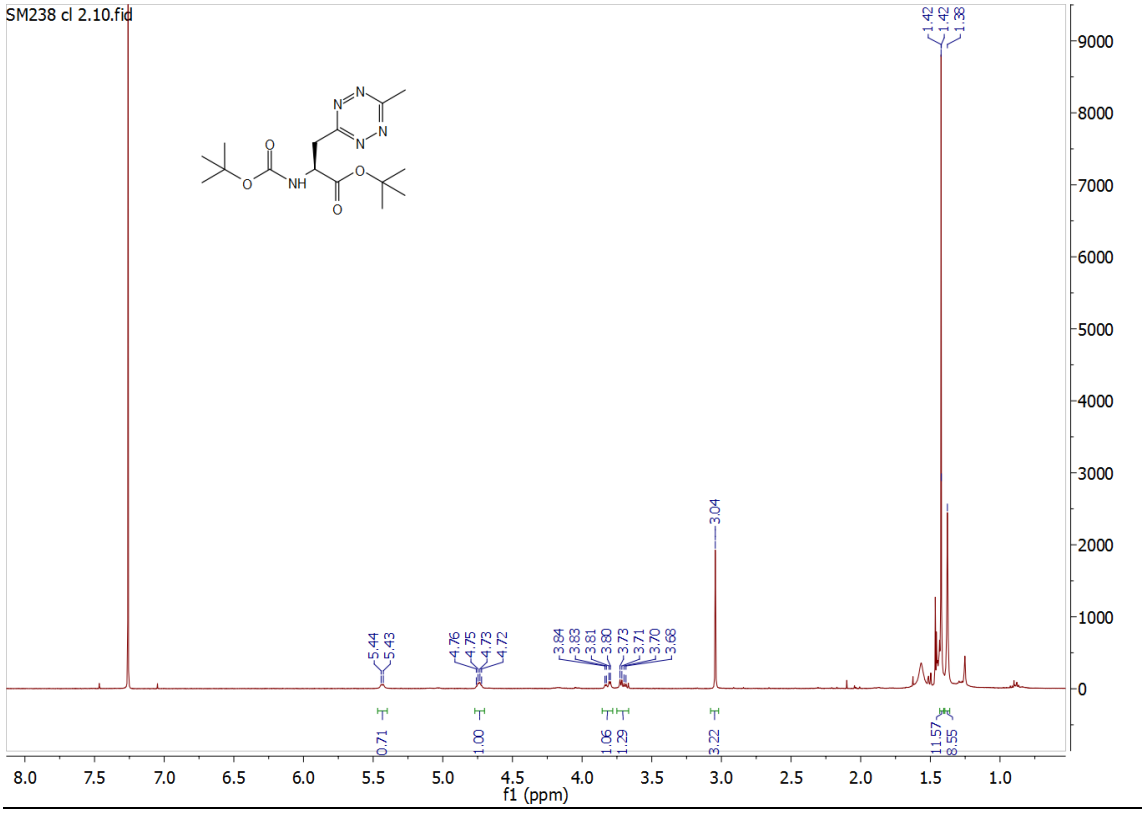
tert-butyl (*tert*-butoxycarbonyl)-*ß*-bromo-L-alaninate

^{1}H:



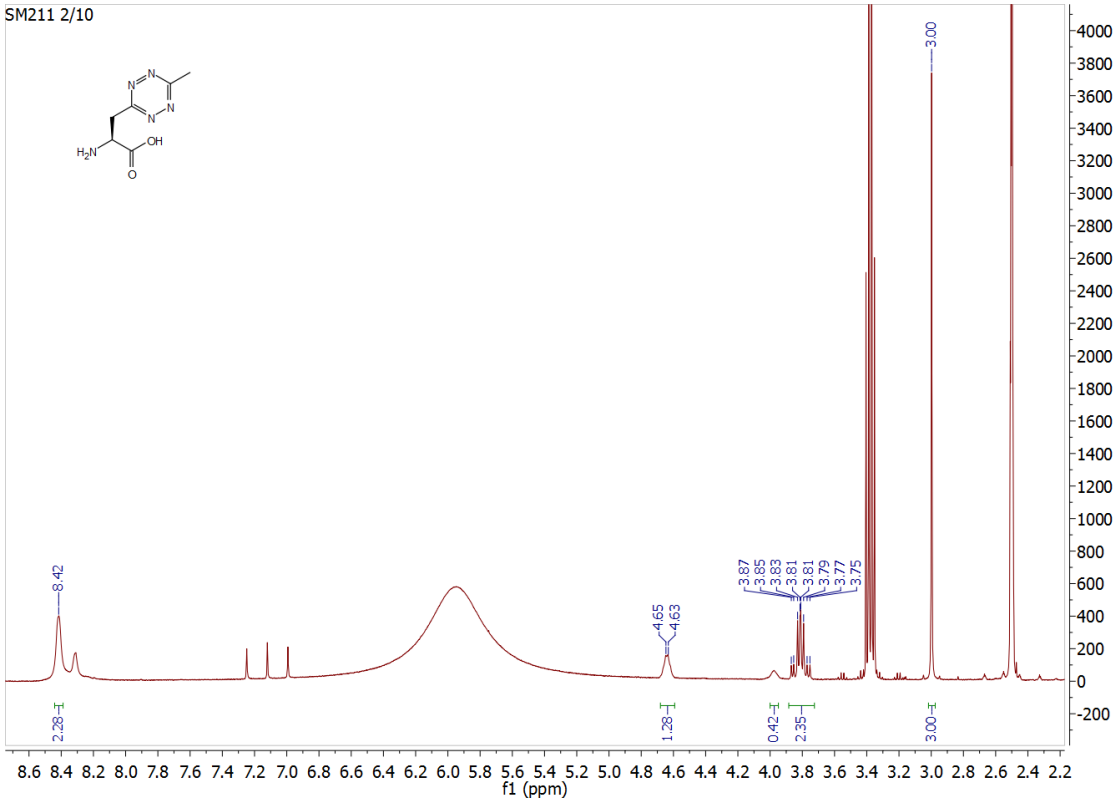
BocTetAOtBu

^{1}H:

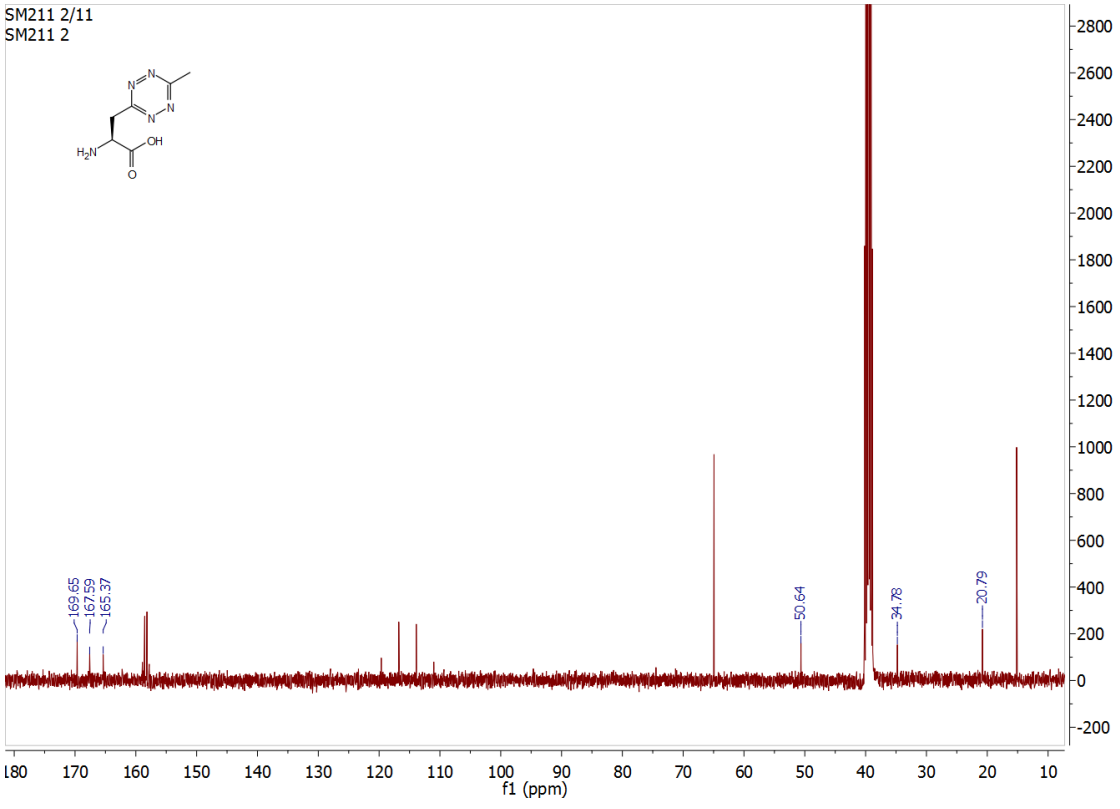


TetA

^{1}H:


SM211 2/10
8.42
4.65
4.63
3.87
3.85
3.83
3.81
3.81
3.79
3.77
3.75
3.00
2.28
1.28
0.42
2.35
3.00
f1 (ppm)

^{13}C:

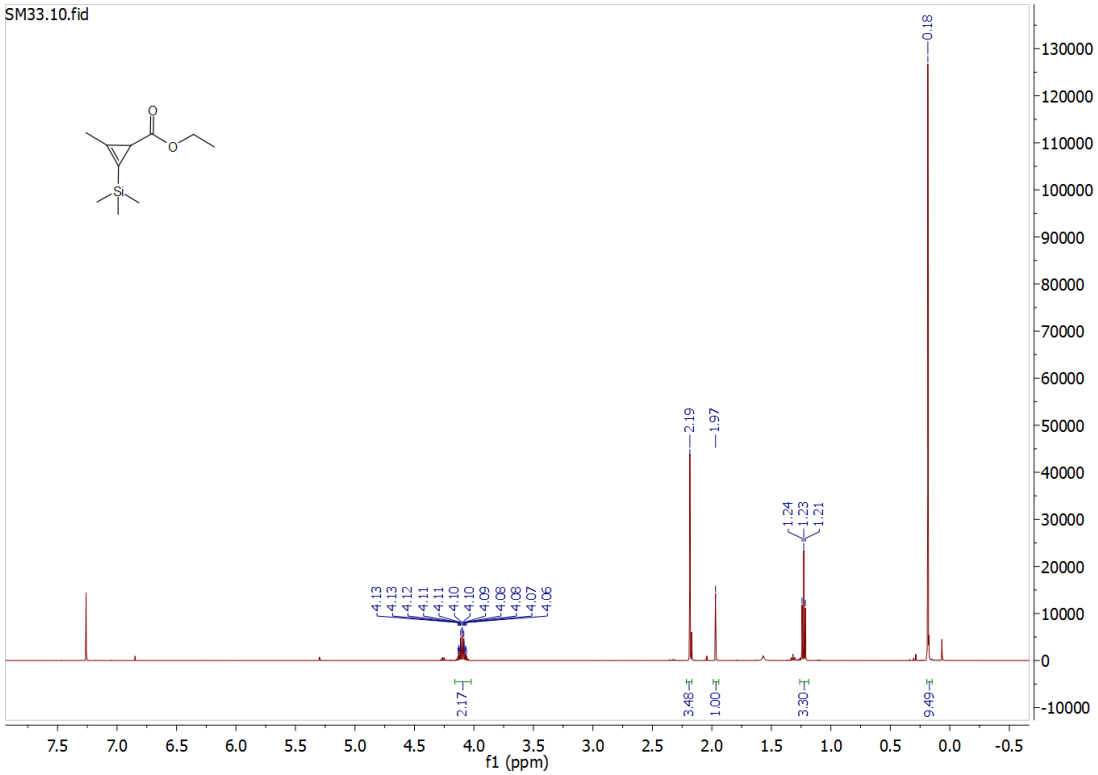

SM211 2/11
SM211 2
169.65
167.59
165.37
50.64
34.78
20.79
f1 (ppm)

IV.1.2.4 CpK

IV.1.2.4.1 TMS-propyne route

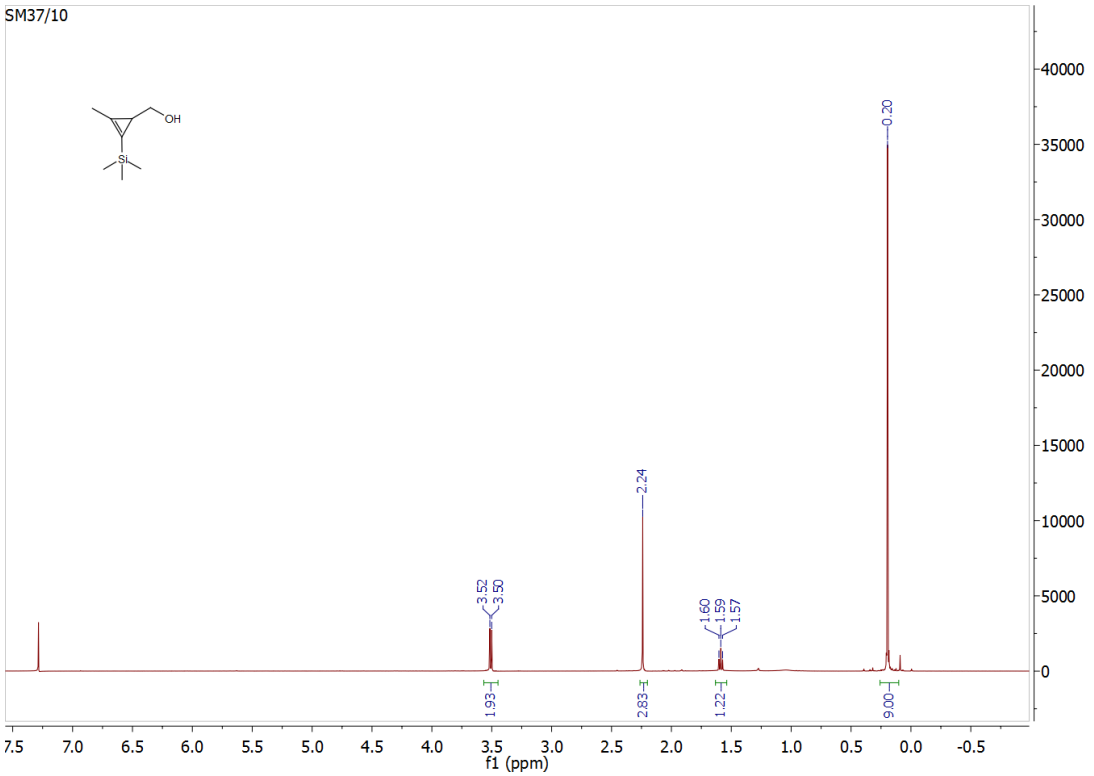
Ethyl 2-methylcycloprop-2-ene-1-carboxylate

^{1}H:



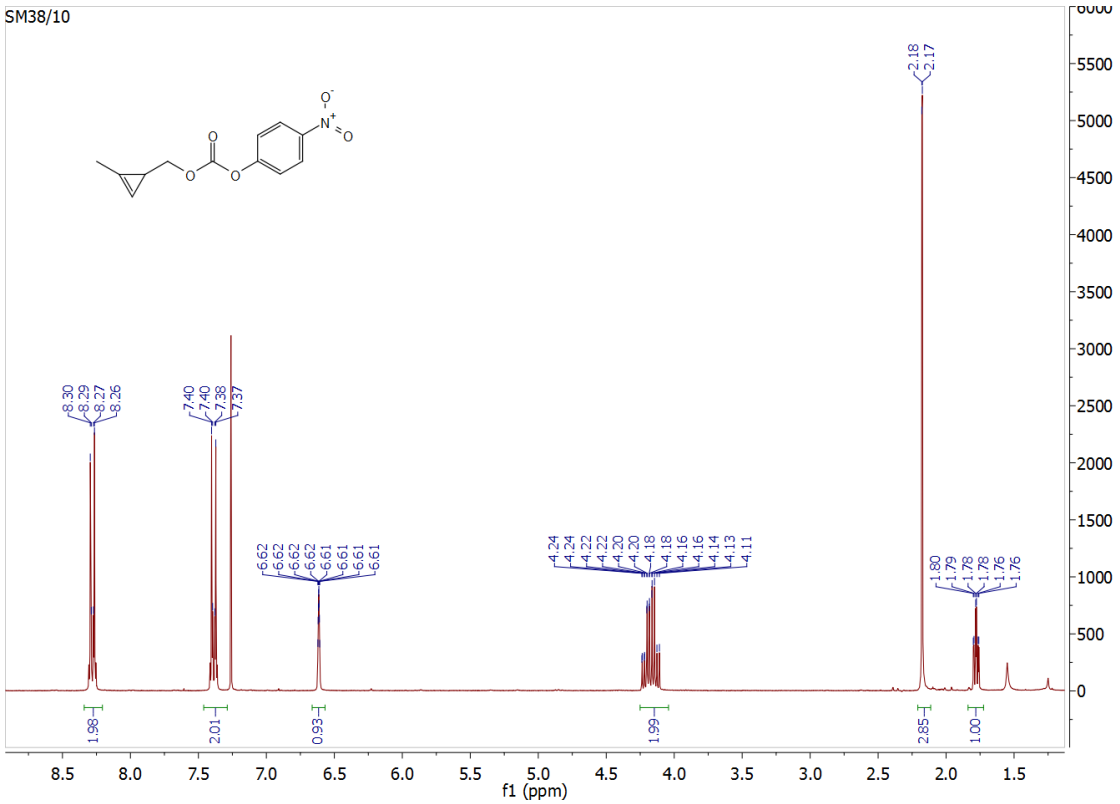
(2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methanol

^{1}H:



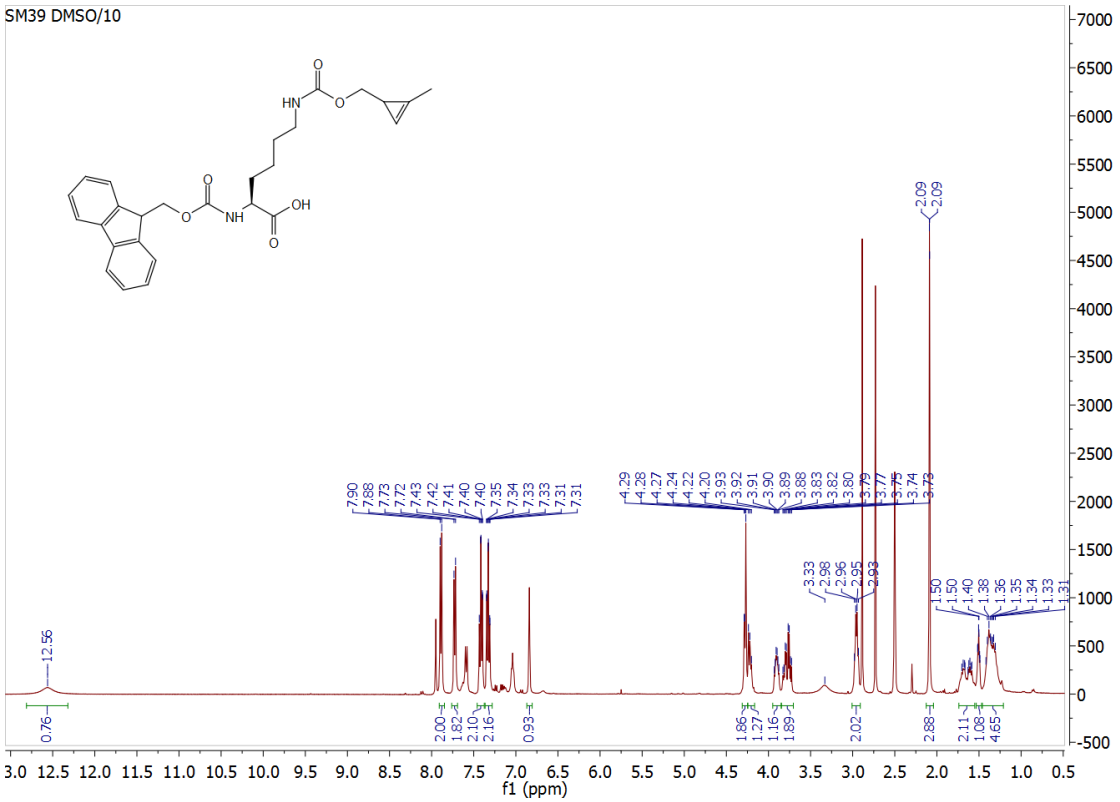
(2-methylcycloprop-2-en-1-yl)methyl (4-nitrophenyl) carbonate

^{1}H:



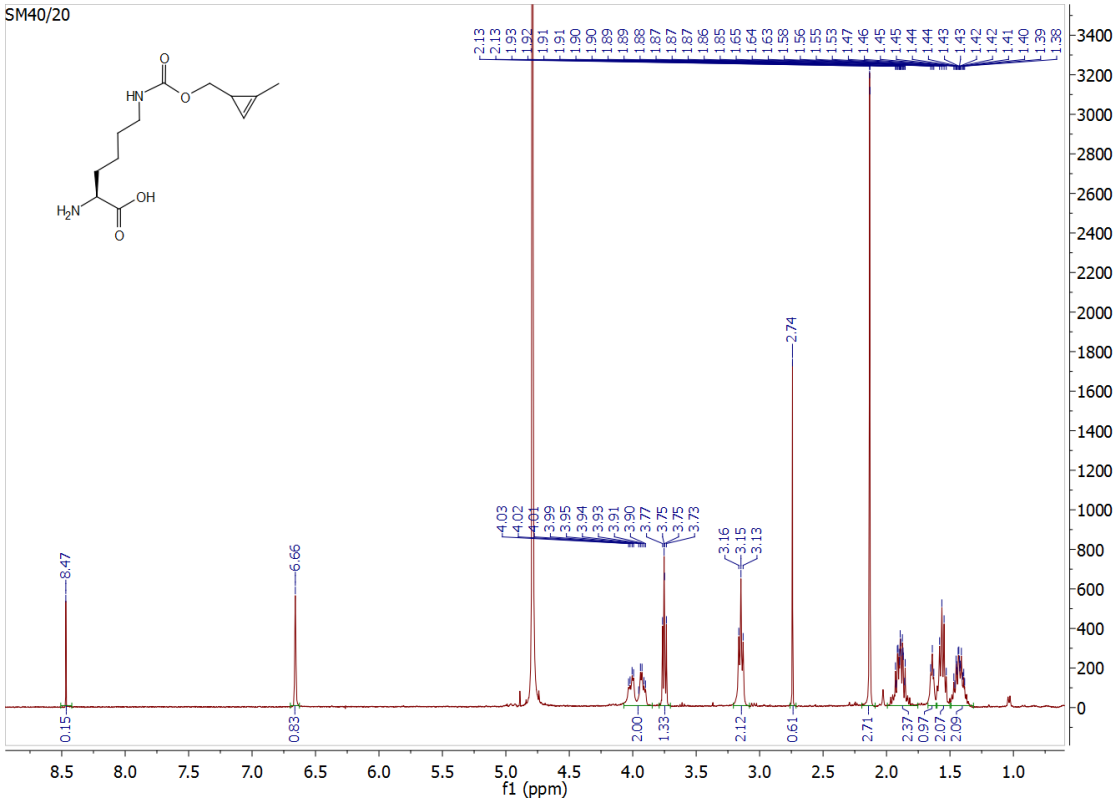
FmocCpK

^{1}H:



CpK

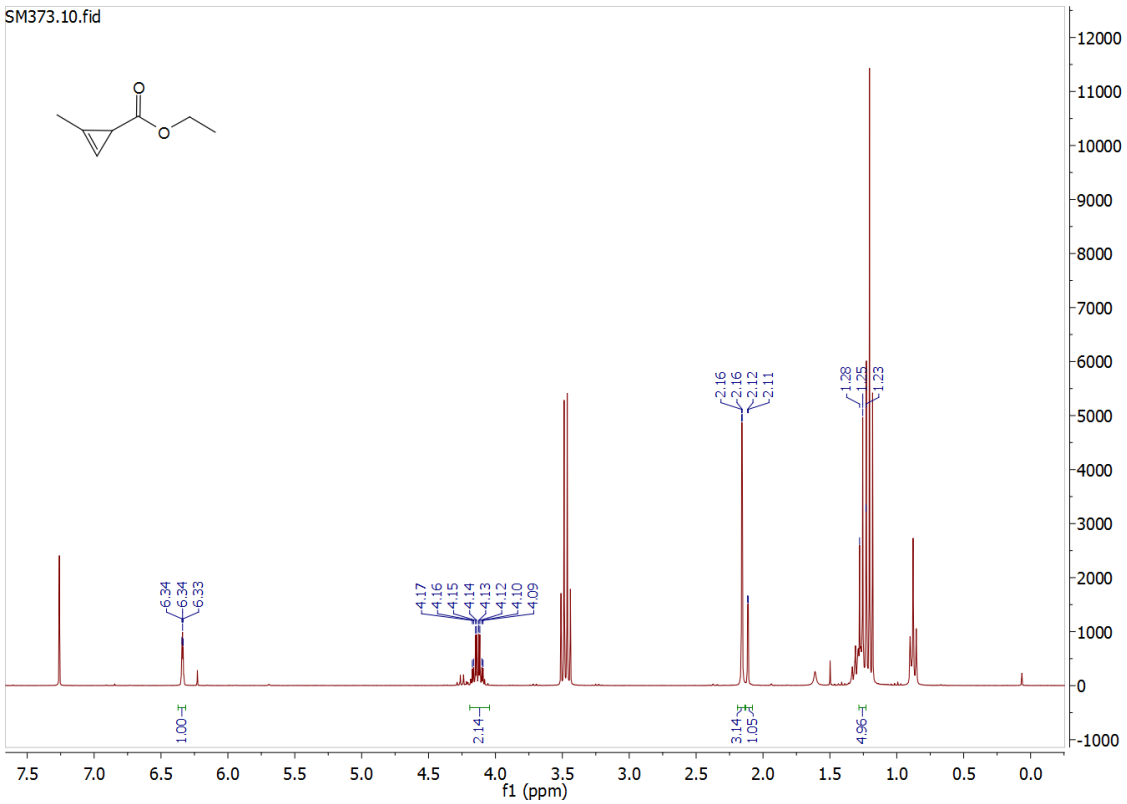
^{1}H:



IV.1.2.4.1 Propyne route

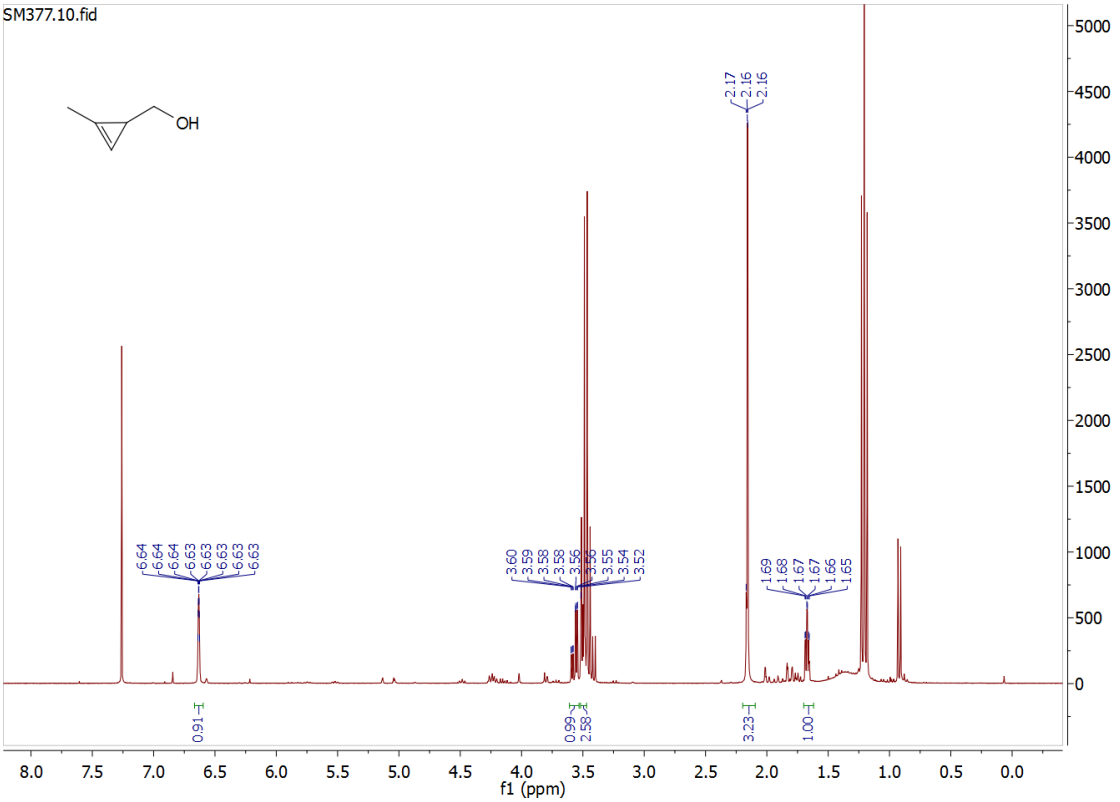
Ethyl 2-methyl-3-(trimethylsilyl)cycloprop-2-ene-1-carboxylate

^{1}H:



(2-methylcycloprop-2-en-1-yl)methanol

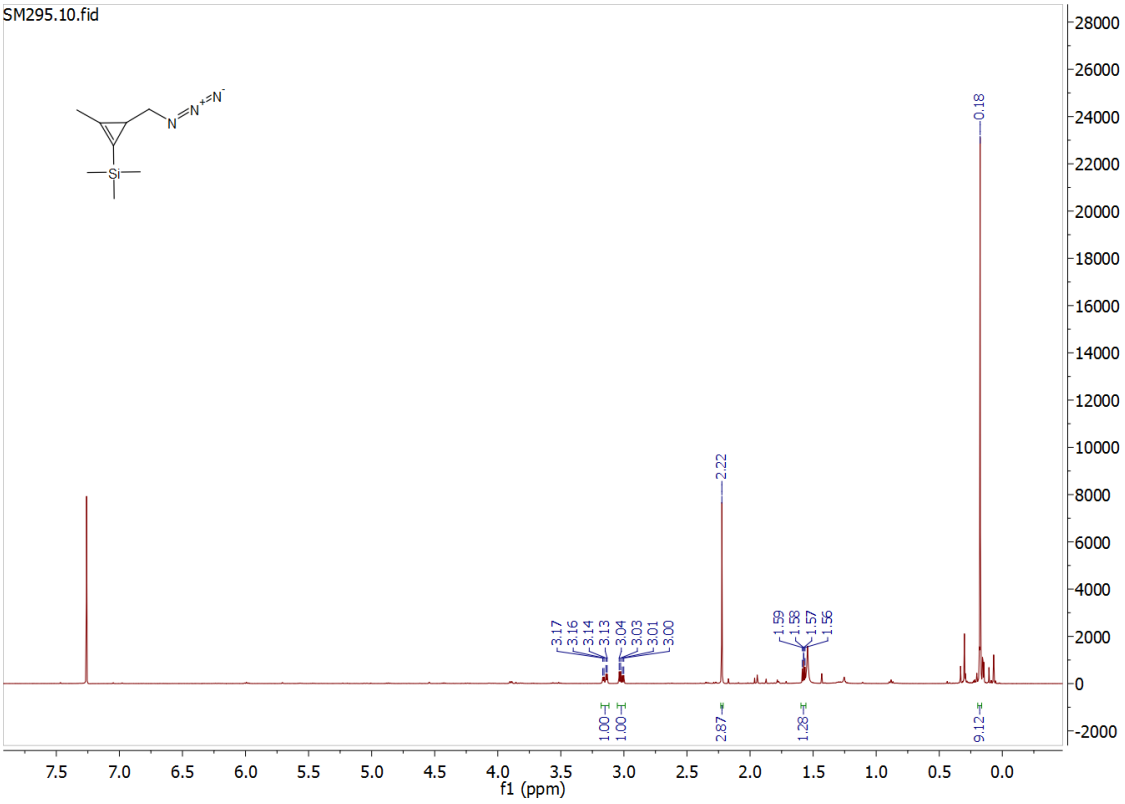
^{1}H:



IV.1.2.5 Cp-PROXYL

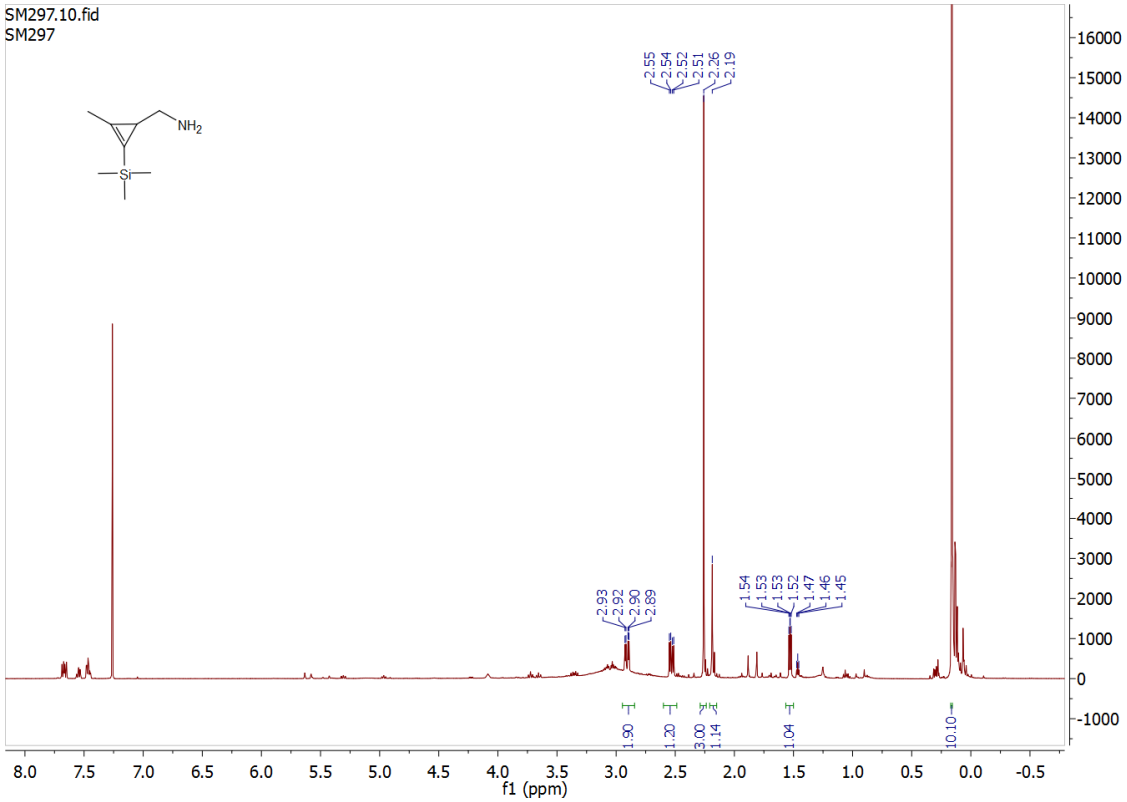
(3-(azidomethyl)-2-methyl-1-(trimethylsilyl))cycloprop-1-e

^{1}H:



(2-methyl-3-(trimethylsilyl)cycloprop-2-en-1-yl)methanamine

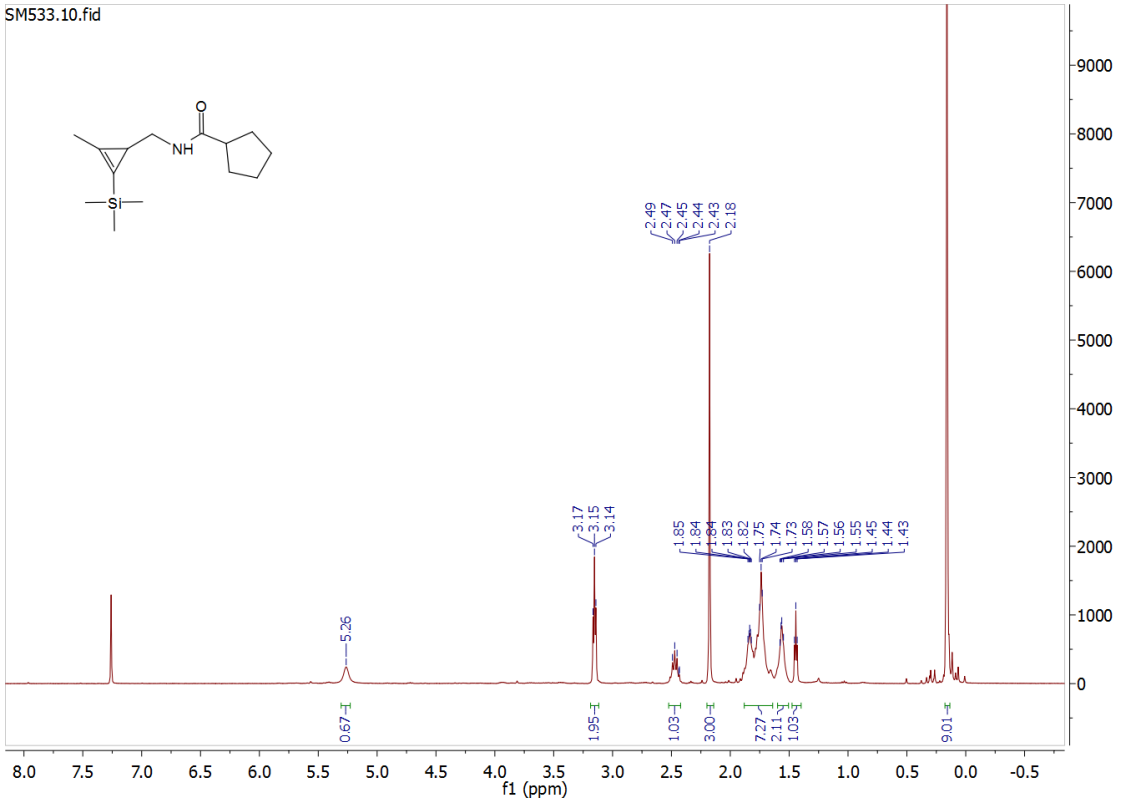
^{1}H:



IV.1.2.5 Cp-SLA

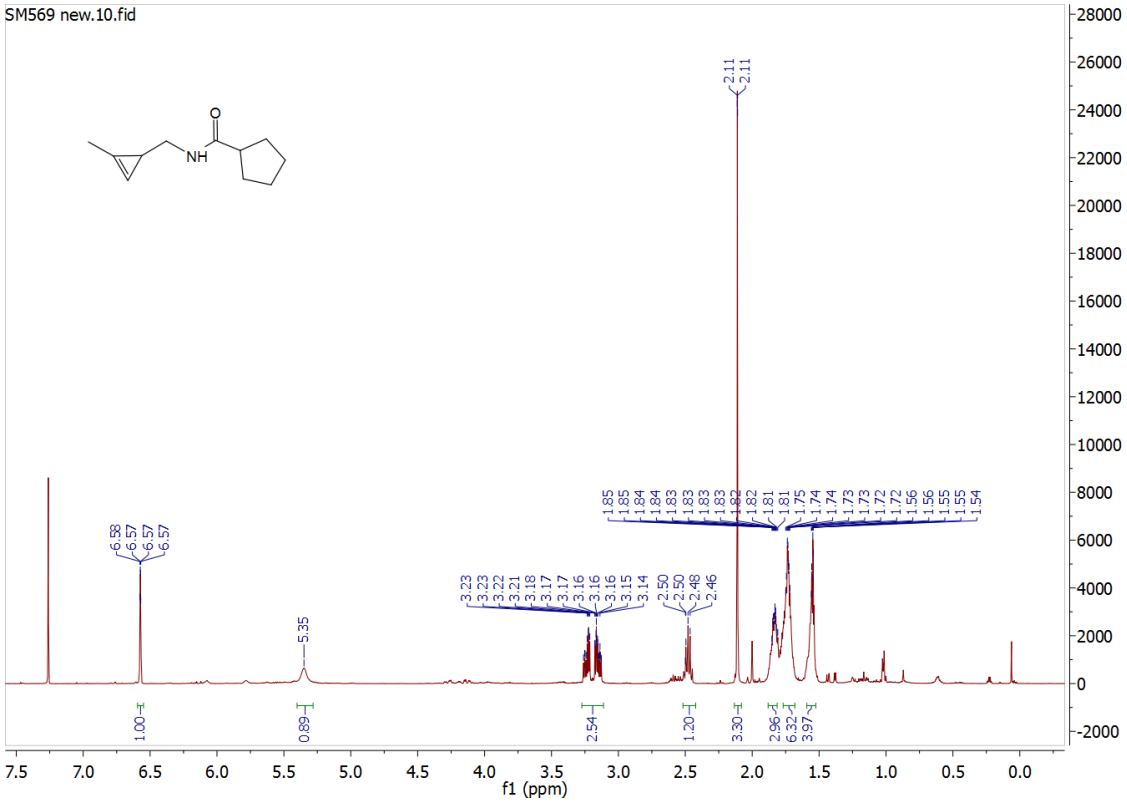
TMS-Cp-SLA

^{1}H:



Cp-SLA

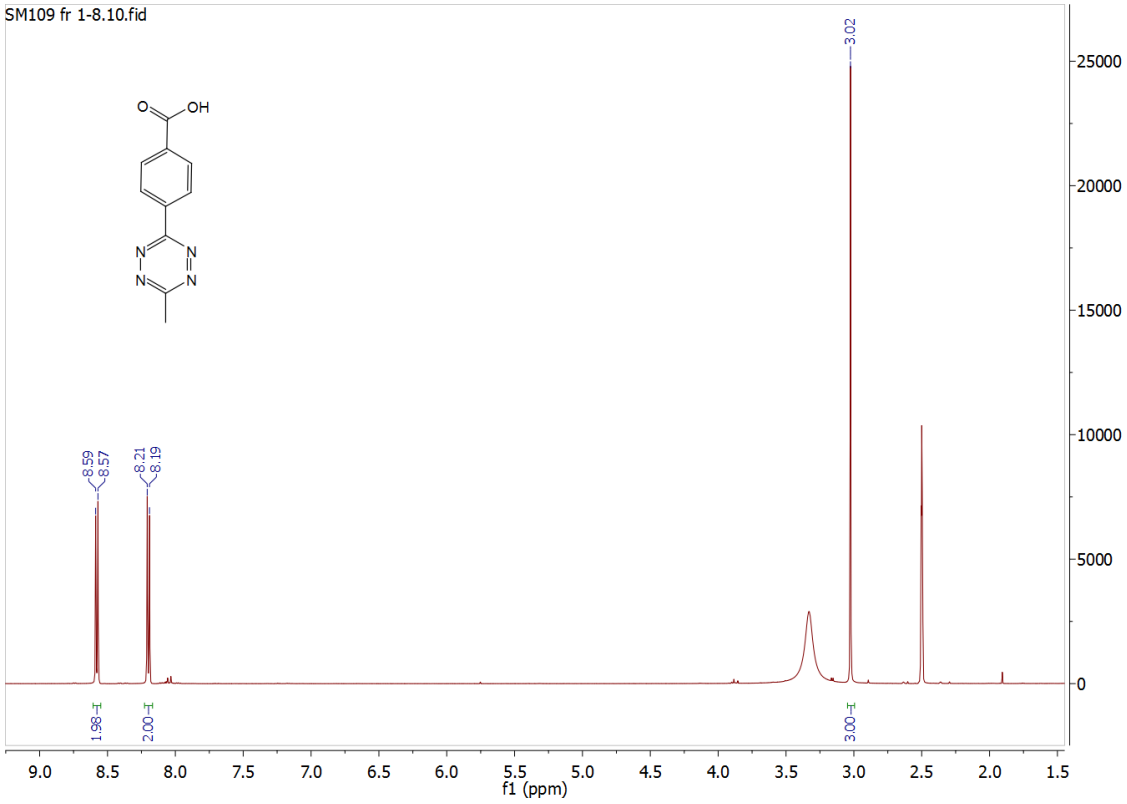
^{1}H:



IV.1.2.6 Tet-TEMPO

3-(4-Carboxyphenyl)-6-methyl-1,2,4,5-tetrazine

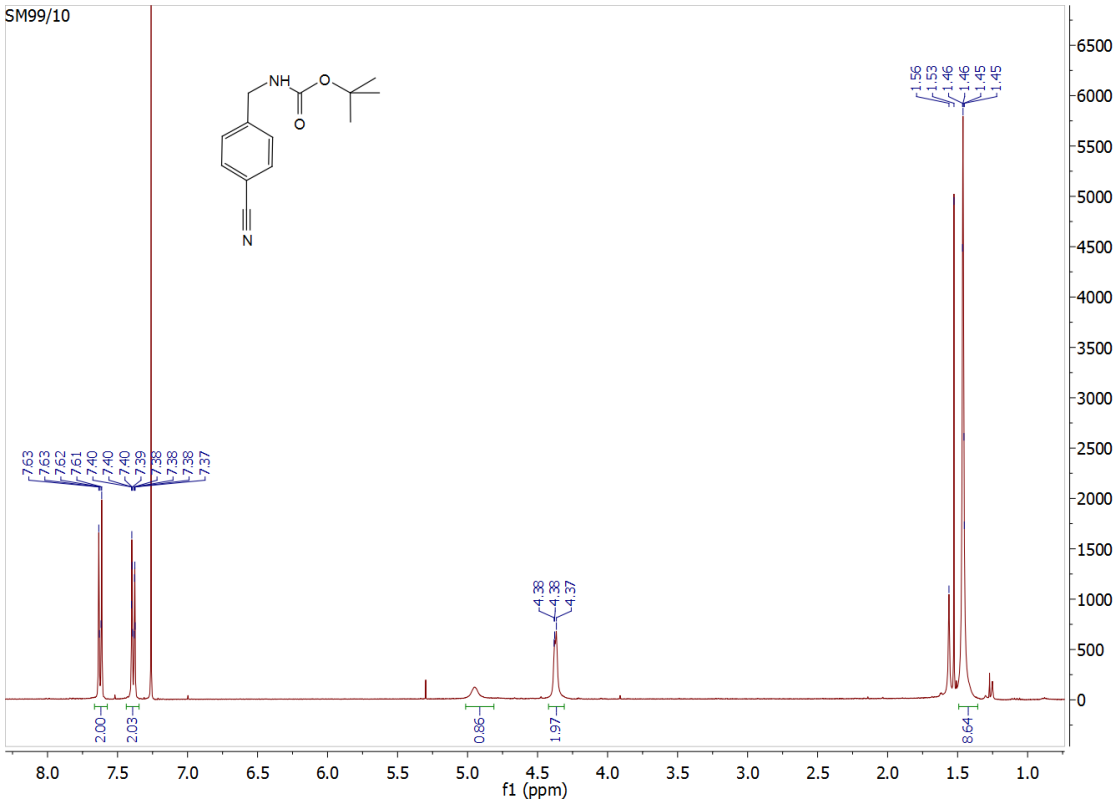
^{1}H:



IV.1.2.7 Tet-PROXYL

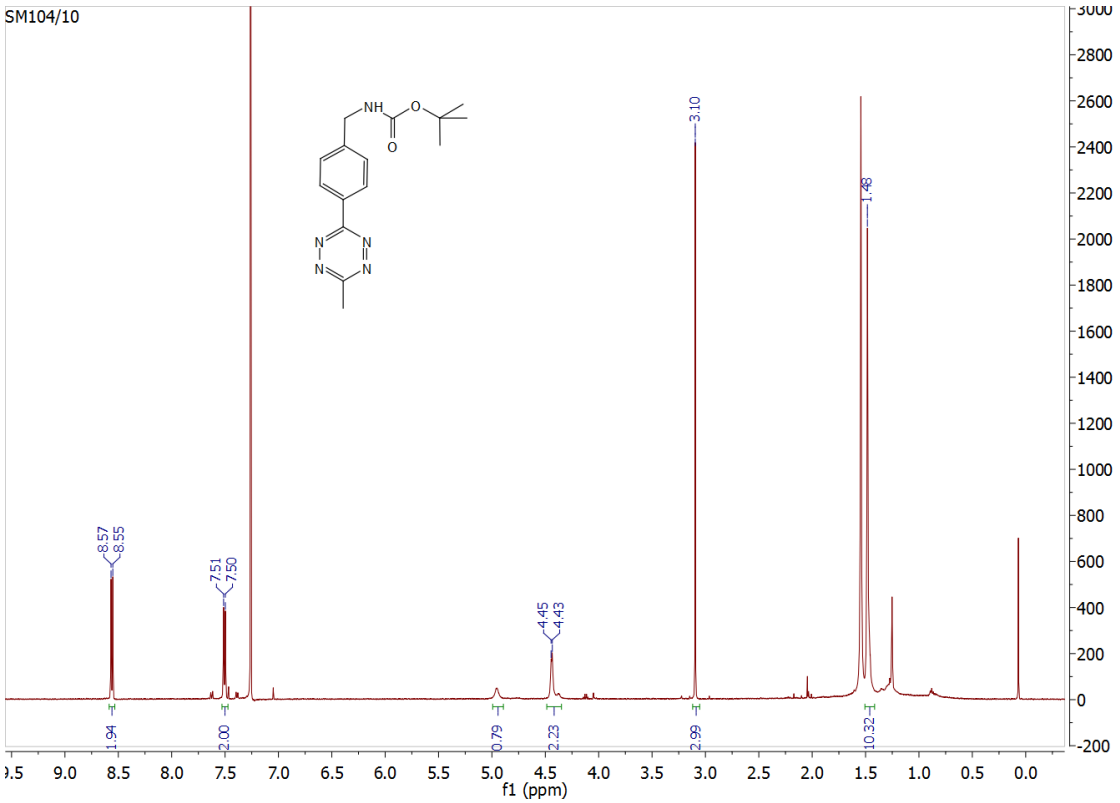
4-(*N*-(tert-butoxycarbonyl)aminomethyl)benzonitrile

^{1}H:



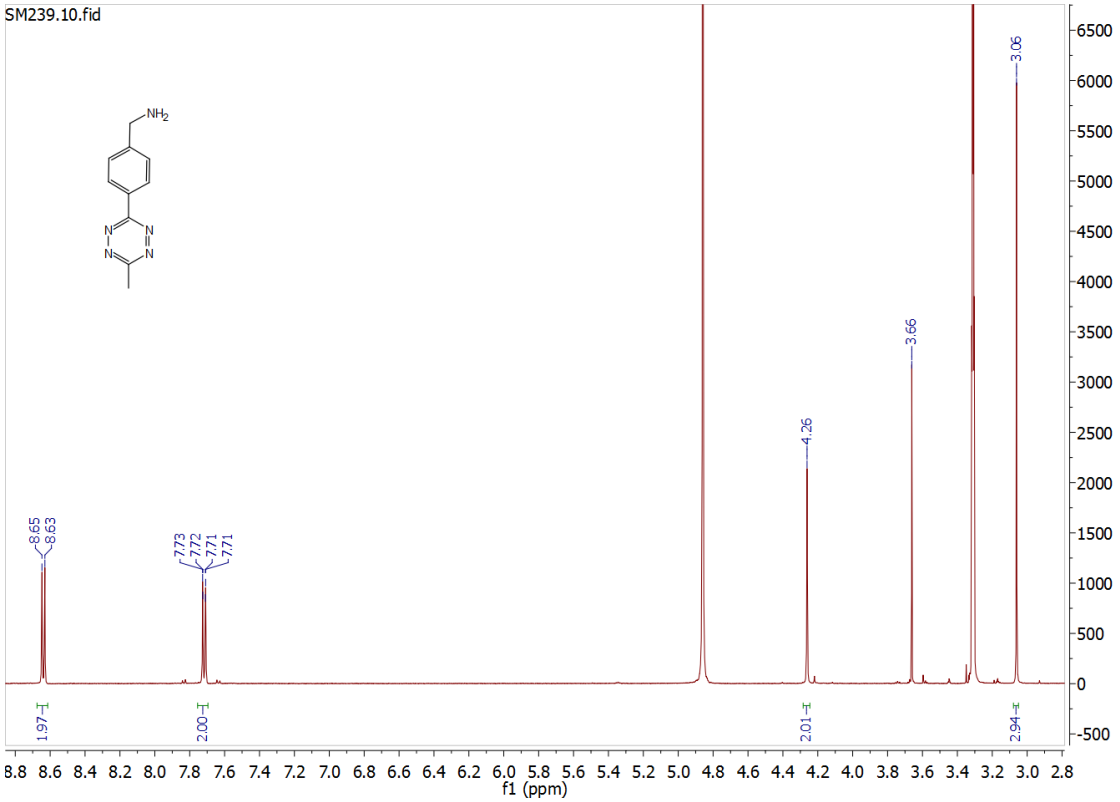
N-(*tert*-butoxycarbonyl)-3-(4-aminomethyl)phenyl-6-methyl-1,2,4,5-tetrazine

^{1}H:



3-(4-aminomethyl)phenyl-6-methyl-1,2,4,5-tetrazine hydrochloride

^{1}H:



V. Eidesstattliche Erklärung

Hiermit erkläre ich eidesstattlich, dass ich die vorliegende Arbeit selbstständig angefertigt und keine anderen als die angegebenen Quellen oder Hilfsmittel verwendet habe. Alle in dieser Arbeit sinngemäß oder wortwörtlich übernommenen Stellen habe ich gekennzeichnet. Diese Arbeit wurde für keinen anderen akademischen Grad eingereicht wie angegeben.

....

Ort, Datum Unterschrift